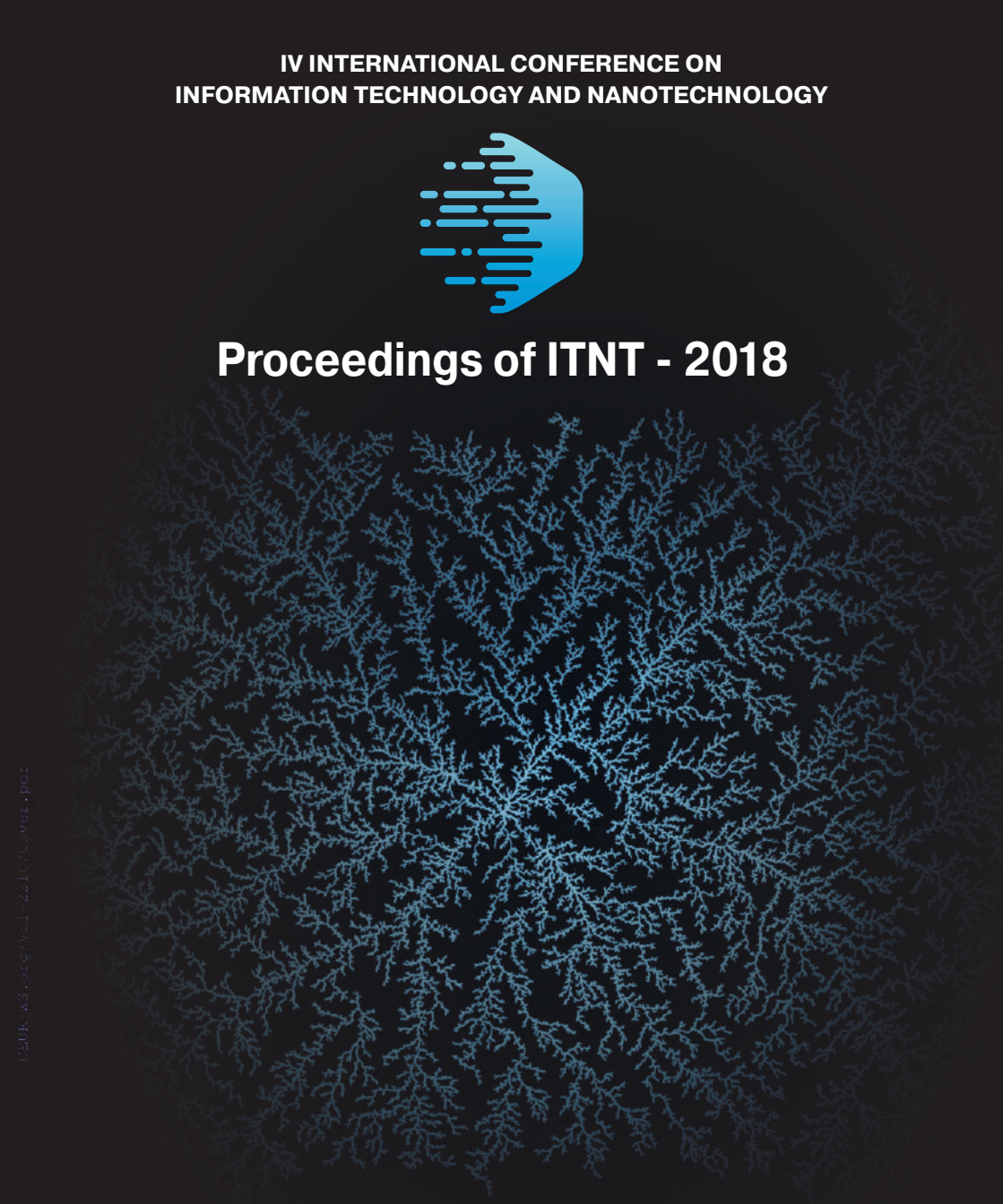
IV INTERNATIONAL CONFERENCE ON
INFORMATION TECHNOLOGY AND NANOTECHNOLOGY

Proceedings of ITNT - 2018



25th annyversary of the Image Processing Systems Institute of RAS - branch of the FSRC "Crystallography and Photonics" RAS

24-27 April, 2018
Samara (Russia)

Samara National Research University
Image Processing Systems Institute of the RAS - branch of the Federal Scientific Research Centre “Crystallography and Photonics” of Russian Academy of Sciences

INFORMATION TECHNOLOGY AND NANOTECHNOLOGY (ITNT-2018)

Section: Image Processing and Earth Remote Sensing
Section: Data Science

Collection of selected papers of the IV International Conference on Information Technology and Nanotechnology
(Samara, 2018, 24-27 April)

978-5-6047396-1-7
Publisher: IP Zaitsev V.D.

Samara
2018

Collection of selected papers of the IV International Conference on Information Technology and Nanotechnology (ITNT-2018). Section. Image Processing and Earth Remote Sensing. Section. Data Science. Samara, Individual Proprietor Zaitsev V.D. 2018, 24-27 April. ISBN 978-5-6047396-1-7.

Compilers of the volume: Vladislav Myasnikov, Vladislav Sergeyev, Vladimir Fursov, Michael Sobolewski
Issuing editor: Denis Kudryashov
Publisher: Individual Proprietor Zaitsev V.D.

The conference is a forum for leading researchers from all over the world aimed to discuss the latest advances in the basic and applied research in the field of Information Technology and Nanotechnology. It is also aimed to attract young people to advanced scientific research and share the latest trends in training and research programs for future ITNT specialists.

CONTENTS

Image Processing and Earth Remote Sensing

CONTENTS

Data Science

Preface

E S Kozlova[1,2]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086
[2]Image Processing Systems Institute of the RAS - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001

This volume contains the papers presented at session "Image Processing and Earth Remote Sensing" within the IV International Conference on Information Technology and Nanotechnology (ITNT-2018) which was held in Samara, Russia, April 24−27, 2018 (itnt-conf.org). The Conference is intended to provide a forum for leading scientists from all over the world to discuss the latest advances in the basic and applied research in the field of Information Technology and Nanotechnology, attract young people to advanced scientific research, and share the latest trends in training and research programs for future ITNT specialists [1]. In addition to the session mentioned above, ITNT-2018 also included the following Sections: "Computer Optics and Nanophotonics", "Mathematical Modeling of Physical-Technical Processes and Systems", "Data Science".

The scientific forum in Samara brought together more than 420 scientists from more than 45 cities in Russia, Germany, United Kingdom, Finland, Switzerland, China, India, Pakistan, Iran, Iraq, Belarus, Moldova, Poland and many others. 407 talks enabled discussion on a wide range of topics. The topics of the papers in this volume include digital image processing, visual recognition and retrieval, motion analysis, scene reconstruction, remote sensing image processing and analysis, multimedia protection and information hiding, and geoinformation technology.

This year we have received 112 submissions addressed to "Image Processing and Earth Remote Sensing". The Program Committee members and the reviewers carefully reviewed each submission. Based on the reviews, Program Committee accepted 88 papers for presentation at the conference and publication. Program Committee published some of the accepted papers in Journal of Physics: Conference Series, in the volume covering all topics of the Conference. The current volume contains 55 papers not included in the Journal of Physics: Conference Series volume.

I thank all the participants of ITNT-2018 for excellent presentations and discussions.

Conference Organizers

- Samara National Research University, Samara, Russia
- Image Processing Systems Institute of the RAS - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Samara, Russia

Program Committee Chair

- Viktor Soifer, Samara National Research University

Organizing Committee Chair

- Evgeniy Shakhmatov, Samara National Research University

Vice-chairs

- Vladimir Bogatyrev, Samara National Research University
- Nikolay Kazanskiy, Image Processing Systems Institute of the RAS
- Eduard Kolomiets, Samara National Research University
- Alexander Kupriyanov, Samara National Research University

Compilers of the volume

- Vladislav Myasnikov, Image Processing Systems Institute of the RAS

- Vladislav Sergeyev, Samara National Research University
- Victor Fedoseev, Samara National Research University

Issuing editor

- Denis Kudryashov, Samara National Research University

The conference was held with the financial support of the Russian Foundation for Basic Research and Federal Agency of Scientific Organizations.

Reference

[1] Savelyev D A 2017 *Computer Optics* **41(5)** 775-785 DOI: 10.18287/2412-6179-2017-41-5-775-785

Preface

E S Kozlova[1,2]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086
[2]Image Processing Systems Institute of the RAS - Branch of the Federal Scientific Research Centre “Crystallography and Photonics” of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001

This volume contains the papers presented at session"Data Science" within the IV International Conference on Information Technology and Nanotechnology (ITNT-2018) which was held in Samara, Russia, April 24−27, 2018 (itnt-conf.org). The Conference is intended to provide a forum for leading scientists from all over the world to discuss the latest advances in the basic and applied research in the field of Information Technology and Nanotechnology, attract young people to advanced scientific research, and share the latest trends in training and research programs for future ITNT specialists [1]. In addition to the session mentioned above, ITNT-2018 also included the following Sections: "Computer Optics and Nanophotonics", "Image Processing and Earth Remote Sensing","Mathematical Modeling of Physical-Technical Processes and Systems".

The scientific forum in Samara brought together more than 420 scientists from more than 45 cities in Russia, Germany, United Kingdom, Finland, Switzerland, China, India, Pakistan, Iran, Iraq, Belarus, Moldova, Poland and many others. 407 talks enabled discussion on a wide range of topics. The topics of the papers in this volume include data mining, machine learning, security, cryptography, high performance computing.

This year we have received 141 submissions addressed to "Data Science". The Program Committee members and the reviewers carefully reviewed each submission. Based on the reviews, Program Committee accepted 104 papers for presentation at the conference and publication. Program Committee published some of the accepted papers in Journal of Physics: Conference Series, in the volume covering all topics of the Conference. The current volume contains 59 papers not included in the Journal of Physics: Conference Series volume.

I thank allthe participants of ITNT-2018 for excellent presentations and discussions.

- Michael Sobolewski, Polish-Japanese Institute of IT, Warsaw, Poland
- Egor Goshin, Samara National Research University

Issuing editor

- Denis Kudryashov, Samara National Research University

The conference was held with the financial support of the Russian Foundation for Basic Research and Federal Agency of Scientific Organizations.

Reference

[1] Savelyev D A 2017 *Computer Optics* **41(5)** 775-785 DOI: 10.18287/2412-6179-2017-41-5-775-785

Real-time multispectral video panorama construction

I A Kudinov[1], O V Pavlov[1], I S Kholopov[1,2] and M Yu Khramov[1]

[1]Scientific and Design Center of Video and Computer technologies, Ryazan State Instrument-making Enterprise, Seminarskaya Street 32, Ryazan, Russia, 390005
[2]Ryazan State Radio Engineering University, Gagarina Street 59/1, Ryazan, Russia, 390005

Abstract. An algorithm for the video panorama construction from distributed multispectral cameras data is described. It is shown that operations of vision enhancement (modified Multiscale Retinex algorithm) and multispectral image fusion are implemented for two independently chosen regions of interest with a frame size of 1024x768 pixels and 30 fps using CUDA technology.

1. Introduction

The real-time automatic generation of high-resolution video panoramas from information of several cameras with partially overlapping fields of view (FoV) is one of the modern trends in the vision systems development. Generally, panorama navigation implies the presence of a user-controlled region of interest (RoI). This approach is an alternative for mechanical drive-based vision systems, as it ensures simultaneous operation of several users with an independent choice of a personal RoI without mechanically moving the camera system. Another advantage of distributed panoramic systems (DPS) is the integration of video cameras into the body / fuselage of the carrier object, what positively affects its aerodynamic properties.

2. Methods of panorama construction

There are two basic approaches for panorama stitching:

- an approach based on finding the correspondence of homogeneous pixel coordinates on the frames of cameras with i and j numbers by detecting and matching keypoints using their descriptors and estimating the homography matrix $\mathbf{H}_{ij}$ [1, 2];
- an approach based on finding the correspondence of homogeneous pixel coordinates by preliminarily calibrating the cameras of the DPS by the test object with an auxiliary wide-angle camera [3] if the fields of view of cameras have small angular sizes of the intersection zone or do not intersect at all.

The advantage of the first approach is operability even in the absence of a priori information about the mutual placement of DPS cameras; the advantage of the second approach is operability in difficult observation conditions and in low-contrast observable scenes.

3. Multispectral images superimposition

It is known [4] that one of the main approaches for increase of situational awareness in poor visibility conditions is the simultaneous use of different spectral ranges cameras: visible TV and infrared (IR) – near IR (NIR), short wave IR (SWIR), medium wave IR (MWIR) and long wave IR (LWIR). Panorama stitching from different spectral frames for each of the methods considered above is

hindered by the different physical nature of the images formed by TV and IR cameras: TV, NIR and SWIR cameras see the light reflected by the object in the wavelength ranges 0.38...0.76, 0.7...1.1 and 0.9...1.7 μm respectively whereas MWIR and LWIR cameras see only the thermal radiation of the object at wavelengths 3...5 and 8...12 μm respectively. Therefore, in order to construct a video panorama in the DPS with different spectral ranges cameras (depending on the selected method of the panorama stitching), either a solution of the problem of keypoint matching on TV and IR frames or the production of a universal (having high contrast in the all operating spectral ranges) pattern for camera calibration is necessary.

3.1. Multispectral images keypoints matching

The analysis of publications [5-9] allows distinguishing of four basic approaches for automatic superimposition of multispectral images:

- based on the transition to the decorrelated colour space and the use the SIFT method: this approach is applicable only for combining RGB frames of the visible range and NIR [5];
- based on the mutual morphological correlation of pre-segmented images [6];
- based on the results of the contour analysis [7],
- based on estimating the homography matrix by manually matches search [8, 9].

Restrictions on the use of these methods are situations where the image in one of the video channels (usually in TV) does not allow selecting of the important details of the scene for matches searching (for example, in near to zero illumination, as well as in dense smoke or fog).

The solution of the superimposition problem can also be achieved by mechanical alignment, which is to ensure the parallelism of the sighting lines and the same angular dimensions of the field of view of the each spectral range cameras, as well as their mutual placement minimizing parallax, but this approach is not applicable for the DPS.

3.2. Multispectral camera calibration with universal pattern

Photogrammetric calibration of DPS cameras is the most universal approach for combining images in the far zone, but requires the use of a test pattern that have high contrast simultaneously in several spectral ranges. Examples of such patterns with typical "chessboard" image are considered in [10]. The results of the calibration allow us to evaluate the matrix of internal and external parameters of the multispectral DPS cameras.

4. Algorithm of the video panorama construction according to information from multispectral cameras

As the images of DPS cameras are characterized by geometric distortions caused by different shooting angles, in order to minimize them it is expedient to construct a panoramic frame on a virtual surface of uniform curvature: a sphere or a cylinder of unit radius [11]. Implemented by the authors the algorithm of the moving along the spherical video panorama RoI filling [12] for the DPS with pre-calibrated multispectral cameras contains the following steps.

1. Initialization: calculating quaternions $\mathbf{q}_{uv0}$, that specify the initial angular direction to the pixels of RoI [12]. If it is necessary to dynamically change the field of view the RoI, the quaternion $\mathbf{q}_{uv0}$ is recalculated in the body of the main operation cycle.

Main operation cycle

2. Estimation of the current angular position of the DPS reference camera by pitch θ and roll φ (for example, according to the data from the inertial measurement unit) and corresponding rotation matrix:

$$\mathbf{R}_{\varphi\theta} = \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}. \quad (1)$$

3. Calculation of the rotation quaternion for a given angular position of the line of sight (center of RoI) $\mathbf{q}_{\text{vis}}$ and quaternions $\mathbf{q}_{uv}$, determining the current angular position of the RoI pixels (u, v):

$$\mathbf{q}_{uv} = \mathbf{q}_{\text{vis}}\mathbf{q}_{uv0}. \quad (2)$$

4. Filling RoI for each spectral range with pixels from cameras by re-projecting points from the surface of a unit virtual sphere to their matrices (with lens distortion compensation [13]).

5. The implementation of the blending procedure [1, 14] for the RoI of each spectral range.

6. Pixel-level image fusion [15] according to selected fusion algorithm.

Since for each pixel of the RoI the processing according to the above algorithm is homogeneous, this allows us to apply the procedure of parallelizing the computations, for example, using the resources of the GPU.

5. Description of the DPS layout

The layout of the DPS is a development of previous authors' work [12] and in addition to grayscale TV cameras contains a LWIR thermal camera with a field of view of 50°×40° (figure 1). The mutual angular position of the fields of view of the TV cameras and the thermal camera in the sector 200°×120° is shown in figure 2. To synchronize frames in time, external sync pulses are applied.

Figure 1. RPM layout cameras: LWIR (bottom center) and five TV.

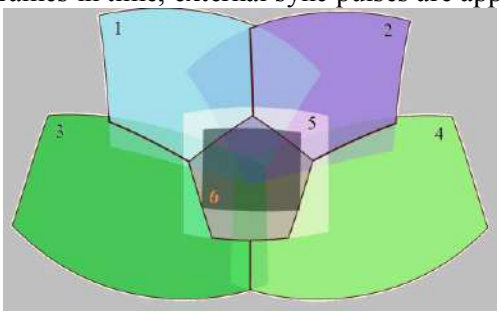


Figure 2. The intersection of the fields of view of the cameras from figure 1.

For filling of the RoI the computations are divided into parallel blocks (64 horizontally and 48 vertically) with 256 threads in each (16 threads horizontally and 16 vertically) using CUDA and CUDA C. As copying of the data of the CPU memory into the GPU memory and back is relatively slow, the number of such operations is minimized in our implementation of the video panorama.

In the DPS layout are implemented:

- angular position control of the line of sight of the operator: according to the data from the head tracking system or (if not) from the joystick;
- independent display of two RoIs with a dynamic change of its field of view from 80°×60° (wide angle) to 10°×7.5° (telephoto angle);
- blending according to the algorithm [14] (figure 3);
- increasing the contrast of the TV image (figure 4) using the modified Multiscale Retinex algorithm [16]: in order to accelerate calculations to estimate the brightness of the background, instead of smoothing by a Gaussian filter a box filter is used;
- RoI display mode selection: TV, grayscale thermal image, false-color thermal image (Jet [17] and Cold-to-Warm [18] color maps are realized), contrasted TV (Multiscale Retinex), image fusion mode – grayscale [19] or false-color [19-21]); results are shown in figures 5 and 6.
- mapping of the mutual angular position of the RoIs of the first and second operators.

As the DPS layout currently contains a single thermal camera, the fusion mode is realized only at angular positions of the RoI containing a part of area 6 (see figure 2). Otherwise, the user in this mode receives information only from the TV cameras. This is illustrated in figure 5, where in the thermal camera and fusion modes the lower rows of the RoI are filled with information from the TV cameras, because at the current position of the line of sight their angular coordinates do not fall in the field of view of the LWIR camera and therefore do not contain data in infrared spectral band.

On the NVIDIA GeForce GTX 560 Ti GPU (384 cores) with the maximum amount of calculations (blending, false-color fusion of contrasted TV and IR channels) and the RoI size 1024×768 pixels for

each of the two operators, the rate of updating the information is 32 Hz. The calculation speed for other modes is given in the table, where for comparison, information on the processing speed is also provided when implementing a video panorama on an Intel Core i5 CPU. All values in the table are rounded to the nearest whole number.

Table 1. The maximum rate of updating information in the RoIs for two users in different operating modes, Hz.

	Mode 1: TV without blending	Mode 2: TV with blending	Mode 3: Mode 2+ Multiscale Retinex	Mode 4: Mode 3 + false-color fusion
GPU	77	39	35	34
CPU	4	2	2	2

As can be seen from the table, the use of parallel computations increases the processing speed by an average of 16 times.

Figure 3. Blending results (observation conditions – cloud cover, 3x digital zoom): left – original frame in RoI, right – after blending.

Figure 4. Multiscale Retinex results (observation conditions – sunny day): left – original frame, right – after Multiscale Retinex.

Figure 5. Grayscale fusion (observation conditions – after rain, twilight): left – LWIR RoI layer, center – TV RoI layer, right – fusion frame.

6. Conclusion

With the use of CUDA technology the algorithm for creating a video panorama based on information from multispectral cameras for two windows of interest with a resolution of 0.7 Mp implements an

independent display of video information and vision enhancement functions (Multiscale Retinex and pixel-level fusion) with a frequency of at least 30 fps.

Figure 6. Fusion results for false-color technique (observation conditions –after rain, partly cloudy): top left – LWIR RoI layer, top right – TV RoI layer, bottom left – false-color fusion frame, bottom right – frame from external RGB camera for comparison.

7. References

[1] Brown M and Lowe D 2007 *Int. J. Comput. Vision* **74(1)** 59-73
[2] Fischler M and Bolles R 1981 *Commun. ACM* **24(6)** 381-395
[3] Shirokov R I and Alekhnovich V I 2014 *Contenant* **4** 10-23
[4] Knyaz V A, Vygolov O V, Vizilter Y V, Zheltov S Y and Vishnyakov B V 2016 *Proc. SPIE* **22** 984022
[5] Brown M and Susstrunk S 2011 *Proc. IEEE CVPR* (Washington DC: IEEE Comput. Soc.) 177-184
[6] Vizilter Y V, Zheltov S Y, Rubis A Y and Vygolov O V 2016 *J. Comput. Syst. Sci. Int.* **55** 598-608
[7] Efimov A I, Novikov A I and Sablina V A 2016 *Proc. 5th Mediterr. Conf. Embedded Comput.* (Bar) 132-137
[8] Bhosle U, Roy S D and Chaudhuri S 2005 *Pattern Recognit. Lett.* **26(4)** 471-482
[9] Efimov A I and Novikov A I 2016 *Computer Optics* **40(2)** 258-265 DOI: 10.18287/2412-6179-2016-40-2-258-265
[10] St-Laurent L, Mikhnevich M, Bubel A and Prévost D 2017 *Quant. Infrared Thermography J.* **14(2)** 193-205
[11] Szeliski R 2006 *Found. Trends Comput. Graphics Vision* **2(1)** 1-104
[12] Kudinov I A, Pavlov O V, Kholopov I S and Khramov M Yu *CEUR Workshop Proc.* **1902** 37-42
[13] Brown D C 1971 *Photogramm. Eng.* **37(8)** 855-866
[14] Burt P and Adelson E 1983 *ACM Trans. Graphics* **2(4)** 217-236
[15] Li S, Kang X, Fang L, Hu J and Yin H 2017 *Inf. Fusion* **33** 100-112
[16] Jobson D J, Rahman Z and Woodell G A 1997 *IEEE Trans. Image Proc.* **6(7)** 965-976
[17] *MATLAB Jet Array* (Access mode: https://www.mathworks.com/help/matlab/ref/jet.html)
[18] Moreland K 2009 *Proc. 5th Int. Symp. Adv. Visual Comput.* (Las Vegas) **II** 92-103
[19] Zheng Y 2011 *Image Fusion and its Applications* (Rijeka: inTech)
[20] Li G, Xu S and Zhao X 2010 *Proc. SPIE* **7710** 77100S
[21] Kholopov I S 2016 *Computer Optics* **40(2)** 266-274 DOI: 10.18287/2412-6179-2016-40-2-266-274

Classification of radar images with different methods of image preprocessing

A A Borodinov[1] and V V Myasnikov[1,2]

[1]Samara National Research University, Moskovskoye shosse34, Samara, Russia, 443086
[2]Image Processing Systems Institute - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001

Abstract. This work is aimed at comparing the classification algorithms and methods of machine learning with various methods of preliminary processing of radar images. Preprocessing step includes speckle noise filtering and object orientation normalization on the image. In comparison, the following classification algorithms were considered: Decision Tree, Support Vector Machine, K Nearest-Neighbor method, Random Forest, AdaBoost where decision tree was used as a weak classifier, Convolutional Neural Network and Residual Neural Network. The principal component analysis was applied to reduce the dimension. The study was carried out on the objects from the base of radar images MSTAR.

1. Introduction

Radar satellite imagery obtained with synthetic aperture radar received its distribution due to independence from weather and natural illumination. The recognition of objects on radar images is used in various fields, such as agriculture, forestry, geology, ecology, and also in the work of the emergency and rescue services. Synthetic aperture radar is usually an air or space radar system that uses the flight path of a platform to simulate an extremely large antenna, which makes it possible to create images of remote sensing of the earth with high resolution.

A certain complexity in the processing of the images obtained is the speckle noise that is present on the radar images. To improve the accuracy of the classification of images, the orientation of the object is also normalized. In most cases, the angle of rotation of the target is unknown. In this paper, we consider three methods commonly used in scientific articles to find the angle of rotation: a method that relies on the use of image moments; a method based on the Hough Transform, in which there is a search for long straight lines in the image, which are taken beyond the boundaries of the object; a method based on finding a rectangle with a minimum area describing the object.

The paper compares classical algorithms for the classification of radar images for various methods of their preliminary processing and methods based on the use of convolutional neural networks [8]. All the experiments were carried out using the generally available database of radar images of MSTAR military equipment [1].

The paper is organized as follows. In the second section, the statement of the classification problem is given. The third paragraph describes methods of preliminary image processing and compares the accuracy of finding the angle of rotation. In the fourth chapter, a method for reducing the

dimensionality of the characteristic space is described. In the fifth chapter, the classification algorithms under investigation are described. The sixth chapter presents the results of experimental studies. The conclusion contains conclusions and plans for further research. In the end of the work, thanks and a list of used literature are indicated.

2. Statement of the classification problem

The task of object recognition on an image can be divided into two main subtasks:

- search for an object in the image and selection of areas of interest;
- recognition and classification of the found object or area of interest. [1]

The first subtask is aimed at finding objects for classification. Often information about the location, size, orientation, availability, and a number of goals is initially missing. In this case, it is necessary to determine the unknown parameters required for further selection of the object or local area of interest.

The second subtask is applied to the entire image and allows you to decide which of the several classes the image being processed belongs to. The decision function for each feature vector relates the object to the corresponding class. In this paper, we consider only the classification problem.

Due to the need to process a large number of images for training and testing, as well as a low performance of some algorithms, there is a need to reduce the dimensionality of the feature space. There are various methods used to solve this problem. These methods include the most popular methods: the principal component analysis, factor analysis, the independent component analysis, Self-organizing map and others. In article [10] the authors use a convolutional neural network to reduce the dimension and compare the results with SVM [11]. In this paper, the principal component analysis is applied.

3. Pre-processing images

The preliminary processing of images consists of two stages. At the first stage, speckle noise is filtered by the Lee filter and threshold processing. At the second stage, one of the methods of normalizing the orientation of the object on the image is applied.

The method based on the moments of the image allows us to obtain the required rotation angle using the following formula, where

$\mu'_{11} = \mu_{11}/\mu_{00} = M_{11}/M_{00} - xy,$
$\mu'_{20} = \mu_{20}/\mu_{00} = M_{20}/M_{00} - x^2,$
$\mu'_{02} = \mu_{02}/\mu_{00} = M_{02}/M_{00} - y^2:$

$$\varphi = \frac{1}{2} arctg \frac{2\mu'_{11}}{\mu'_{20}+\mu'_{02}} + \nu\frac{\pi}{2} \ (\nu = 0,1,\dots) \quad (1)$$

In the method based on finding the minimum area of the rectangle describing the object in the image, the rotation angle is calculated as the angle between the horizontal and one of the sides of the rectangle. It is assumed that the rectangle will describe the object on the image in such a way that the longest side of the object will be described by the long side of the rectangle. To isolate the boundaries of an object, the Canny algorithm is used. The objects from the MSTAR database, described by the minimum area rectangle are shown in Figure 1.

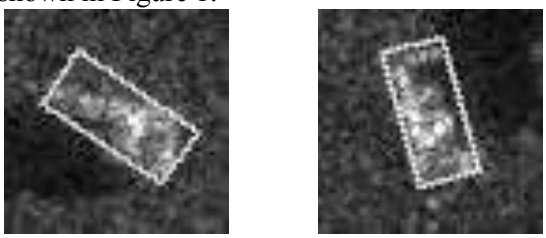


Figure 1. On the left is an example of an incorrect finding of a rectangle, the right selection of an object is shown on the right.

Figure 2 shows the scheme of the preprocessing algorithm.

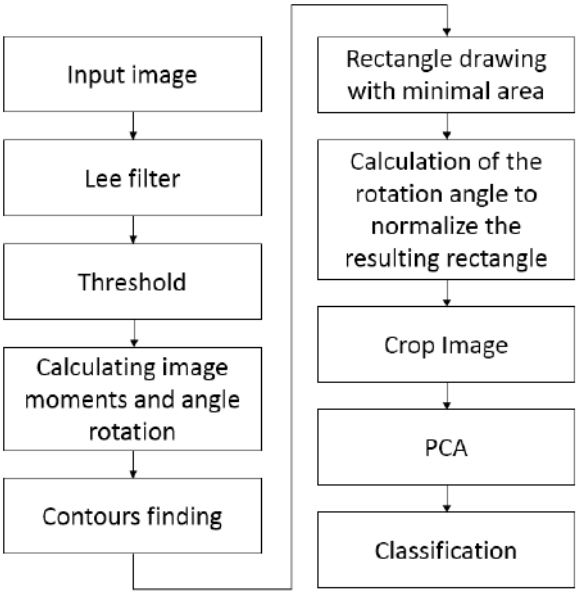


Figure 2. Preprocessing algorithm using a method based on finding the rectangle of the minimum area.

The method based on the Hough Transform involves finding long straight lines on the image that can be the boundaries of the object. The rotation angle is calculated for the line corresponding to the longest, straight and continuous face of the object. Like the previous one, this method is suitable for normalizing the orientation of an object with direct boundaries. To isolate boundaries, one can use the Canny algorithm or the Sobel operator. Figure 3 shows the scheme of the preliminary processing algorithm with the described method.

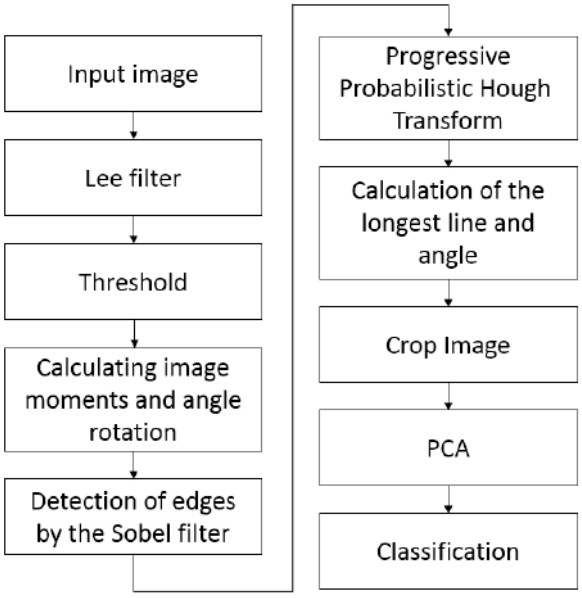


Figure 3. Preprocessing algorithm using the method based on finding the rectangle of the minimum area.

An example of this method is shown in Figure 4.

Table 1 shows the percentage of the number of images deviated from the true rotation angle by a specified angle to the total number of images.

The images obtained after normalization of the object are cut off.

Figure 4. The first two images show an incorrect finding of the face of the object, the third and fourth demonstrate the correct operation of the method.

Table 1. Comparison of the results of the methods used to normalize the orientation of the object.

Deviation from the true angle of rotation	Image moments, %	Rectangle with min. area, %	Hough Transform, %
>30	4.276	9.922	6.079
20 – 30	8.697	10.678	6.399
10 – 20	37.056	25.778	9.862
5 – 10	25.916	20.367	14.251
2 – 5	14.049	18.621	30.163
1 – 2	4.945	7.390	15.707
<1	5.061	7.244	17.539

4. Principal component analysis

Principal Component Analysis (PCA) is one of the most widely used methods for reducing the dimensionality of a feature space with the loss of the least amount of information. This method reduces to calculating the eigenvalues of the covariance matrix of the analyzed image. [2] Algorithms for calculating the covariance matrix operate in the line-by-line mode of reading the image, which allows achieving high performance and low requirements for the amount of RAM. [3]

5. Classification algorithms

5.1. K-Neighbors

K-nearest neighbors algorithm refers to metric classification algorithms with training sample Ω_O. Such algorithms refer object u to that class $y \in Y$, for which the total weight of the nearest objects from the training sample is maximal:

$$a(u,\Omega_O) = arg\max_{y\in Y}\Gamma_y(u,\Omega_O), \text{ и } \Gamma_y(u,\Omega_O) = \sum_{i=1}^{K}\left[y_u^{(i)} = y\right]\omega(i,u), \qquad (2)$$

where the weight function $\omega(i,u)$ estimates the degree of importance of the i-th neighbor for the classification of the object u. The function $\Gamma_y(u,\Omega_O)$is an estimate of the closeness of the object u to the class y. The importance function is chosen to be non-negative and not increasing in i. The selection criteria are due to the fact that the smaller the distance between the sampled objects u and $x_u^{(i)}$, the greater the probability of a correct classification. In the algorithm k of the nearest neighbors, the object u is referred to a class with more elements among the k nearest neighbors $x_u^{(i)}$, $i = \overline{1,k}$:

$$\omega(i,u) = [i \le k]\omega_i, a(u,\Omega_O,k) = arg\max_{y\in Y}\sum_{i=1}^{k}\left[y_u^{(i)} = y\right]\omega_i. \qquad (3)$$

The article used the Minkowski metric. It can be regarded as a generalization of the Euclidean and Manhattan distances. For the parameter $p = 2$, the Minkowski distance is generalized to the Euclidean distance, and for $p = \infty$- to the Chebyshev distance. This metric is defined by the following formula:

$$r(x,y) = \left(\sum_{i=1}^{n}|x_i - y_i|^p\right)^{1/p}. \qquad (4)$$

The drawbacks of metric algorithms include storage of the entire training sample.

5.2. Decision Tree (C4.5, CART)

Decision Tree is a structure of a hierarchical type, in which branches a partition of the feature space is defined, and the sheets are elementary classification functions. There are various methods for constructing trees. In this paper, the algorithms C4.5 [4] and CART [5] are considered.

C4.5 builds the source tree based on the input sample Ω_0. If all objects in the sample belong to the same class or the sample is small, then the tree is a sheet marked with the most common class in the sample. Otherwise, a split criterion is selected that divides the sample into two or more samples. Then the criterion is chosen for the obtained partitions. This procedure is recursively applied to each sample received. One of the criteria used, which is used in the work, is to minimize the entropy value of the obtained sample partitions. The resulting source tree is then trimmed to avoid retraining. Based on the received tree, a decision function is constructed for classifying objects.

In the CART algorithm, a binary decision tree is recursively constructed. The tree is created to the maximum size without using the stopping rule, and then it is clipped. The algorithm builds not one but a sequence of nested truncated trees. The best division is selected based on the sliding control. The partition criterion is based on the Gini index.

5.3. SVM

The support vector machine is one of the most reliable methods among all known algorithms and is most often used for comparison with new algorithms. The function separating the classes is a separating hyperplane. The algorithm maximizes the shortest distance between the points closest to the points on the hyperplane [6]. In this paper, the radial basis function is used as the separating function.

5.4. AdaBoost

A popular boosting algorithm AdaBoost was suggested by Freund and Schapire [7]. The main idea of the algorithm is to train a set of weak learners, such as small decision trees, on the prepared data. After this, the predictions from all of them are combined by weighted majority voting to obtain the final prediction. At the boosting iteration, the algorithm changes the weights for each sample from the training sample. Initially, the weighting factors are equal. For each subsequent iteration, the sample scales change individually. The weight of data that has been classified incorrectly is increasing. Thus, the algorithm concentrates training on the hard-to-class data.

5.5. Random forest

Random Forest is a set of decision trees. The decision to classify is made by voting the majority. Each decision tree is built independently of the others. For each tree, a subsample of the training sample is selected. For splitting the tree, the best attribute is selected. Typically, the tree is built before the exhaustion of the sample, the leaves of the tree should contain representatives of only one class.

5.6. Convolutional Neural Network

A convolutional neural network (CNN) consists of one or more convolutional layers and subsequent one or more fully connected layers. CNN is often used to recognize objects, images, as well as to detect and segment objects. The use of GPUs made it possible to significantly accelerate the training of the neural network and to reduce the time for conducting experimental studies. Classification algorithms based on convolutional neural networks make it possible to achieve better results in image recognition and stability to rotate and shift the object in the image.

5.7. Residual Neural Network

Residual networks allow the construction of much deeper networks than conventional convolutional networks. The article [8] describes the use of an ensemble of six residual neural networks, each of which had up to 152 layers.

The residual neural network consists of residual blocks, including a pair of layers, where the input of the first layer is added to the output of the second layer. Figure 5 shows the residual block.

6. Evaluation of classification results

To assess the results of the classification, the sliding control method is used. Cross-validation is a statistical method for assessing the generalization of the quality of classification. It is a more reliable and thorough assessment method, compared to the usual sequential division of a data set into a

training and test sample. With cross-validation, the data is repeatedly divided into training and test sets and fed to the classifier's input.

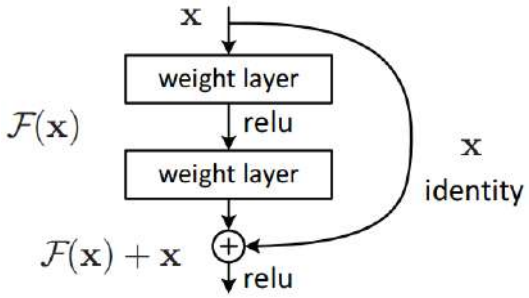


Figure 5. Residual block in the residual neural network.

The paper uses a modified method of sliding control with multiple partitioning [9] in which the entire volume of data is divided into a specified number of parts of N (equal to 1). The number of iterations of learning in this algorithm corresponds to the number of blocks N. There is also a stratification of classes and samples, allowing to reduce the dispersion of estimates of sliding control. This leads to a decrease in the confidence interval and a more accurate classification quality. Applying class stratification makes it possible to break each class in a given ratio. At each iteration of the algorithm, K parts are randomly selected as the training sample and L parts as the test sample. This partition can be described as follows:

$$\mathrm{N} \geq \mathrm{K} + \mathrm{L}, \Omega_{\mathrm{O}} \cup \Omega_{\mathrm{T}} = \Omega', \Omega' \subseteq \Omega, \Omega_{\mathrm{O}} \cap \Omega_{\mathrm{T}} = \emptyset, \Omega_{\mathrm{O}} = \bigcup_{i=0}^{\mathrm{K}-1} \Omega_{\mathrm{O}i}, \Omega_{\mathrm{T}} = \bigcup_{j=0}^{\mathrm{L}-1} \Omega_{\mathrm{T}j} \quad (5)$$

Where Ω_{O} is a training sample, Ω_{T} is a test sample, Ω is the original sample.

For each partition obtained, the classifier is set up on the training sample and the quality value of the classifier is calculated on the test sample.

The functional quality of the algorithm on the sample has the following form:

$$\mathrm{CV}(\mu, \Omega') = \frac{1}{\mathrm{N}} \sum_{p=1}^{\mathrm{N}} \frac{1}{\mathrm{N}} \sum_{q=1}^{\mathrm{N}} \mathrm{Q}\left(\mu\left(\Omega' \backslash \Omega_{\mathrm{O}pq}\right), \Omega_{\mathrm{O}pq}\right), \quad (6)$$

where μ is the learning method.

7. Experiments

All the experimental studies were conducted on Intel Core i5-6600, 16 GB of RAM, NVIDIA GeForce GTX 970. All classification algorithms were written in Python 3.6. Also used were frameworks and libraries scikit-learn, scikit-image, openCV, numpy, tflearn, tensorflow. As objects of classification, samples of military equipment from the public database of radar images MSTAR.

For recognition, the magnitudes of the images of BMP-2, BTR-60, BTR-70 and T-72 were used. Prior to the classification, preliminary processing and cropping of images to a size of 60×60 pixels described in paragraph 3 was carried out. Classification by neural networks was carried out without preliminary image processing.

The target shooting angle is 15 and 17 degrees. The initial sample consists of 3438 images. For all images from the general sample, the dimension was reduced.

The value of the classification quality will be calculated as the average relative number of correctly classified objects from the test sample Ω_{T}. For the sliding control method, we specify the number of partitions and the number of iterations $\mathrm{N} = 10$, $\mathrm{K} = 6$ and $\mathrm{L} = 4$, dividing the total sample in the ratio 6:4. The method of the main components will reduce the dimensionality, retaining a significant part of the radar image information necessary for the classification of objects. The results of the classification, together with the results obtained in [12] and [13] are shown in Table 2.

The algorithms AdaBoost and Random forest significantly outperformed the results of CART and C4.5. Perhaps a further selection of parameters will help to achieve values that showed the method of nearest neighbors and the method of support vectors. Although the method based on Hough Transform showed the smallest average deviation of the obtained angle from a known value, the accuracy classification with this pre-treatment method gave way to the classification accuracy using the other two methods because of the high mean square deviation.

Table 2. Results of classification.

Classifier	Image moments	Rectangle with min. area	Hough Transform	Known rotation angle	Without preprocessing
KNeighbors	0.93837	0.91904	0.89862	0.97689	-
CART	0.72500	0.74856	0.70908	0.82137	-
C4.5	0.75218	0.75196	0.71672	0.84055	-
SVM	0.93263	0.93299	0.90887	0.97071	-
AdaBoost	0.90029	0.87232	0.87769	0.94365	-
Random forest	0.90320	0.89113	0.86926	0.94851	-
CNN	-	-	-	-	0.97516
ResNet	-	-	-	-	0.96535

Classification algorithms based on the use of convolutional neural networks have surpassed the other classification algorithms studied in the case of methods of object orientation normalization. The method of nearest neighbors and the method of support vectors for a known angle of rotation of the target show a similar classification accuracy. Due to the fact that in practical classification tasks the angle of object rotation is rarely known, the application of classification algorithms based on convolutional neural networks will allow achieving better results even without preliminary image processing.

8. Conclusion

In this paper, we compare the classification algorithms in the problem of object recognition on radar images. Various methods of preliminary image processing and object orientation normalization were presented and studied. The experimental studies carried out have shown the dependence of the classification accuracy on the value of the root-mean-square deviation of the image rotation angle. It is shown that the application of classification algorithms based on convolutional neural networks allows obtaining the best classification accuracy even without preliminary image processing.

A further area of research is the development of a classification algorithm based on a convolutional neural network and the selection of the best hyperparameters, which allows to exceed the obtained classification accuracy.

9. References

[1] Ross T, Worrel S, Velten V, Mossing J and Bryant M 1991 Standard SAR ATR evaluation experiments using the MSTAR public release data set *Proc. SPIE - The International Society for Optical Engineering* **3370** 566-573
[2] Turk M 1998 Pentland, Eigenfaces for recognition *J. Cognit. Neurosci* **3(1)** 71-86
[3] Kuznetsov A V and V V Myasnikov 2014 A comparison of algorithms for supervised classification using hyperspectral data *Computer Optics* **38(3)** 494-502
[4] Salzberg S L 1994 C4.5: Programs for Machine Learning *Machine Learning* **16(3)** 235-240
[5] Wu X 2008 Top 10 algorithms in data mining *Knowledge and Information Systems* **14(1)** 1-37
[6] Cortes C and Vapnik V 1995 Support-Vector Networks *Machine Learning* **20(3)** 273-297
[7] Freund Y and Schapire R E 1997 A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting *Journal of Computer and System Sciences* **55(1)** 119-139
[8] He K, Zhang X, Ren S and Sun J 2016 Deep residual learning for image recognition *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* **2016** 770-778
[9] Vorontsov K V 2004 Combinatorial substantiation of learning algorithms *Computational Mathematics and Mathematical Physics* **44(11)** 1997-2009
[10] Zherdev D A, Minaev E Y, Procudin V V and Fursov V A 2017 Object recognition using real and modelled SAR images *Procedia Engineering* **201** 503-510
[11] Zherdev D A, Kazanskiy N L and Fursov V A 2014 Object recognition by the radar signatures

of electromagnetic field scattering on base of support subspaces method *Computer Optics* **38(3)** 503-510
[12] Borodinov A A and Myasnikov V V 2018 Comparison of classification algorithms for various methods of preprocessing radar images of the MSTAR base *Proceedings of SPIE* **10696** 1069614-7
[13] Borodinov A A and Myasnikov V V 2017 Comparison of classification algorithms in the task of object recognition on radar images of the MSTAR base *CEUR Workshop Proceedings* **1901** 37-41

Acknowledgments
This work was supported by the Russian Foundation for Basic Research (RFBR) grant 17-29-03190, 18-01-00748 A and by the Federal Agency of scientific organization (Agreement 007-GZ/43363/26).

Correlative features for the classification of textural images

M A Turkova[1] and A V Gaidel[1,2]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086
[2]Image Processing Systems Institute - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001

Abstract. The work is aimed to study various correlative image features in order to solve the classification problem of textural images, with further selection of the most informative features for increasing the reliability of classification. In the present paper series of experiments are performed on a set of 4480 real digital images from the open Kylberg Texture Dataset v. 1.0. A set of programs necessary for computations was developed. The autocorrelative image features are calculated, and the approximating coefficients from the regression equation are obtained. Due to the removal of the least informative features, the share of correctly recognized images increased in 1.04 times and amounted to 90.11% for 28 image classes.

1. Introduction

At present, the task of computer vision and, consequently image recognition, cause a particular interest in many areas of science and technology. Identifying a number of characteristic features which help to classify a texture is essential for the textural images recognition. A variety of parameters can act as the image features. Optimal and universal descriptors development continues. Among the most common classification signs for textural images are morphological features, fractal dimension, Markov's parameters, samples of the autocorrelation function and others [1].

At the beginning of the classification process each image can be characterized by a large number of parameters. Since the classifier for solving the classification problem is not ideal, the error it introduces essentially depends on the quality of the feature space. Thus, for proper classification, it is necessary to reduce the subset of features by removing uninformative and redundant features, thereby reducing the dimensionality of the feature set. Usage of the optimal feature subset will increase the reliability, at the same time reducing the computational complexity and time for classification.

Despite the urgency of the data processing problem, an optimal method for identifying and selecting features has not been found yet. In paper [2] a novel feature selection strategy resistant to problems with asymmetric data and based on Genetic Programming is proposed. It works well with both balanced and unbalanced data. The results of the experiments made possible reducing the size of the feature space up to 83%, which allow one to increase the effectiveness of the learning process. More details about the process of selecting features and using the samples of the correlation function as characteristics are described in [3].

A two-stage feature selection approach which makes full use of interactions is proposed in [4]. On the first stage, the problem of feature selection is decomposed into the sum of information interaction. Then, higher-order interactions are used to select interaction-preserving feature subset. On the second stage, the design of experiments is employed to identify significant interactions from the subset of

features. A flexible algorithm for selecting a subset of features under user control based on clustering was developed in work [5]. This is a two-step hybrid process which reduces the computational complexity of the feature selection, especially for large sets of features.

In [6] the feature selection is carried out in two stages: in the first step, the correspondence of each objective function feature is evaluated. If the attribute is irrelevant, then it is discarded. In the second step, the number of features is minimized by removing duplicate or low-performance functions. In [7] the process of selecting features is carried out using a heuristic evolutionary algorithm, which is the example of competition in a tribe. The main idea of the method is that only the most significant signs win and participate in the classification.

In this paper, there is a data set consisting of different classes of textural images. A number of features is formed for each image, based on the readings of the correlation function and the coefficients from the regression equations. A filtering method with a marked set of training data is used for the feature selection. A measure of the Euclidean distance is chosen as a search strategy.

2. Texture descriptors

2.1. Autocorrelation function as a characteristic of a texture image

To form a set of texture image features we use the often applicable texture descriptor – the autocorrelation function, as in [8]:

$$\rho(x,y) = \frac{1}{(N_i - |x|)(N_j - |y|)} \sum_i \sum_j I(i,j) I(i+x, j+y) \left(\frac{1}{N_i N_j} \sum_{i=1}^{Ni} \sum_{j=1}^{Nj} I(i,j)^2\right)^{-1} \quad (1)$$

where $I(i,j)$ – pixel value, $N_i \times N_j$ – size of image, $\rho(0,0) \equiv 1$ [1]. It was shown in [9] that linear combinations of autocorrelation function counts are the best signs in the class of quadratic features.

The coefficients of the autocorrelation function determine the relationship of the pixels in the image. As the distance between points increases, the correlation weakens and the value of the coefficients does not make a significant contribution to the calculations. Before the study begins, it is necessary to normalize the autocorrelation function to get rid of the high-frequency noise influence. Thus, we obtain an expression for the coefficients of the autocorrelation function.

2.2. Regression equation coefficients as a characteristic of the texture image

The autocorrelation function $\rho(x,y)$ obtained in the previous section can be unnecessarily complicated for describing the texture. However, if the existing function is approximated by a parametric function, then its parameters can be used as a more reliable texture descriptor.

Suppose that the autocorrelation function at a point (x_i, y_i) is known and is equal to ρ_i. Let us approximate a function by the following regression equations:

$$\rho_1(x,y) = \exp(a_1 x^2 + a_2 y^2 + a_3), \quad (2)$$

$$\rho_2(x,y) = \exp\left(b_1 \sqrt{x^2 + y^2} + b_2\right), \quad (3)$$

where a_1, a_2, a_3, b_1, b_2 – the coefficients to be found. We take the logarithm of (2) and (3) and obtain:

$$\ln\left(\rho_1(x,y)\right) = a_1 x^2 + a_2 y^2 + a_3, \quad (4)$$

$$\ln(\rho_2(x,y)) = b_1 \sqrt{x^2 + y^2} + b_2. \quad (5)$$

The total error for (4) and (5), respectively, is:

$$E_1 = \sum_i (\ln\left(\rho_1(x_i, y_i)\right) - \rho_i)^2 = \sum_i (a_1 x_i^2 + a_2 y_i^2 + a_3 - \rho_i)^2,$$

$$E_2 = \sum_i (\ln\left(\rho_1(x_i, y_i)\right) - \rho_i)^2 = \sum_i (b_1 \sqrt{x_i^2 + y_i^2} + b_2 - \rho_i)^2.$$

It is necessary to select the parameters a_1, a_2, a_3, b_1, b_2 in such a way for the error to be minimal. To do this we take the partial derivative with respect to each parameter and equate it to zero, obtaining a system of linear equations with respect to unknown parameters:

$$\begin{cases} \dfrac{\partial E_1}{\partial a_1} = 0 \Rightarrow 2\sum_i \left(a_1 x_i^2 + a_2 y_i^2 + a_3 - \rho_i\right) x_i^2 = 0, \\ \dfrac{\partial E_1}{\partial a_2} = 0 \Rightarrow 2\sum_i \left(a_1 x_i^2 + a_2 y_i^2 + a_3 - \rho_i\right) y_i^2 = 0, \\ \dfrac{\partial E_1}{\partial a_3} = 0 \Rightarrow 2\sum_i \left(a_1 x_i^2 + a_2 y_i^2 + a_3 - \rho_i\right) = 0. \end{cases}$$

Solving the resulting system of linear equations:

$$\begin{cases} a_1 \sum_i x_i^4 + a_2 \sum_i x_i^2 y_i^2 + a_3 \sum_i x_i^2 = \sum_i \rho_i x_i^2, \\ a_1 \sum_i x_i^2 y_i^2 + a_2 \sum_i y_i^4 + a_3 \sum_i y_i^2 = \sum_i \rho_i y_i^2, \\ \sum_i x_i^2 + a_2 \sum_i y_i^2 + a_3 \sum_i 1 = \sum_i \rho_i, \end{cases}$$

we obtain expressions for the coefficients a_1, a_2, a_3 that are used in the regression equation (2). Similarly, we obtain values for the coefficients b_1, b_2 from equation (3) [1].

3. Experimental studies

3.1. Input data

As the input data set 28 classes of textural images from the open image database Kylberg Texture Dataset v. 1.0 are used [10]. Each class contains 160 unique monochrome photographs without rotations measuring 576×576 pixels. All images are normalized with a mean value of 127 and a standard deviation of 40 (Figures 1*a* and 1*b*).

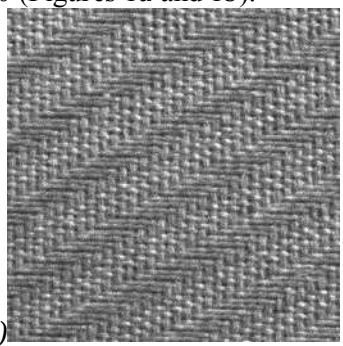
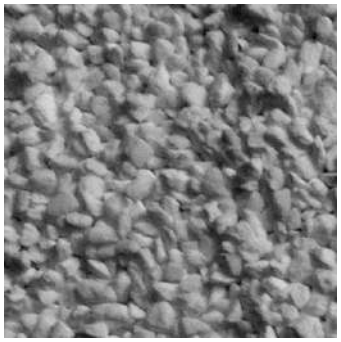
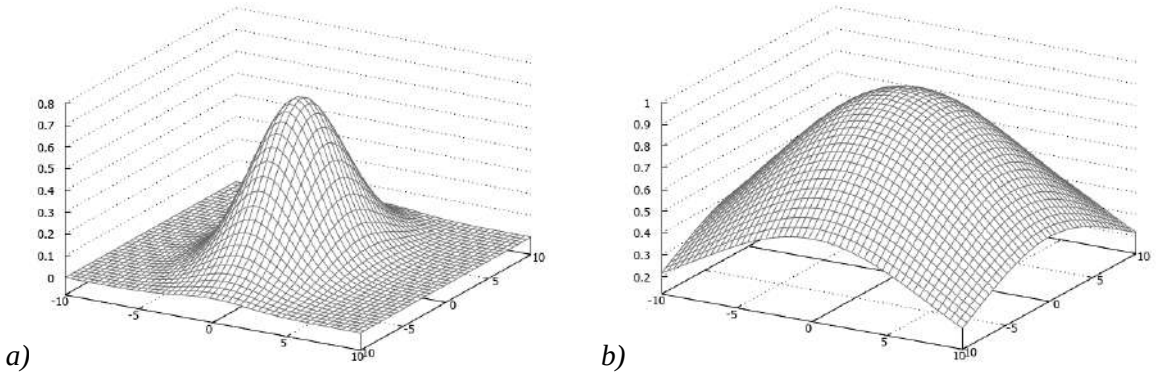
a)  *b)* 

Figure 1. Examples of images on which experiments were performed: seat (*a*), pearl sugar (*b*).

For each image a vector of 30 features is constructed. The first 25 features are formed as the samples of the correlation function $\rho(x, y)$ (1), $x = \overline{0,4}$, $y = \overline{0,4}$, but instead of the value $\rho(0,0) \equiv 1$ the count $\rho(-1,1)$ is used. Features 25-27 are the coefficients a_1, a_2, a_3 from the regression equation (2), 28 and 29 are the coefficients b_1, b_2 from equation (3). The form of the regression equations (2) and (3) for the specific values obtained during the computational experiment is shown in Figures 2*a*-*b* and 3*a*-*b*, respectively. The values of the coefficients from the regression equations (2) and (3) are presented in table 1.

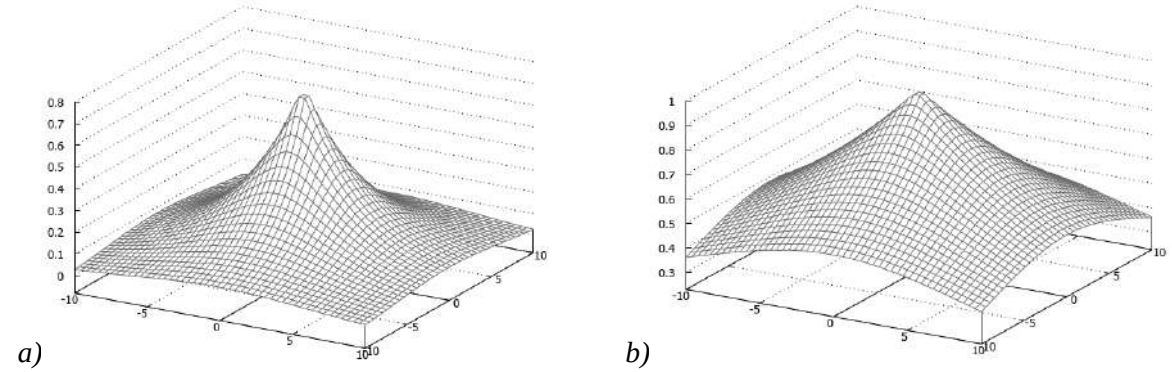
Table 1. The values of the coefficients of the regression equations for the graphs presented in Figures 1*a* and 1*b*.

Coefficients	a_1	a_2	a_3	b_1	b_2
Seat	-0.08999	-0.03208	-0.30889	-0.23781	-0.23272
Pearl sugar	-0.00835	-0.00652	-0.04151	-0.02968	-0.02982



a) *b)*

Figure 2. The graph of the regression equations (2): for Figure 1*a* (*a*), for Figure 1*b* (*b*)



a) *b)*

Figure 3. The graph of the regression equations (3): for Figure 1*a* (*a*), for Figure 1*b* (*b*)

Also, the "Class" field is added to the characteristics vector, which has a range of values from 1 to 28 and will be used for the model training.

3.2. The scheme of the computational experiment

Before the beginning of the research it is necessary to scale out the feature set, since the numerical values are presented in different scales, according to the following formulas. To normalize the sample it is necessary to divide each feature x_i into the sample rate:

$$x_i = x_i(\sum_{k=1}^{K} x_k)^{-1/2}, i = \overline{1,K},$$

where K is the total number of features. For standardization, we use the formula:

$$x_i = (x_i - M(i))R(i,i)^{-1/2}, i = \overline{1,K},$$

where $M(i)$ is the mean value and $R(i,i)$ is the standard deviation.

To classify images, it is necessary to form a training sample. In this paper we form it according to the principle of leave-one-out cross-validation:at each iteration step of the training there is a sample of $K-1$ elements, where K is the total sample size. Due to this, we increase the accuracy of classification (it is defined as the ratio of the number of correctly defined classes to their total number), although we increase the computational complexity of the algorithm. We distribute an image with an unknown class to a particular group based on the nearest neighbour, finding the vector closest to the desired vector by the set of characteristics.

During the first computational experiment we form the training sample consisting of the object by the cross-validation method. We successively determine the class of each image not included in the training sample. As a feature set, we use all the 30 features in the first case, in the second – only features with numbers 0-24, in the third – 25-29. During the second computational experiment, the training sample is formed in a similar way, but the attribute space is formed by the full search method

as a subset of the cardinalities one and three, as well as the random search method with capacities of four and five.

To compare the results of computational experiments we will classify the vectors for each image and evaluate the quality criterion J, assigned as the proportion of correctly recognized objects to their total number. Object classification should be carried out in its characteristic space, obtained according to the requirements of the computational experiments described above.

3.3. Results of experiments

The obtained results for the classification of texture images in accordance with the task of the first computational experiment are presented in table 2.

Table 2. Results of the first computational experiment.

Number of features	Cardinality	Reliability, %	Time, s
0, 1, …, 29	30	86.93	876.71
0, 1, …, 24	25	69.82	732.74
25, 26, 27, 28, 29	5	53.19	72.59

Table 2 shows that the highest proportion of correctly classified objects is achieved using a full set of features. However, in this case, the presence of non-informative features is possible, which will reduce the reliability of the classification. Also, a large subset of characteristics significantly increases the computational complexity of the algorithm and, as a consequence, increases the calculation time. Therefore, it is necessary to select features, which will reduce re-training, increase reliability and shorten the classification time.

The most informative features for the classification problem were signs under numbers 4, 0 and 5. The individual qualities of the five best attributes are given in table 3.

Table 3. Individual qualities of features.

Number of features	Reliability, %
4	44.02
0	35.09
5	32.54
9	30.84
24	30.22

We form feature subsets of cardinality three by the method of complete enumeration. The average time necessary for classifying the images using the generated subsets was 72.36 seconds. Table 4 shows that the first two best subsets of the features consist of only the samples of the correlation function, the third one also includes a sign derived from the coefficient of the regression equation. The reliability of the classification using the most effective feature subset of cardinality three is 1.13 times lower than the classification using the full feature set. However, the time spent on classification is 11.38 times lower.

Table 4. The feature selection result using a subset of cardinality three.

Number of features	Reliability, %
4, 14, 19	77.05
3, 4, 19	76.38
4, 19, 29	76.38

We use the method of random search to construct subsets of cardinalities four and five. The classification using a feature subset of cardinality four is 1.05 times as low as the classification proposed in the first computational experiment. However, when using a subset of cardinality five reliability of classification is 1.04 times higher.

Table 5 shows that in the formation of a characteristic subset the most informative features also participate: in the subset of cardinality four are the sign at number four, in the subset of cardinality of five, the sign with the number zero and four. The average classification time was 67.14 and 73.44 seconds when using subsets of cardinalities four and five, which gives a gain in time of 13.06 and 11.94 times, respectively.

Table 5. The feature selection results using subsets of cardinalities four and five.

Number of features	Reliability, %
4, 14, 28, 29	82.77
4, 14, 19, 25	82.59
4, 19, 26, 29	82.59
0, 2, 3, 25, 29	90.11
0, 2, 3, 25, 27	89.69
0, 2, 3, 4, 29	89.22

Thus, after carrying out series of experiments, the usage of the feature subset of cardinality five was the most effective, with the help of which it was possible to achieve a quality classification of more than 90%. Also the need to reduce feature set by removing the least informative ones was confirmed experimentally.

4. Conclusion

In the recent work, the objects were classified on the base of the calculated features. A number of features were formed using the samples of the correlation function and the coefficients from the regression equations. The selection of features was carried out with the methods of complete and random search.

The features that are the most informative and necessary in the construction of the classifier were revealed, they turned out to be the readings of the correlation function $\rho(1,0)$, $\rho(0,1)$, $\rho(1,1)$. Their inclusion in a variety of attributes for further classification will increase its reliability, allowing one to remove less informative features. The features, which are coefficients from the regression equations, were not among the top ten, which indicates their little information in the task of classifying images. Among the features that have the least impact on the classification process, are the coefficients a_3 and a_2 from the regression equations $\rho_1(x,y) = \exp(a_1x^2 + a_2y^2 + a_3)$ and $\rho_2(x,y) = \exp(b_1\sqrt{x^2+y^2} + b_2)$, as well as the correlation function $\rho(0,2)$.

The highest reliability of the image classification, which amounted to 90.11% was achieved with the use of a feature subset of cardinality five, which included the samples of the correlation function $\rho(0,1)$, $\rho(0,3)$, $\rho(0,4)$ and the coefficients a_1 and b_2 from the regression equations.

In this way, the hypothesis about the need to exclude the least informative features from the set was experimentally confirmed. When using the full set of attributes for image classification, the achieved reliability was 86.83%, which is 1.04 times as low as the maximum reliability obtained in the course of the computational experiment. Also, the time spent on classification decreased by 11.94 times by reducing the cardinality of feature set.

5. References

[1] Petrou M 2006 *Image processing: Dealing with texture* (John Wiley & Sons, Ltd)
[2] Viegas F, Rocha L, Goncalves M, Mourao F, Salles T, Andrade G and Sandin I 2018 A Genetic Programming approach for feature selection in highly dimensional skewed *Neurocomputing* **273** 554-569
[3] Gaidel A V, Zelter P M, Kapishnikov A V and Khramov A G 2014 Computed tomography texture analysis capabilities in diagnosing a chronic obstructive pulmonary disease *Computer Optics* **38(4)** 843-850
[4] Tang X, Dai Y, Sun P and Meng S 2018 Interaction-based feature selection using Factorial Design *Neurocomputing* **281** 47-54

[5] Goswami S, Das A K, Chakrabarti A and Chakraborty B 2017 A feature cluster taxonomy based feature selection technique *Expert System with Applications* **79** 76-89
[6] Rahmaninia M and Moradi P 2017 OSFSMI: Online Stream Feature Selection Method based on Mutual Information *Appl. Soft Comput.* **68** 733-746
[7] Maa B and Xia Y 2017 A tribe competition-based genetic algorithm for feature selection inpattern classification *Appl. Soft Comput.* **58** 328-338
[8] Gaidel A V 2015 A method for adjusting directed texture features in biomedical image analysis problems *Computer Optics* **39(2)** 287-293 DOI: 10.18287/0134-2452-2015-39-2-287-293
[9] Gaidel A V 2016 Matched polynomial features for the analysis of grayscale biomedical images *Computer Optics* **40(2)** 232-239 DOI: 10.18287/2412-6179-2016-40-2-232-239
[10] Kylberg G 2011 The Kylberg Texture Dataset v. 1.0: External report (Blue series) *Centre for Image Analysis, Swedish University of Agricultural Sciences and Uppsala University* **35**

Acknowledgments
The work was partially funded by the Russian Foundation of Basic Research grant 16-41-630761 p_a, by the Federal Agency for Scientific Organizations under agreement No. 007-GZ/Ch3363/26 and by the Russian Federation Ministry of Education and Science.

Parameterization of the nonlinear interpolator invariant to four directions contours for multidimensional digital signals

A I Maksimov[1] and M V Gashnikov[1]

[1]Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. In this paper, a parameterization of the nonlinear interpolator invariant to four directions contours for the digital signals is described. The interpolator automatically selects one of the possible calculation methods for each signal sample based on the presence of the contour in a sample. Training procedure is performed to calculate the optimal value of the interpolator parameter. A criterion for minimizing the post-interpolation residues energy and a criterion for minimizing of the post-interpolation residues entropy are used to find the optimal value. The proposed interpolator is applied to the task of signal compression and the task of combining heterogeneous signals. In order to examine interpolators, computational experiments are carried out on a test set. The advantage in terms of the quadratic error of suggested interpolator over prototypes is shown.

1. Introduction

During recent years, the size of processed multidimensional digital signals has been expanding faster than the storage capacities. That is why compression methods nowadays are in a great demand.

Today, a large number of compression methods have been developed. JPEG method [1] is, perhaps, the most common of them. It uses statistical coding [2, 3] of transformants obtained by the two-dimensional discrete cosine transform [4]. Other popular compression methods are based on two-dimensional orthogonal transformations [5], fractals [6] and wavelets [7].

The listed methods have one common flaw – they are resource-intensive, which makes it difficult to use them in real-time systems, for example, in on-board imaging systems of satellites, drones and other aircrafts. These systems need compression methods that have little computational complexity. Methods which perform signal processing in a spatial domain, without transitioning to spectral components will be perfect for such application.

Methods based on differential pulse-code modulation (DPCM) [7-10] satisfy these requirements. Decorrelation of the signal in DPCM compression is achieved via the use of difference representation. However, the scope of DPCM is not limited to real-time systems, for example, DPCM acts as a stage in other compression methods. For example, JPEG algorithm uses DPCM compression at one of the stages of operation. Thus, the task of enhancement of DPCM and increasing its efficiency is topical.

In this paper, we propose a new parameterized interpolator that makes it possible to increase the efficiency of DPCM compression by switching between different interpolation methods for each processed sample, depending on the presence of a contour in the vicinity of it.

The work is arranged in the following way. First, a general differential pulse-code modulation method, main types of DPCM interpolators and quantizers are presented. The next paragraph outlines

the interpolation and training procedures of proposed interpolator. The third paragraph shows the results of experimental study of proposed interpolator.

2. Differential pulse-code modulation

DPCM compression is the sequence of following procedures – interpolation, quantization, reconstruction, and coding. The generalization of this method is presented below.

2.1. General DPCM compression

Let $C(\bar{v})$ be the original (compressible) multidimensional digital signal where $\bar{v}$ is the vector of its arguments. Differential pulse-code modulation can be represented as the following sequential procedure:

1. First, a reference value $\bar{C}(\bar{v})$ is interpolated;
2. Then, the differential signal $f(\bar{v})$ is calculated as the difference between reference and original value of the signal (1):

$$f(\bar{v}) = C(\bar{v}) - \bar{C}(\bar{v}), \quad (1)$$

where $C(n,m)$ – is the original value of the signal.

3. The resulting differential signal is quantized (2):

$$f_q(\bar{v}) = Q(f(\bar{v})), \quad (2)$$

where Q –is the quantization function.

4. The reconstructed signal value $\hat{C}(\bar{v})$ id calculated as the sum of the quantized differential signal and the reference value. The reconstructed value (3) is used to further interpolation:

$$\hat{C}(\bar{v}) = \bar{C}(\bar{v}) + f_q(\bar{v}); \quad (3)$$

5. After processing the entire signal, the quantized differential signal $f_q(n,m)$ is statistically encoded.

2.2. DPCM compression: Interpolation

High DPCM compression ratio is achieved when the difference signal is close to zero. Thus, it is necessary to interpolate the reference value as close to the reference one as possible. There are three main types of DPCM interpolators: averaging [8], nonlinear [9] and adaptive [10].

Averaging (linear) interpolators calculate the reference value the average over several previous signal samples. The simplest linear interpolator is the one that chooses the previous sample as the reference value. Let us describe the work of interpolators for the case of a two-dimensional original (compressible) signal $C(\bar{v}) = C(n,m)$.

Figure 1 shows the work of the interpolator, averaging over the two previous samples [8]. The interpolated value in this case is formed as follows:

$$\bar{C}(n,m) = \left[\frac{\hat{C}(n-1,m)+\hat{C}(n-1,m)+\hat{C}(n-1,m-1)+\hat{C}(n+1,m-1)}{4}\right]. \quad (4)$$

Interpolators of this type are robust to noise, since they use averaging over several samples, but they can have a huge error on the contours.

Non-linear interpolators have a small error on the contours. The main idea of such interpolators is to interpolate the value of the signal "along" the contour using the values of the previous signal samples.

Graham interpolator [9] is a non-linear one. It is able to interpolate "along» horizontal and vertical contours. As the reference value the sample corresponding to the smallest of the following differences (5,6,7) is taken:

$$\bar{C} = \begin{cases} \hat{C}(n,m-1), \; if \; \lambda_n > \lambda_m \; ; \\ \hat{C}(n-1,m), \; if \; \lambda_n \le \lambda_m, \end{cases} \quad (5)$$

here

$$\lambda_n = \left|\hat{C}(n,m-1) - \hat{C}(n-1,m-1)\right|, \quad (6)$$

$$\lambda_m = \left|\hat{C}(n-1,m) - \hat{C}(n-1,m-1)\right|, \quad (7)$$

where λ_n – is the horizontal difference; λ_m – is the vertical difference. Figure 2 shows, how this interpolator works.

The interpolator invariant to the four directions contours [10] is the enhancement of Graham interpolator. It is able to interpolator "along" horizontal, vertical and two diagonal contours. To select the reference value, 12 differences between the 8 previous samples are calculated. The interpolation is performed as follows :

$$\bar{C} = \begin{matrix} \hat{C}(n,m-1),\ if\ \lambda_| = \min(\lambda); \\ \hat{C}(n-1,m),\ if\ \lambda_- = \min(\lambda); \\ \hat{C}(n-1,m-1),\ if\ \lambda_/ = \min(\lambda); \\ \hat{C}(n+1,m-1),\ if\ \lambda_\backslash = \min(\lambda), \end{matrix} \quad (8)$$

here

$$\lambda_| = \left|\hat{C}(n,m-1) - \hat{C}(n-1,m-1)\right| + \left|\hat{C}(n-1,m-1) - \hat{C}(n-2,m-1)\right| + \left|\hat{C}(n-1,m+1) - \hat{C}(n-2,m+1)\right|, \quad (9)$$

$$\lambda_- = \left|\hat{C}(n-1,m-1) - \hat{C}(n-1,m-2)\right| + \left|\hat{C}(n-1,m) - \hat{C}(n-1,m-1)\right| + \left|\hat{C}(n-1,m+1) - \hat{C}(n-1,m)\right|, \quad (10)$$

$$\lambda_/ = \left|\hat{C}(n,m-1) - \hat{C}(n-1,m)\right| + \left|\hat{C}(n-1,m) - \hat{C}(n-2,m-1)\right| + \left|\hat{C}(n-1,m-1) - \hat{C}(n-2,m)\right|, \quad (11)$$

$$\lambda_\backslash = \left|\hat{C}(n,m-1) - \hat{C}(n-1,m-2)\right| + \left|\hat{C}(n-1,m) - \hat{C}(n-2,m-1)\right| + \left|\hat{C}(n-1,m+1) - \hat{C}(n-2,m)\right|, \quad (12)$$

where $\lambda_|$ – is the vertical difference; λ_-– is the horizontal difference; $\lambda_/$– is the 45° angle diagonal difference; $\lambda_\backslash$–is the 135° angle diagonal difference.

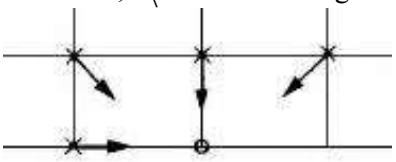


Figure 1. Averaging over four previous samples interpolator.

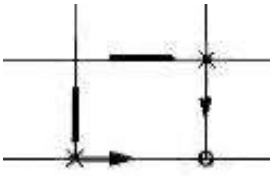


Figure 2. Graham interpolator.

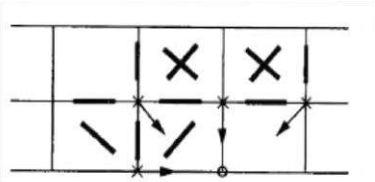


Figure 3. Interpolator invariant to the four directions contours.

Since nonlinear interpolators do not use averaging, they are subject to noise. Adaptive interpolators use the advantages of both averaging and non-linear interpolators. The adaptive interpolator at each point switches between the averaging and nonlinear method of interpolation of the reference value depending on the local features of the signal. To select an interpolation method, a criterion is chosen whose value at each point determines the choice of the interpolation method. Each interpolation method is meant to be used when it produces the smallest interpolation error - the averaging interpolator on the "smooth" sections of the signal, the nonlinear one - on the contours. The work of this interpolator is shown in Figure 3.

Adaptive Graham interpolator [10] is an example of adaptive interpolators. In each processed sample, the difference of the differences from formulas (6) and (7) is calculated. The calculated criterion is compared with the thresholds that are set for each processed signal by the preliminary training procedure. The interpolation procedure can be written as follows :

$$\bar{C} = \begin{cases} \hat{C}(n,m-1),\ if\lambda < \lambda_-; \\ \hat{C}(n,m-1) + \hat{C}(n-1,m),\ if\ \lambda \in [\lambda_-, \lambda_+]; \\ \hat{C}(n-1,m),\ if\lambda > \lambda_+, \end{cases} \quad (13)$$

$$-C_{max} \le \lambda_- \le 0 \le \lambda_+ \le C_{max}, \quad (14)$$

here

$$\lambda_n = \left|\hat{C}(n,m-1) - \hat{C}(n-1,m-1)\right|, \quad (15)$$

$$\lambda_m = \left|\hat{C}(n-1,m) - \hat{C}(n-1,m-1)\right|, \quad (16)$$

$$\lambda = \lambda_n - \lambda_m, \quad (17)$$

where λ is the contour criterion; λ_- is the lower threshold for switching between interpolation methods; λ_+ is the upper threshold for switching between interpolation methods; C_{max} is the maximum value of the signal. Figure 4 shows the work of adaptive Graham interpolator.

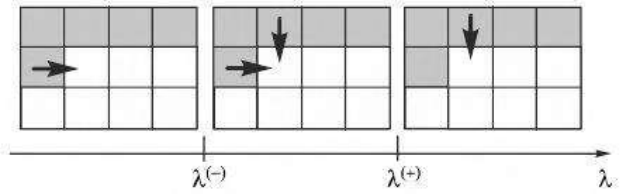


Figure 4. Adaptive Graham interpolator.

The disadvantage of this interpolation method is the need for a preliminary training procedure to calculate the thresholds which allow switching between interpolation methods.

2.3. DPCM compression: Quantization

Quantization [8,11] is the process of dividing the signal value range into a finite number of ranges, each of which is represented by one specific value.

Quantization allows increasing the compression ratio of the processed signal by introducing an error in it. The quantization of the difference signal has a special feature - a small number of quantization levels, therefore, they should be chosen thoroughly. In the present study, two quantization scales are used - uniform [12] and Max [13] quantization scales.

All intervals of the uniform quantization scale are equal, its representative values are located in the middle of the intervals and at equal distances from each other. The uniform quantization scale is a scale with a controllable error [14]. The maximum error introduced into the quantized signal is determined by the width of the scale interval. Uniform scale quantization can be written as follows:

$$f_q(\vec{v}) = sign\big(f(\vec{v})\big) \times \left[\frac{|f(\vec{v})| \times E_{max}}{2E_{max}+1}\right], \tag{18}$$

where E_{max} – is the maximum error.

The advantage this quantization scale is the level proportionality of the quantized signal and the original one.

Intervals of Max quantization scale are not equal. Its levels are constructed on the basis of the mean-square quantization error minimization criterion [13]. With a known distribution density of the difference signal, expression (19) for the standard-mean quantization error can be written as follows:

$$\mathcal{E}_q^2 = \sum_{l=0}^{L-1} \int_{d_l}^{d_{l+1}} (C - r_l) p(C)\, dC, \tag{19}$$

where $\mathcal{E}_q^2$ – is the quantization root-mean-square error; L – is the number of quantization levels; d_l – is the border value; r_l – is the representative value; $p(C)$- is the differential signal distribution density.

Since the difference signal is discrete, (19) can be written as:

$$\mathcal{E}_q^2 = \sum_{l=0}^{L-1} \sum_{C=d_l}^{d_{l+1}} (C - r_l) p(C). \tag{20}$$

To construct the Max scale, it is necessary to choose d_l and r_l which would minimize the error (20). Expressions (21,22) for them are found by taking the partial derivatives of expression (20):

$$r_i = \frac{\sum_{C=d_i}^{d_{i+1}} Cp(C)}{\sum_{C=d_i}^{d_{i+1}} p(C)}, \tag{21}$$

$$d_j = \frac{r_{j-1}+r_j}{2}. \tag{22}$$

In practice, Max scale is constructed from the uniform one iteratively. Its border values are calculated with the use of its representative values, and vice versa. The resulting Max scale levels are more frequent on the signal value area where the values most often appear.

3. Parameterization of the nonlinear interpolator invariant to four directions contours

We propose a new adaptive interpolator. It is a parameterization of the interpolator invariant to the four directions contours. Interpolation and trainig procedures are developed for proposed interpolator.

During training procedure a preliminary pass through the original signal is required to calculate the threshold value, upon which the interpolation method is chosen.

The contour criterion for the given parameterized interpolator is the difference module between the reference values obtained by means of the interpolator averaging over the four previous readings (4) and the interpolator invariant to the four directions contours (8). If the difference absolute value between the averaging and nonlinear methods is small, averaging interpolation is chosen, since it is noise-resistant. If the difference absolute value is large, then a contour is detected and nonlinear interpolation is chosen. Let us describe the work of interpolator for the case of a two-dimensional original signal $C(\bar{v}) = C(n,m)$.

3.1. Interpolation procedure

The interpolation procedure for proposed interpolator can be represented as follows:

$$\bar{C} = \begin{cases} (\hat{C}(n-1,m) + \hat{C}(n,m-1)\hat{C}(n-1,m-1) + \hat{C}(n+1,m-1))/4\,, \\ \quad if\left|min(\lambda) - \lambda_m\right| \le \Lambda; \\ \widehat{C}\ (n,m-1),\ if\ \lambda_| = min(\lambda) \wedge |\lambda_| - \lambda_m| > \Lambda; \\ \widehat{C}\ (n-1,m),\ if\ \lambda_- = min(\lambda) \wedge |\lambda_- - \lambda_m| > \Lambda; \\ \hat{C}(n-1,m-1),\ if\ \lambda_/ = min(\lambda) \wedge |\lambda_/ - \lambda_m| > \Lambda; \\ \hat{C}(n+1,m-1),\ if\ \lambda_\backslash = min(\lambda) \wedge |\lambda_\backslash - \lambda_m| > \Lambda, \end{cases} \quad (23)$$

here

$$\lambda_| = \left|\hat{C}(n,m-1) - \hat{C}(n-1,m-1)\right| + \left|\hat{C}(n-1,m-1) - \hat{C}(n-2,m-1)\right| + |\widehat{C}(n-1,m+1) - \hat{C}(n-2,m+1)|, \quad (24)$$

$$\lambda_- = |\widehat{C}(n-1,m-1) - \hat{C}(n-1,m-2)| + |\hat{C}(n-1,m) - \hat{C}(n-1,m-1)| + |\widehat{C}(n-1,m+1) - \hat{C}(n-1,m)|, \quad (25)$$

$$\lambda_/ = |\widehat{C}(n,m-1) - \hat{C}(n-1,m)| + |\hat{C}(n-1,m) - \hat{C}(n-2,m-1)| + |\widehat{C}(n-1,m-1) - \hat{C}(n-2,m)|, \quad (26)$$

$$\lambda_\backslash = |\widehat{C}(n,m-1) - \hat{C}(n-1,m-2)| + |\hat{C}(n-1,m) - \hat{C}(n-2,m-1)| + |\widehat{C}(n-1,m+1) - \hat{C}(n-2,m)|, \quad (27)$$

$$\lambda_\mathrm{m} = \left[\frac{\hat{C}(n-1,m)+\hat{C}(n-1,m)+\hat{C}(n-1,m-1)+\hat{C}(n+1,m-1)}{4}\right], \quad (28)$$

where Λ – is the interpolation threshold. Figure 5 shows the work of proposed interpolator. The procedure for calculation of the threshold value Λ is described in the next section of the article.

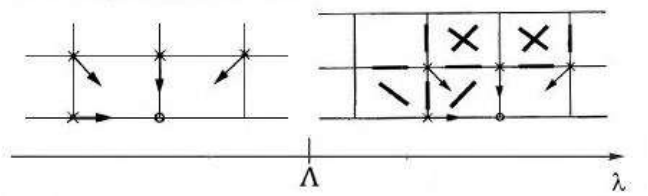


Figure 5. Parameterized nonlinear interpolator invariant to four directions contours.

3.2. Post-interpolation residues energy minimization-based training procedure

To determine the value of the threshold Λ used for switching interpolation methods, it is necessary to solve the following optimization problem (29):

$$\delta(\Lambda) = \sum_{(m,n)\in\omega} |C(n,m) - \hat{C}(n,m)| \to min_\Lambda. \quad (29)$$

where δ – is the interpolation error (the energy of post-interpolation residues); ω – is the set of signal samples coordinates.

To solve this problem, a preliminary signal pass is required during which the matrix Δ (30-32) is filled. Its elements contain the total interpolation error (the sum of the difference modules between the signal sample and the interpolated value) for which criterion equals λ:

$$\Delta(1,\lambda) = \sum_{m,n\in w(\lambda)} |C(n,m) - min_{\lambda}\{\hat{C}(n-1,m), \hat{C}(n,m-1), \hat{C}(n-1,m-1), \hat{C}(n+1,m-1)\}|, \quad (30)$$

$$\Delta(2,\lambda) = \sum_{m,n\in w(\lambda)} |C(n,m) - \left[\frac{\hat{C}(n-1,m)+\hat{C}(n,m-1)+\hat{C}(n-1,m-1)+\hat{C}(n+1,m-1)}{4}\right]|, \quad (31)$$

$$-C_{max} \le \lambda \le C_{max}, \quad (32)$$

where $w(\lambda)$ – is the set of λ values. The error vector is filled via recursive procedure:

$$\delta(\lambda_{max}) = \sum_{\lambda=0}^{C_{max}} \Delta(1,\lambda), \quad (33)$$

$$\delta(\lambda) = \delta(\lambda+1) - \Delta(1,\lambda) + \Delta(2,\lambda), C_{max} - 1 \ge \lambda \ge 0. \quad (34)$$

After this operation, the value of the threshold Λ can be found from expression (29). This vector contains $C_{max} + 1$ elements so one can find the required threshold value by a simple search. Thus, the optimization problem (29) is solved.

It should be noted that the proposed interpolator, including the described optimization procedure, can be used not only for compression, but also for other signal processing tasks, for example, for solving the problem of combining heterogeneous signals which have different resolution, number of components, etc.

3.3. Post-interpolation residues entropy minimization-based training procedure

When solving the compression problem, instead of (29), it is also worth considering a criterion for minimizing the amount of compressed data. In this paper we consider the quantized post-interpolation residues entropy of (18) as an estimation of the amount of compressed data:

$$H(\Lambda) = -\sum_{q=-C_{\max}}^{C_{\max}} N_q(\Lambda) \ln N_q(\Lambda) \to \min_{\Lambda} \quad (35)$$

where $N_q(\Lambda)$ is the amount of the quantized post-interpolation residuals with a value q.

To solve the optimization problem (35), a three-dimensional matrix is used:

$$\mathbf{N}_{\lambda,q}^{(i)},\ i \in \{1,2\},\ 0 \le \lambda < C_{\max}. \quad (36)$$

Each element of $\mathbf{N}_{\eta,q}^{(i)}$ contains the amount of quantized post-interpolation residues (18) with a value q, given the feature value (24) λ and interpolator number i, which is taken from (23). i =0 means that the first row of (23) is taken as the interpolator, for i =1 all other rows of (23) are taken as interpolator.

Matrix $\mathbf{N}_{\lambda,q}^{(i)}$ is used to calculate the amount of quantized post-interpolation residuals (18) $N_q(\lambda)$ with value q for all possible values of the threshold λ:

$$N_q(0) = \sum_{\lambda=0}^{C_{\max}-1} \mathbf{N}_{\lambda,q}^{(1)},\ N_q(\lambda+1,q) = N_q(\lambda,q) - \mathbf{N}_{\lambda,q}^{(1)} + \mathbf{N}_{\lambda,q}^{(2)}. \quad (37)$$

The number of values of the quantized post-interpolation residues (18) equal to q makes it possible to calculate a one-dimensional array of entropy values (35) for all possible values of the threshold λ. The length of this array is small (C_{max} of the elements, i.e. equal to the maximum value of the signal). The number of the minimum entropy value (35) can be found by the direct search:

$$\Lambda = \arg\min_{\lambda}\left(H(\lambda)\right). \quad (38)$$

Thus, the optimization problem (35) is solved.

4. Experimental study of DPCM interpolators

For the developed interpolator based on criterion (29) a series of computational experiments was carried out, during which the dependencies of the root-mean-square error (RMS) on the compression ratio of the quantization scale were obtained. The uniform (18) and the Max scale (21, 22) were used. The mathematical expression for the root-mean-square is as follows:

$$\mathcal{E} = \frac{1}{\delta_C}\sum_{\bar{v}\in V}\sqrt{(C(\bar{v}) - \tilde{C}(\bar{v}))^2}, \quad (39)$$

where δ_C – is the original signal variance; V– is the set of signal samples; $C(\bar{v})$ – is the original signal values; $\tilde{C}(\bar{v})$ – is the processed (decompressed) signal values.

The results obtained for the developed interpolator were compared with the others. For comparison, prototypes on the basis of which a parameterized interpolator was developed were taken. They are averaging over the four previous samples (4) and invariant to the contours of the four directions (8-12) interpolators. As a test signal set Waterloo Grayscale Set 1 [15] was used. Figure 6 shows the test set. The computational experiment was carried out in the following way: the test signals were DPCM-compressed [16] and decompressed with a constant compression ratio of a uniform quantization scale and a selected interpolator, after which the error was calculated. The results for the selected interpolator were averaged over all test set. Then the procedure was repeated for the next compression ratio of the scale. Thus, the dependencies of the root-mean-square error on the compression ratio of a uniform quantization scale were constructed for all three interpolators. A similar series of experiments was carried out for the Max quantization scale. Figures 7 and 8 show the results of the computational experiments.

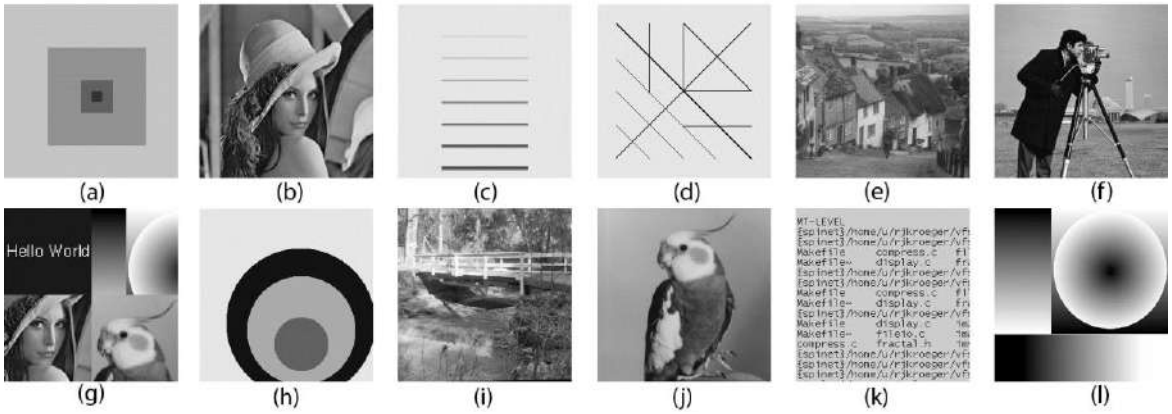


Figure 6. Waterloo Greyscale Set 1 test signals, signals (a), (c), (d), (h) are inverted.

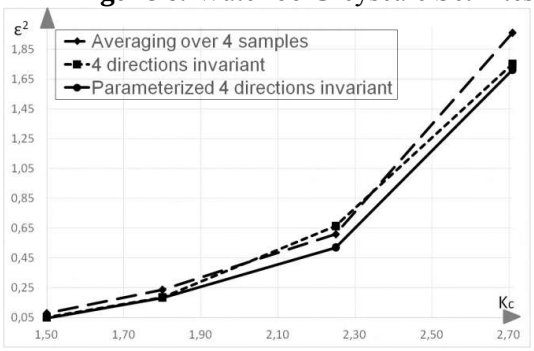


Figure 7. Dependency of RMS on compression ratio for uniform quantization scale for each interpolator.

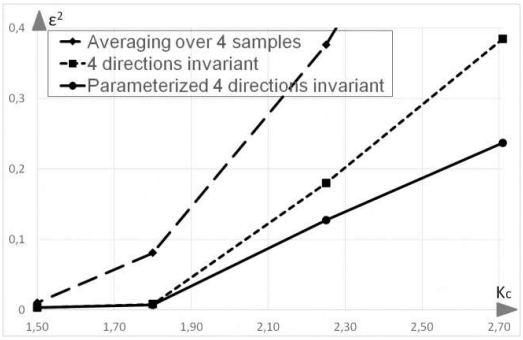


Figure 8. Dependency of RMS on compression ratio for Max quantization scale for each interpolator.

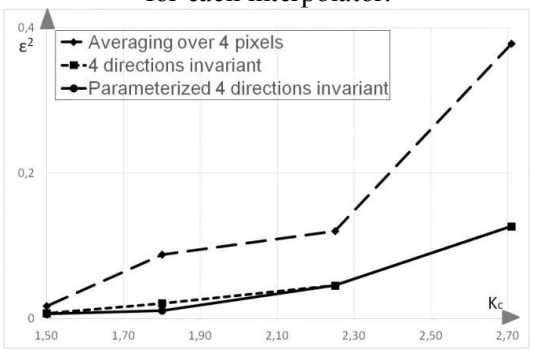


Figure 9. Dependency of RMS on compression ratio for uniform quantization scale for Fig.6c.

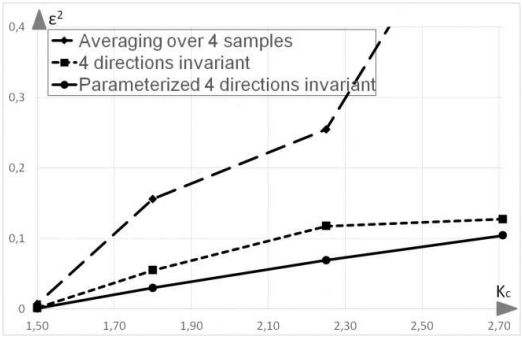


Figure 10. Dependency of RMS on compression ratio for Max quantization scale for Fig.6c.

As can be seen from the presented dependencies, the developed parameterization outperforms its prototypes. During computational experiments, the type of signal was found, with which the developed parameterization gives the least RMS. This result is shown in Figures 9 and 10.

5. Conclusions

This article presents a new adaptive interpolator for DPCM-signal compression, based on the parameterization of the interpolator, which is invariant to four directions contours. Its interpolation and training procedures are presented. Training procedure includes a preliminary pass through the signal.

A series of computational experiments was carried out. The developed parameterization outperformed its prototypes. During the computational experiments, the type of signals on which the proposed interpolator shows the best results was determined.

6. References

[1] Gupta V 2014 Enhanced Image Compression Using *Wavelets International Journal of Research in Engineering and Science (IJRES)* **2(5)** 55-62
[2] Sayood K 2012 *Introduction to Data Compression* (The Morgan Kaufmann Series in Multimedia Information and Systems)
[3] Wallace G 1991 The JPEG Still Picture Compression Standard *Communications of the ACM,* **34(4)** 30-44
[4] Huffman D A 1952 A Method for the Construction of Minimum Redundancy Codes *Proc. IRE,* **40** 1098-1101
[5] MacKay D 2003 *Information Theory, Inference, and Learning Algorithms* (Cambridge Univ. Press)
[6] Plonka G M 2005 Fast and numerically stable algorithms for discrete cosine transforms *Linear Algebra and its Applications* **394(1)** 309-345
[7] Woon W M 2000 Achieving high data compression of self-similar satellite images using fractal *Proceedings of IEEE International Geoscience and Remote Sensing Symposium* 609-611
[8] Pratt W 2007 *Digital Image Processing* (Wiley)
[9] Gonzalez R 2008 *Digital Image Processing* (Pearson Education)
[10] Soifer V A 2010 *Computer Image Processing. Part II: Methods and algorithms* (VDM Verlag)
[11] Gashnikov M V 2017 Parameterized adaptive predictor for digital image compression based on the differential pulse code modulation *Proceedings of SPIE The International Society for Optical Engineering* 1034110 DOI: 10.1117/12.2268530
[12] Sayood K 2006 *Introduction to data compression* (Morgan Kaufmann Publishers)
[13] Salomon D 2007 *Data Compression. The Complete Reference* (Springer-Verlag)
[14] Lin S 2004 *Error Control Coding: Fundamentals and Applications, second edition* (New Jersey: Prentice-Hall, inc)
[15] *Image Repository: The Waterloo Fractal Coding and Analysis Group* (Access mode: http://links.uwaterloo.ca/Repository.html) (17.11.2017)
[16] Gashnikov M V 2016 Parameterization of nonlinear Greham predictor for digital image compression *Computer Optics* **40(2)** 225-231 DOI: 10.18287/2412-6179-2016-40-2-225-231

Acknowledgements

The reported study was funded by RFBR according to the research projects № 18-01-00667, № 18-07-01312.

On the Problem of Development of Methods and Algorithms Based on the Object-Oriented Logic Programming for Intelligent Video Monitoring of Laboratory Rats

A A Morozov[1] and O S Sushkova[1]

[1]Kotel'nikov Institute of Radio Engineering and Electronics of RAS, Mokhovaya str. 11-7, Moscow, Russia, 125009

Abstract. The problem of the video monitoring the laboratory rats by the means of the object-oriented logic programming is considered. The main task of the video monitoring is the analysis of the behavior of the animals in cognitive testing. An essential feature of the video records is in that the experiments are conducted in the same cage where the animal lives, that is, the background of the cage is sawdust. The color of the animals is about the same as the color of the sawdust; thus the detection of the animals is not a simple task. An additional difficulty is in that the videos were recorded simultaneously with electroencephalograms (EEG) in the animals; thus the head of the rat is connected with EEG cable that moves and causes false detections of recognition algorithms. In the paper, development of low-level algorithms for video analysis as well as logical methods for the analysis of the animal behavior is discussed. The methods and algorithms are implemented in the Actor Prolog object-oriented logic language.

1. Introduction

In recent years, automation of neurophysiological experiments on laboratory animals was recognized as an important direction in the computer vision and intelligent video monitoring [1, 2]. The methods of computer vision enable automation of routine of animal behavior analysis and, that is more important, make the results of the analysis independent from the human factor. Usually, a neurophysiological research is based on the relative change of the quantity of given events (for instance, actions performed by the laboratory animal) under given experimental conditions, but not on the absolute quantity of these events. Moreover, often it is not possible to estimate the exact quantity of the events because the behavior of the animal is not clearly expressed; in this case, the recognition of the events depends on the experience and subjective opinion of the experimenter. The state of the experimenter including his fatigue and his current ideas about the importance of various experimental events can influence the results of the recognition too. Therefore it is important to provide the constancy and uniformity of the recognition of required elements of animal behavior. This uniformity can be provided by the methods of automatic video analysis (intelligent video monitoring).

Another important problem that can be solved using intelligent video monitoring is standardization of neurophysiological experiments and providing reproducibility [3] of the experimental results by independent researchers in various laboratories.

There are free and commercial software available for automation of video processing and laboratory animal behavior analysis in biomedical experiments [4–8]. Nevertheless, new problems arise constantly in the neurophysiological experiments that require video processing that is out of capabilities of the existed software. It is expedient to use high-level programming languages specialized for intelligent video monitoring for solving these problems. In this paper, a video monitoring problem of such kind and video analysis methods used for solving this problem are considered.

The videos considered below are produced in neurophysiological experiments on the study of a convulsive electrical activity of the brain cortex. In the experiments, videos of a behavior of laboratory rats were recorded simultaneously with EEG signals. A comparison of EEG data with the behavior of the animals is necessary because sharp motions of the animals can result in EEG artifacts that are very similar to the epileptic discharges. Thus, the first task of the video monitoring is recognition of the sharp motions of the animals and using this information for proper interpretation of the results of the experiments. The second task of the video monitoring is the analysis of the behavior of animals in cognitive testing (in the tests on social recognition and recognition of a new object). An essential feature of the video records is in that the experiments are conducted in the same cage where the animal lives, that is, the background of the cage is sawdust. The color of the animals is about the same as the color of the sawdust; thus the detection of the animals is not a simple task.

Initial experiments with the video analysis have demonstrated that the methods of object detection implemented in commercial software based on the analysis of brightness, analysis of color, and background subtraction cannot provide stable recognition of the laboratory rats on the sawdust background. Thus, we have applied more sophisticated texture-oriented methods implemented in the Actor Prolog logic programming system for the recognition of the animals. The texture-oriented methods provide stable detection of the animals at the expense of a decrease of spatial resolution of the detection. In action, it implies a loss of information about the coordinates of the contour of the animal and impossibility of using modern model-based tracking methods for recognition of postures and actions of the animals. We have developed logical methods for analysis of the behavior of laboratory rats based on the information about the coordinates and velocity of objects that can be obtained using the low-level video processing means implemented in the Actor Prolog system [9–16]:

(i) The coordinates and velocity of the centroid of the animal are computed using the texture-based methods of image processing.

(ii) The coordinates of the EEG cap that connects the animal with the EEG cable are computed using the color-based methods.

(iii) The exact coordinates of auxiliary objects placed in the cage are computed using the color-based methods too.

We have developed a logical definition (a set of logical rules) of the exploratory behavior of the laboratory rats that provides an acceptable quality of recognition of the required behavior in the cognitive tests.

The logical approach to the definition and analysis of laboratory animal behavior is described in the first section of the paper. The experimental conditions and peculiarities of the video data to be processed are described in the second section. A description of an experimental program for the video analysis implemented in the Actor Prolog logic language and the results of the experiment are discussed in the third section.

2. The Logic Programming Approach to the Intelligent Video Monitoring

The idea of using mathematical logic and logic programming for intelligent video surveillance was developed in research projects W4 [17], VidMAP [18], VERSA [19], LTAR [20], RoboSherlock [21], Actor Prolog [16], etc. The idea is in that one applies logical formulae/rules for description and recognition of objects, situations, and events. One can explain the advantage of the logical approach to the intelligent video monitoring in the following way. The activity and behavior notions [22] differ in that the behavior of an object is the activity of the object related to the context information about the place, time, object attributes, etc. The information about the context allows deciding, for instance, whether the behavior of the object is abnormal and/or dangerous. Thus the analysis of the behavior is a more complicated problem than the analysis of the activity. It is necessary to describe and analyze the information about the context of the activity and the mathematical logic is perhaps the best instrument that can be used for this purpose.

The Actor Prolog language is an object-oriented logic language, that is, it combines expressiveness of the logical and object-oriented approaches to the programming [23–26]. This combination has increased the area of application of the logic programming. In particular, the object-oriented features enable to effort an opportunity to solve the problem of storing and processing big arrays of binary data (such as audio/video data) in the logic languages. The problem is in that plain logic languages do not implement data arrays directly, but use lists and structures for storing data because these data structures correspond to the Skolem functions in the first-order Predicate Calculus [27]. In the object-oriented logic languages, the arrays of data can be encapsulated in the instances of some specialized built-in classes. This enables fast and effective processing of the big data arrays in the logic languages.

The method of object-oriented logic programming of intelligent video surveillance was developed for the analysis of people behavior and recognition of abnormal activity [9–16]. The idea of the method is in the following:

(i) The stages of low-level and high-level processing of the video stream are separated.

(ii) The stage of the low-level video processing includes background subtraction, extraction of blobs, computing trajectories of the movements of the blobs, etc. The low-level processing is performed directly upon the video data arrays using special built-in classes of the logic language. The built-in classes are implemented in a procedural programming language to increase the speed of the data processing.

(iii) The stage of the high-level video processing includes analysis of trajectories/graphs of the blob movements. The algorithms of the high-level analysis are implemented in the Actor Prolog logic language in a form of logical rules. The graphs of the blob movements are described using the terms of the logic language: structures, lists, and underdetermined sets [23].

(iv) The logic programs written in Actor Prolog are translated to Java [14, 28]. The Java language is used as an intermediate language in the translation scheme to provide high performance and stable work of the intelligent video monitoring software.

New built-in classes of the Actor Prolog language were developed to use the language for the intelligent video monitoring of the laboratory animals. These new built-in classes implement new means of the low-level video processing that is necessary to analyze simultaneously blobs of different kinds extracted using different methods. The standard methods of blob extraction based on the background subtraction are not applicable in the case of sawdust background because the sawdust surface is changed permanently that is an effect of the movements of the rat and this is a cause of multiple false results during the background subtraction.

It turned out, that the methods of blob extraction based on the analysis of the brightness and hue of objects are of little use in the recognition of the rats in the cage with the sawdust background too. The reason for this is in that the color of the rats is about the same as the color of the sawdust. Theoretically speaking, the computer vision methods can differentiate the hues of the rats and the sawdust, but, in practice, the illumination inside the cage often is non-uniform andthe color of the rats and sawdust background is influenced by the shadows and reflections of the light from the colored objects and plastic walls of the cage. As a result of this, the standard methods of blob extraction based on the brightness and color also mistakes.

3. The Experimental Conditions

Let us consider a problem of intelligent video monitoring laboratory rats by the example of a neurophysiological experiment on the study of cognitive possibilities of the animal. By the terms of the experiment, one puts new objects into the cage with the animal. The animal explores the new objects and the experimenter estimates the total time spend by the animal to explore the objects. After a time, the experiment is to be repeated with the same objects. If the time spent by the animal to explore the objects is less than one in the first test, the experimenter can conclude that the animal remembers the objects. If the time is about the same, it means that the animal forgot the objects.

The exploratory activity of the rat is usually manifested in that it approaches the object and sniffs around the object (see Figure 1). It is difficult to describe the exploratory activity of the rat in a formal way; thus a simplified approach to detection of the exploratory activity is often used in the neurophysiological experiments: the distance between the rat and the object is estimated and the cumulative time when the rat is situated close enough to the object is calculated. Sometimes one estimates just the number of approaches of the rat to the object. This simplified method leads to mistakes and is not applicable to the experiment under consideration because the cage is small and the rat is situated not far from the object about all the time. At that the rat can sniff around the object or ignore it; it can just lie or dig the sawdust near the object.

Figure 1. A laboratory rat investigates a new object in the cage. The logic program has detected three blobs in the video: EEG cap that connects the head of the rat with EEG cable; a new green object; and a new blue object placed in the cage.

An additional problem is in that we cannot detect the contour of the animal in the background of sawdust. Thus we cannot recognize in a reliable way the face and forelegs of the animal. Instead of this, we estimate the following attributes of the experimental setting with the help of the low-level image processing procedures:

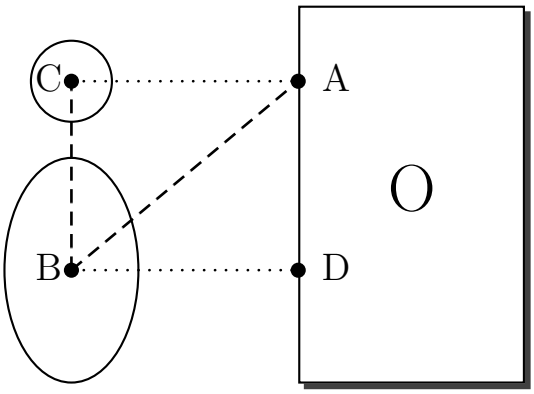


Figure 2. The attributes of the image to be used for detection of the exploratory activity of the rat.

(i) The coordinates and the velocity of the centroid of the blob related to the body of the rat are estimated using the texture-oriented method of image analysis that is sensitive to the smoothness of the object surfaces. The body of the rat is well visible in the sawdust background because the rat hair has a smoother surface in comparison with the sawdust.

(ii) The coordinates of EEG cap are detected using the color in the HSB space.

(iii) Exact coordinates of the objects placed in the cage are estimated using the color information too. The only difference is in that the coordinates of the object are estimated permanently in all the frames of the video and then averaged to avoid mistakes when the rat covers the object.

Let B be the centroid of the blob corresponding to the body of the rat (see Figure 2). Let C be the centroid of the blob corresponding to EEG cap of the rat. Let O be the object that is nearest to the C point. Let A be a point on the contour of the object O that is nearest to the C point. Let D be a point in the contour of the same object O that is nearest to the B point. The following combinations of these attributes have been recognized as useful for the detection of the exploratory activity of the rats during the experiments:

(i) If the distance between the A and C points is less than 1 cm, the rat probably sniffs the object. If the distance between the A and C points is more than 4 cm, the rat probably does not explore the object. If the distance lies in the interval 1–4 cm, additional analysis is necessary to determine whether the rat explores the object or not. The analysis is complicated by the fact that the rat can lie down sideways during the sniffing of the object; in this case, the $A - C$ distance can be big enough.

(ii) If the angle between the $B - C$ and $B - A$ lines is more than 50 degrees, the rat probably does not investigate the object. This heuristic rule reflects the fact that the rat turns his face to the object during the sniffing.

(iii) If the velocity of the B point is more than 3 cm per second, the rat probably just walks around the object, but does not explore it.

(iv) If the ratio between the $D - B$ and $A - C$ distances is less than 1.3, the rat probably does not explore the object. This rule reflects the fact that the rat usually stands aside from the object and pulls his face to the object during the sniffing.

These heuristic rules were implemented in a logic program for detection of the exploratory activity of the rats and tested in neurophysiological experiments.

4. The Logic Program for the Video Analysis

The Actor Prolog language has no built-in means of the fuzzy logical inference, however one can easily implement a kind of a fuzzy logical reasoning using the standard top-down resolution and standard built-in arithmetical operations. In particular, one can define the heuristic rules of rat behavior analysis described above in the following way:

```
PREDICATES:
determ:
sniffing_is_detected(REAL,REAL,REAL,REAL)    - (i,i,i,i);
imperative:
fuzzy_metrics(REAL,REAL,REAL) = REAL         - (i,i,i);
```

The *sniffing_is_detected* predicate succeeds if an exploratory behavior of the rat is detected. The predicate has four input arguments: distance $A - C$; distance $D - B$; the angle between the $B - C$ and $B - A$ lines; and the velocity of the B point.

```
CLAUSES:
sniffing_is_detected(AC,_,_,_):-
     AC < 0.01,!.
sniffing_is_detected(AC,_,_,_):-
     AC > 0.04,!,
     fail.
sniffing_is_detected(AC,DB,A,V):-
     M1== ?fuzzy_metrics(DB/AC,1.3,0.1),
     M2== 1-?fuzzy_metrics(A,50.0,10.0),
     M3== 1-?fuzzy_metrics(V,0.03,0.01),
     P== M1*M2*M3,
     P > ?power(0.5,3).
```

The *fuzzy_metrics* function is auxiliary one. It is used for definition of fuzzy thresholds. The function has three input arguments: the value to be checked; the threshold; and the width of the interval of uncertainty [14]:

```
fuzzy_metrics(X,T,H) = 1.0 :-
     X >= T + H,!.
fuzzy_metrics(X,T,H) = 0.0 :-
     X <= T - H,!.
fuzzy_metrics(X,T,H) = V :-
     V== (X - T + H) * (1 / (2*H)).
```

In Figure 3, the user interface of the logic program intended for the analysis of the rat behavior is shown. Control elements of the dialog window of the program allow selection of the video files and examination of the video frames at various rates in the forward and backward directions. The information on the total distance of movement, average velocity, cumulative time expended by the rat for investigation of new objects in the cage is demonstrated in the top of the dialog window. In the graphics window, the program demonstrates the frames of the video, detected blobs, and auxiliary information related to the behavior analysis. The blob related to the body of the rat is indicated by the orange color. The blob related to the EEG cap is indicated by the cyan color. The green and blue blobs correspond to the new objects placed in the cage in the course of the experiment. The logical rules defined above are used for the recognition of the exploratory activity of the rat.

The logic program was tested on five rats. The results of the tests are given in Table 1 : the total distance passed by the rat during the test; the average velocity of the rat; and the time of the exploratory activity of the rat (in percents of the total time of the test).

All five videos were marked manually to estimate the quality of the algorithm of detection of the exploratory activity.

The results of the automatic recognition were compared with the results of the manual marking to compute the sensitivity and specificity of the detection (see Table 2). The results of the application of standard algorithms SVM [29] and ANFIS [30] are given in the table for reference.

The tests demonstrated that the logic program ensures about the same sensitivity and specificity ($\approx 80\%$) as the SVM and ANFIS algorithms do on the base of the same data.

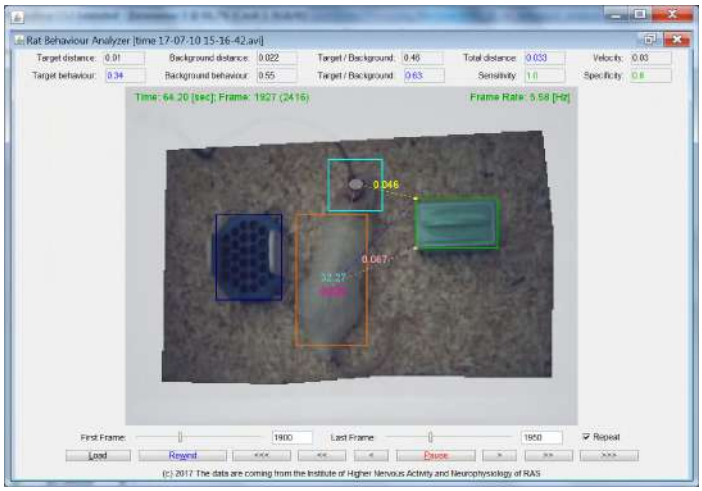


Figure 3. The user interface of the logic program intended for intelligent monitoring behavior of laboratory rats. The program estimates cumulative time expended by the rat for investigation of new objects in the cage.

Table 1. The results of analysis of laboratory rat behavior.

No.	Total time [s]	Distance [m]	Velocity [cm/s]	Exploratory behavior [%]
1	181	1.572	0.9	63
2	189	2.97	1.6	52
3	184	3.17	1.7	36
4	181	3.355	1.9	28
5	171	1.12	0.7	16

Table 2. The sensitivity and specificity of detection of the exploratory activity of the rats.

No.	The logic program		SVM		ANFIS	
	sens.	spec.	sens.	spec.	sens.	spec.
1	92	72	86	73	85	88
2	91	78	90	78	90	85
3	78	79	74	82	71	85
4	72	82	66	84	59	86
5	68	91	70	91	67	92

This quality of detection is comparable with the quality of manual detection and is enough for conduction of the neurophysiological experiment. Note that an advantage of the logic programming approach to the detection is in that it does not require a preliminary training of the program and, that is more important, logical rules are understandable for the experimenter and can be manually fixed and/or improved in any time. The values of the attributes used in the rules (the thresholds of angles and distances, etc.) can be assigned manually or computed (automatically or semi-automatically) on the base of statistical analysis of marked videos.

5. Conclusions

A method of intelligent video monitoring the laboratory animals based on the object-oriented logic programming is developed. The method is intended for intelligent video monitoring the laboratory rats in non-standard experimental conditions when one cannot apply existed software for automation of neurophysiological experiments. In particular, this method is applicable for analysis of the behavior of rats in cages with a sawdust background when the contrast of the images is low and the illumination is non-uniform. In the framework of the method, heuristic rules and fuzzy definitions are used for describing and recognition of the animal behavior. The method is implemented on the base of the Actor Prolog [16] object-oriented logic language.

6. References

[1] Robinson L and Riedel G 2014 *Journal of neuroscience methods* **234** 13-25
[2] Tscharke M and Banhazi T M 2016 *Journal of Agricultural Informatics* **7** 23-42
[3] Spruijt B M, Peters S M, de Heer R C, Pothuizen H H and van der Harst J E 2014 *Journal of neuroscience methods* **234** 2-12
[4] van Dam E A, van der Harst J E, ter Braak C J, Tegelenbosch R A, Spruijt B M and Noldus L P 2013 *Journal of neuroscience methods* **218** 214-224
[5] Noldus L P, Spink A J and Tegelenbosch R A 2001 *Behavior Research Methods* **33** 398-414
[6] Ohayon S, Avni O, Taylor A L, Perona P and Egnor S R 2013 *Journal of neuroscience methods* **219** 10-19
[7] Weissbrod A, Shapiro A, Vasserman G, Edry L, Dayan M, Yitzhaky A, Hertzberg L, Feinerman O and Kimchi T 2013 *Nature Communications* **4** 2018
[8] Giancardo L, Sona D, Huang H, Sannino S, Manag`o F, Scheggia D, Papaleo F and Murino V 2013 *PloS one* **8** e74557
[9] Morozov A A and Sushkova O S 2016 Real-time analysis of video by means of the Actor Prolog language *Computer Optics* **40**(6) 947-957 DOI: 10.18287/2412-6179-2016-40-6-947-957
[10] Morozov A A, Sushkova O S and Polupanov A F 2017 Object-oriented logic programming of 3D intelligent video surveillance: The problem statement *IEEE 26th International Symposium on Industrial Electronics (ISIE), Edinburgh, United Kingdom* 1631-1636
[11] Morozov A A, Sushkova O S and Polupanov A F 2017 Towards the distributed logic programming of intelligent visual surveillance applications *Proceedings Advances in Soft Computing: 15th Mexican International Conference on Artificial Intelligence* **2** 42-53
[12] Morozov A A 2015 *Pattern Recognition and Image Analysis* **25** 481-492
[13] Morozov A A and Polupanov A F 2015 Development of the logic programming approach to the intelligent monitoring of anomalous human behaviour *OGRW* (Koblenz: University of Koblenz-Landau) 82-85
[14] Morozov A A and Polupanov A F 2014 Intelligent visual surveillance logic programming: Implementation issues *CICLOPS-WLPE (Aachener Informatik Berichte)* 31-45
[15] Morozov A A, Vaish A, Polupanov A F, Antciperov V E, Lychkov I I, Alfimtsev A N and Deviatkov V V 2015 Development of concurrent object-oriented logic programming platform for the intelligent monitoring of anomalous human activities *BIOSTEC* **511** 82-97
[16] Morozov A A and Sushkova O S 2018 *The intelligent visual surveillance logic programming* (Access mode: http://www.fullvision.ru)
[17] Haritaoglu I, Harwood D and Davis L S 1998 W^4: Who? When? Where? What? A real time system for detecting and tracking people *FG (Japan)* 222-227

[18] Shet V, Harwood D and Davis L 2005 VidMAP: Video monitoring of activity with Prolog *IEEE AVSS* 224-229
[19] O'Hara S 2008 V*ERSA-video event recognition for surveillance applications* (M.S. thesis. University of Nebraska at Omaha)
[20] Artikis A, Sergot M and Paliouras G 2010 A logic programming approach to activity recognition *International Workshop on Events in Multimedia* 3-8
[21] Worch J H, Bálint-Benczédi F and Beetz M 2016 *KI - Künstliche Intelligenz* **30** 21-27
[22] Borges P V K, Conci N and Cavallaro A 2013 *IEEE Transactions on Circuits and Systems for Video Technology* **23** 1993-2008
[23] Morozov A A 1999 Actor Prolog: an object-oriented language with the classical declarative semantics *IDL* (Paris, France) 39-53
[24] Morozov A A 2002 On semantic link between logic, object-oriented, functional, and constraint programming *MultiCPL* 43-57
[25] Morozov A A 2007 Operational approach to the modified reasoning, based on the concept of repeated proving and logical actors *CICLOPS* 1-15
[26] Morozov A A 2003 *Pattern Recognition and Image Analysis* **13** 640-649
[27] Chang C L and Lee R C T 1973 *Symbolic logic and mechanical theorem proving* (New York: Academic Press)
[28] Morozov A A, Sushkova O S and Polupanov A F 2015 A translator of Actor Prolog to Java *RuleML DC and Challenge*
[29] Cristianini N and Shawe-Taylor J 2000 *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods* (Cambridge: Cambridge University Press)
[30] Jang J S R 1993 *IEEE Transactions on Systems, Man, and Cybernetics* **23** 665-685

Acknowledgments
Authors are grateful to Natalia V. Gulyaeva, Ilya G. Komoltsev, Anna O. Manolova, Margarita R. Novikova, and Irina P. Levshina (IHNA and NPh RAS) for the experimental video data, to Yury V. Obukhov (IRE RAS) for the problem statement, and to Alexander F. Polupanov (IRE RAS) for the help in the research. We thank anonymous reviewers for the advertence and useful remarks on the paper. This research is funded by Russian Science Foundation (project No. 16-11-10258).

Greedy algorithms of feature selection for multiclass image classification

E F Goncharova[1] and A V Gaidel[1,2]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086
[2]Image Processing Systems Institute - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001

Abstract. To improve the performance of remote sensing images multiclass classification we propose two greedy algorithms of feature selection. The discriminant analysis criterion and regression coefficients are used as the measure of feature subset effectiveness in the first and second methods respectively. The main benefit of the built algorithms is that they estimate not the individual criterion for each feature, but the general effectiveness of the feature subset. As there is a big limitation on the number of real remote sensing images, available for the analysis, we apply the Markov random model to enlarge the image dataset. As the pattern for image modelling, a random image belonging to one of the 7 classes from the UC Merced Land-Use dataset has been used. Features have been extracted with help of MaZda software. As the result, the largest fraction of correctly classified images accounts for 95%. Dimension of the initial feature space consisting of 218 features has been reduced to 15 features, using the greedy strategy of removing a feature, based on the linear regression model.

1. Introduction

Multiclass or multinomial classification is a significant and complicated step, which can be applied in solving various computer vision tasks. Large number of techniques has been developedto perform the task of multinomial image classification. Some of them apply neural networks, while the others tend to adapt the classical methods of machine learning to improve the quality of the classification results.

In this paper we present two greedy algorithms of feature selection to improve the performance of multiclass image classification. An image itself can be described by various numerical characteristics. For example, the MaZda software for texture analysis [1] estimates almost 300 histogram and texture features, moreover, it includes procedures for their reduction and classification. It should be noticed that not all the extracted features have similar influence on image distinguishing. Redundancy features can affect the performance of classification badly and require additional computational cost.

The feature selection methods have been widely developed in recent years. Some researchers propose feature selection methods based on clustering process. In paper [2] the algorithm is built in the following way: firstly, objects are clustered, than the features which provide the biggest distance between clusters' centroids areappended to the subset of the most informative features. In [3] authors present the novel approach of dimensionality reduction for hyperspectral image classification. To reduce the numbers of variables they use inter band block correlation coefficient technique and QR

decomposition. The support vector machines algorithm has beenapplied to fulfill the classification task. Classification accuracy for images from different databases is between 83 and 99%.

In this work we use MaZda software to extract more than 200 texture features per image. The most informative features are selected with the help of two greedy strategies, based on the discriminant analysis and linear regression model, respectively. The proposed algorithms enable us to select descriptors which have the strongest effect on multinomial image classification. As there is a huge limitation on the number of images, available for the analysis, we also consider the algorithm of image modelling based on the applying of Markov random fields [4].

The experiments are carried out on images belonging to 7 land use classes, 100 for each class, from the UC Merced Land-Use dataset, which provides aerial optical images. To measure the significance of feature subset we estimate the classification error, using *k*-nearest-neighbor scheme. To estimate the effectiveness of image synthesis, we compare the description of the generated and source image in the best feature subset, using the Euclidean distance between two feature vectors.

2. Feature extraction

An image is characterized by its intensity matrix $I^{(M\times N)}$, where $M \times N$ is an image size.

$$I(m,n) = \frac{R(m,n) + G(m,n) + B(m,n)}{3},\ m = \overline{1,M},\ n = \overline{1,N}, \quad (1)$$

R, G, B is an intensity of red, green, and blue component of the image resolution cell having coordinates (m,n) respectively. $I(m,n)$ ranges in value from 0 to $\mathbf{I}-1$, where $\mathbf{I}-$ is a maximum grey level.

To extract the features we compute numerical descriptors of an image, which, eventually, are going to be used to perform the feature selection procedure and further classification. The MaZda software is applied to form the set of features, describing input images [1].

The histogram is calculated via the intensity of each image pixel, calculated by (1), regardless to the spatial correlation among the pixels. The following descriptors are computed: mean intensity, variance, skewness, kurtosis, and percentiles.

The next type of features includes the textural characteristics, calculated with the gray-level spatial dependence matrix. It is built according to the following rule:

$$P_{d_1 d_2}(i,j) = \left\{(m,n) \in \{1,2,...,M\} \times \{1,2,...,N\} \mid I(m,n) = i, I(m+d_1, n+d_2) = j\right\}, i, j = \overline{0, L-1}.$$

Thus, the following features are calculated for five different distances in four directions: angular second moment, contrast, entropy, and correlation.

Features from the other group are calculated based on the autocorrelation function, which describes the dependency between image pixels. The calculated features, as well as the previous ones, are estimated for five different distances in four directions.

3. Methods of feature selection

3.1. Formulatin of feature selection task

The main idea of feature selection process is to improve the classification performance. Thus, let $\mathbf{\Omega}$ be a set of objects for recognition. The set $\mathbf{\Omega}$ is divided into L non-overlapping classes.

To fulfill the classification task we should create the mapping function $\tilde{\Phi}(x)$, which identifies the feature vector x, $x \in \Phi^K$ (K – number of features), with its class. $\tilde{\Phi}(x)$ should be as similar to the ideal mapping function $\Phi(x)$ as possible. $\Phi(x)$ is the ideal mapping function, which is aware of the information about the real object's class.Classification is considered an instance of supervised learning, that is why $\tilde{\Phi}(x)$ is created on the basis of a training set of data $\mathbf{U} \subseteq \mathbf{\Omega}$, containing object with the known class labels.

The aim of feature selection step is to extract the subset of the most informative features, which provides the least classification error.

To classify the feature vectors we apply*k*-nearest-neighbor scheme. According to this method, an object is classified by a majority vote of its neighbors. The classifier assigns the class of the x vector to the class of its *k*-nearest neighbors. The distance between two feature vectors is calculated as the Euclidean distance (2):

$$\rho(x,y)=\sqrt{\sum_{i=1}^{K}(x_i-y_i)^2},\ x\in \mathsf{R}^K ?y\in \mathsf{R}^K, \tag{2}$$

where K is a number of features.

The nearest neighbor error rate is assessed by the formula (3).

$$\varepsilon=\frac{\left|\left\{x\in\tilde{\mathbf{U}}\mid\Phi(x)\neq\tilde{\Phi}(x)\right\}\right|}{\left|\tilde{\mathbf{U}}\right|}, \tag{3}$$

We should notice that $\tilde{\mathbf{U}}$, which is a test set, should be independent of the training set, i.e. $\tilde{\mathbf{U}}\cap\mathbf{U}=\varnothing$. In order to avoid overfitting of classification model and get more accurate results, the leave-one-out cross-validation technique is applied.

Normalization of data is a crucial step of classification process. As different features can be measured in varied scales they affect the classification performance differently. To avoid this problem, all the features in dataset should be standardized. Therefore, the feature vectors get zero mean and unit variance. To achieve this goal we should estimate the expected value $\overline{x}(i)$ and variance $\sigma_{x(i)}$ for each feature.

$$\overline{x}(i)=\frac{1}{|\mathbf{\Omega}|}\sum_{x\in\mathbf{\Omega}}x(i),\ \overline{x}(i)\in\square\ ,$$

$$\sigma_{x(i)}=\frac{1}{|\mathbf{\Omega}|}\sum_{x\in\mathbf{\Omega}}\left(x(i)-\overline{x}(i)\right)^2,\ \sigma_{x(i)}\in\square\ .$$

Thus, each feature can be standardized by applying formula (4).

$$\forall x\in\mathbf{\Omega}\quad x(i)=\frac{x(i)-\overline{x}(i)}{\sqrt{\sigma_{x(i)}}},\ i=\overline{1,K}\ . \tag{4}$$

3.2. Greedy adding algorithm based on the discriminant analysis

When we have several classes, feature selection aims on choosing the features which provide the strongest class separability. In discriminant analysis theory the criterion of separability is evaluated using within-class, between-class, and mixture scatter matrices. Let x be a random vector, belonging to the feature space. Therefore, to measure the importance of the current feature space we should evaluate the degree of isolation of vectors, belonging to different classes [5].

The feature selection method based on the discriminant analysis criterion was proposed in paper [6]. There the separability of two classes was assessed with help of the discriminant criterion[7]. In this work we generalize that technique to the case of several classes.

A within-class scatter matrix (5) shows the scatter of points around their respective class mean vectors (6), and is calculated as follows:

$$\overline{x}_j=\frac{1}{\left|\mathbf{U}\cap\mathbf{\Omega}_j\right|}\sum_{x\in\mathbf{U}\cap\mathbf{\Omega}_j}x,\ \overline{x}_j\in\square^{|\mathbf{Q}|}. \tag{5}$$

$$R_j=\frac{1}{\left|\mathbf{U}\cap\mathbf{\Omega}_j\right|}\sum_{x\in\mathbf{U}\cap\mathbf{\Omega}_j}\left(x-\overline{x}_j\right)\left(x-\overline{x}_j\right)^T,\ R_j\in\square^{|\mathbf{Q}|\times|\mathbf{Q}|}. \tag{6}$$

Prior probability of class $\mathbf{\Omega}_j$ is expressed by $\mathrm{P}\left(\mathbf{\Omega}_j\right)=\dfrac{\left|\mathbf{U}\cap\mathbf{\Omega}_j\right|}{\left|\mathbf{U}\right|}$. The mixture scatter matrix is the covariance matrix of all samples among all the classes, it is defined by:

$$R_{mix}=\frac{1}{\left|\mathbf{U}\right|}\sum_{x\in\mathbf{U}}\left(x-\overline{x}_{mix}\right)\left(x-\overline{x}_{mix}\right)^T,\ R_{mix}\in\mathbb{R}^{|\mathbf{Q}|\times|\mathbf{Q}|},$$

where $\overline{x}_{mix}=\sum_{i=0}^{L-1}\overline{x}_i P\left(\mathbf{\Omega}_i\right),\ \overline{x}_{mix}\in\mathbb{R}^{|\mathbf{Q}|}$ is a mean vector of mixture distribution.

Thus, the discriminant criterion is formulated as

$$J\left(\mathbf{Q}\right)=\frac{\operatorname{tr}R}{\sum_{l=0}^{L-1}\mathrm{P}\left(\mathbf{\Omega}_l\right)\operatorname{tr}R_l}.$$

Criterion $J\left(\mathbf{Q}\right)$ tries to assess the influence of feature set $\mathbf{Q}$ on the within-class compactness and inter-class separability.

To select the most informative features we propose greedy adding strategy. On the first step of the algorithm current set of features is empty $\mathbf{Q}_{(0)}=\varnothing$. On the step i, we observe all the sets, formed as follows $\mathbf{Q}_{(i,j)}=\mathbf{Q}_{(i-1)}\cup\left\{j\right\}$, and calculate the criterion $J_{ij}=J\left(\mathbf{Q}\right)$. We choose the feature subset which provide the maximum value of criterion J_{ij}:

$$\mathbf{Q}_{(i)}=\mathbf{Q}_{(i-1)}\cup\left\{\underset{j\in[1;K]\cap\mathbf{Z}\setminus\mathbf{Q}_{(i-1)}}{\arg\max}\ J_{i,j}\right\}=\mathbf{Q}_{(i-1)}\cup\left\{\underset{j\in[1;K]\cap\mathbf{Z}\setminus\mathbf{Q}_{(i-1)}}{\arg\max}\ J\left(\mathbf{Q}_{(i-1)}\cup\left\{j\right\}\right)\right\}.$$

Then the above steps are repeated until we get the required number of features.

3.3. Greedy algorithm of feature removing based on the regression model

The second algorithm develops the method, examining in paper [6]. The regression analysis studies the relationship between the output (dependent) variable and one, or more, independent descriptors. For the binary classification the number of class can be considered as the dependent variable, which is influenced by feature vector. In the case of multinomial classification we cannot use the number of class as an output, thus we present the function $\Psi^l\left(x\right):\mathbf{\Xi}\to\left[0;1\right]\cap\mathbf{Z}$, which determines whether the feature belongs to the class l or not. The function is defined by:

$$\Psi^l\left(x\right)=\begin{cases}1, y\left(x\right)=l,\\ 0, y\left(x\right)\neq l.\end{cases}$$

Thereby, $\Psi^l\left(x\right)$ is a dependent variable, which is affected by the feature vector $x\in\mathbf{\Xi}\left(\mathbf{Q}\right)$. To assess the degree of feature vector influence we should build L linear regression equations:

$$\Psi^l=X\theta^l+\varepsilon^l,\ l=\overline{0,L-1},$$

where $\Psi^l=(\Psi_1^l\quad\Psi_2^l\quad\ldots\quad\Psi_n^l)^T$ – the output vector; X – "object-feature" matrix; $\theta^l=\left(\theta_0^l\quad\theta_1^l\quad\ldots\quad\theta_{|\mathbf{Q}|}^l\right)^T$ – regression weights; $\varepsilon^l=\left(\varepsilon_1^l\quad\varepsilon_2^l\quad\ldots\quad\varepsilon_n^l\right)^T$ – error vector.

The unknown parameters are estimated by applying the method of least squares:

$$\left(\Psi^l-X\theta^l\right)^T\left(\Psi^l-X\theta^l\right)\to\min_{\theta^l}.$$

Therefore, L vectors θ^l which characterize the coefficients in linear regression are found for each of L classes. Vector $\hat{\theta}=\left(\hat{\theta}_1\quad\hat{\theta}_2\quad\ldots\quad\hat{\theta}_{|\mathbf{Q}|}\right)^T$ is expressed by:

$$\hat{\theta}_i = \sum_{l=0}^{L-1} (\theta_i^l)^2,\ i = \overline{1, |\mathbf{Q}|}. \quad (7)$$

The measure of the influence of each feature is evaluated according to the vector $\hat{\theta}$ element. To select the most informative features we propose greedy removing strategy. The initial feature subset includes all the features $\mathbf{Q}_{(0)} = \mathbf{Q}$. Than we sequentially remove the worst feature from the current subset and rebuild the linear regression as follows: on the step j of the algorithm we create L linear regression models $\tilde{\Psi}^l_{(j)} = X^l_{(j)} \theta^l_{(j)}$, then vector $\hat{\theta}_{(j)}$ is calculated for the current feature subset $\mathbf{Q}_{(j)}$. The feature with the minimal value of $\hat{\theta}_{(j)}(k)$ is removed from the current subset:

$$\mathbf{Q}_{(j+1)} = \mathbf{Q}_{(j)} \setminus \left\{ \arg\min_{k \in [1;K] \cap \mathbf{Z} \cap \mathbf{Q}_{(j)}} \left| \hat{\theta}_{(j)}(k) \right| \right\}.$$

The steps of the algorithm are iterated until a required number of features is obtained.

4. Image modelling

4.1. Markov random fields

Imagemodellingisperformedwithhelpof Markovrandomfields. Let $F = \{F_i \mid i \in S\}$ be a multivariate random variable, which is defined on the discrete set of index $S = \{1, 2, ..., N\}$. F_i is a random variable that takes values $\{F_1 = f_1, F_2 = f_2, ..., F_n = f_n\}$. The probability of random variable F_i taking the value of f_i is denoted as $P(f_i)$. Thus, F is a random field.

Configuration $f = (f_1, f_2, ..., f_n)$ is a specific realization of random variable F. Let $\mathrm{N} = \{\mathrm{N}_i \mid \forall i \in S\}$ be a neighborhood system, where N_i – set of elements neigh bouring i. Thus, the nodes that influence the local characteristics of the i-node are included in its neighbourhood system.

Markov random fields satisfy the following formula (8):

$$\forall f \ P(F_i = f_i \mid F_j = f_j, i \neq j) = P(F_i = f_i \mid F_j = f_j, j \in \mathrm{N}_i). \quad (8)$$

Hence, Markov random fields imply conditional independence [8]. According to (8) F_i only depends on the nodes included in its neighbourhood N_i. Thus, if nodes in the neighbourhood are known than the values of F_j for $j \neq i$ and $j \notin \mathrm{N}_i$ do not affect F_i [9].

4.2. Image modelling

Suppose that the set of indices S defines the set of points on the 2D plane. The discrete image is a realization of 2D random variable F, defined in the points S. Following by the conditional independence of F, wecanassumethattheintensityvalueofeachimagepixel can be predicted on the basis of several nodes, included in its neighbourhood.

Thus, we can present the following strategy of image modelling using the Markov random fields. On the each step k of the algorithm the neighbourhood system N_i is created for i pixel of the image $G_{(k)}(i)$. Than this neighbourhood system is compared with the neighbourhood of the correspond pixel, belonging to the input image $G_{in}(i)$. $G_{in}(i)$ is a sample real image for synthesis.The pixel's values are set as follows:

$$G_{(k+1)}(i) = G_{in}\left(\arg\min_{i_* \in S} \left(\rho\left(\mathrm{N}_{i_*}(G_{in}), \mathrm{N}_i\left(G_{(k)}\right) \right) \right) \right).$$

The distance $\rho(x, y)$ is defined by formula (2). The initial image $G_{(0)}$ is approximated by the white noise.

In this paper we propose to use causal 5-neighbourhood system. This neigh bourhood pattern is shown in figure 1. The peculiarity of this type of neigh bourhood is that it contains only those nodes that precede the current output pixel. That means that $\mathrm{N}_i\left(G_{(k)}\right)$ includes already assigned pixels.

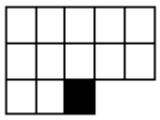


Figure 1. The instance of causal 5-neighbourhood system. The qurrently processing pixelis marked byblacksquare.

5. Experimental results

5.1. Experiments of feature selection

The experiments were carried out on the images from the UC Merced Land-Use dataset, which consists of the aerial optical images, belonging to different classes (agricultural field, forest, beach, etc.), 100 for each class. Each image measures 256×256 pixels. In this work we analyzed images, belonging to 7 classes (agricultural field, forest, buildings, beach, golf course, chaparral, and freeway). figure 2 illustrates sample images belonging to the mentioned above classes.

To get the correct classification results the Leave-one-out cross-validation technique was applied. The total number of features, extracted with the MaZda accounts for 218.

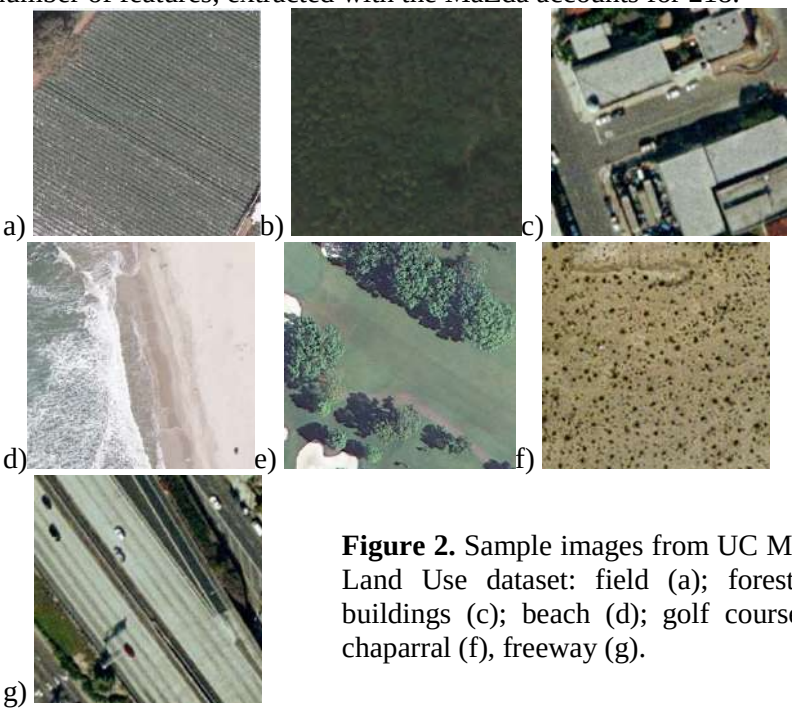


Figure 2. Sample images from UC Merced Land Use dataset: field (a); forest (b); buildings (c); beach (d); golf course (e), chaparral (f), freeway (g).

The results obtained with the discriminant and regression analysis methods are shown in table 1.

The most informative groups of features, selected with the two proposed strategies, along with the classification error (3), obtained on these groups, are presented in tables 2 and 3.

Having analyzed the results, we can conclude that the greedy removing algorithm, based on the linear regression model, performed best on this multinomial classification task. The lowest classification error rate of 0.05 was achieved in feature space, consisting of the 15 features from the 218 initial.

Table 1. The features selected with the greedy algorithms, based on discriminant and regression analysis respectively, in descending order of priority.

Discriminant analysis		Regression analysis	
Feature number	Feature name	Feature number	Feature name
37	S11SumVarnc	96	S202DifEntrp
30	S01DifEntrp	74	S02DifEntrp
24	S01InvDfMom	85	S22DifEntrp
40	S11DifVarnc	107	S30DifEntrp
…	…	…	…
79	S22InvDfMom	171	S44Entropy
34	S11SumOfSqs	215	S55Entropy
32	S11Contrast	217	S55DifEntrp

Table 2. Groups of the most informative features, selected with the discriminant analysis.

K	Features	ε
3	37, 30, 24	0.74
4	37, 30, 24, 40	0.47
5	37, 30, 24, 40, 38	0.63
6	37, 30, 24, 40, 38, 2	0.63
…	…	…
47	37, 30, 24, 40, 38, 2, 55, 13, 42, 209, 44, …, 91, 69, 80, 16, 118, 85	0.32
48	37, 30, 24, 40, 38, 2, 55, 13, 42, 209, 44, …, 91, 69, 80, 16, 118, 85, 5	0.32

Table 3. Groups of the most informative features, selected with the regression analysis.

K	Features	ε
3	96, 74, 85	0.16
4	96, 74, 85, 107	0.32
5	96, 74, 85, 107, 140	0.42
6	96, 74, 85, 107, 140, 63	0.16
…	…	…
15	96, 74, 85, 107, 140, 63, 151, 129, 173, 118, 110, 154, 66, 99, 143	0.05
16	96, 74, 85, 107, 140, 63, 151, 129, 173, 118, 110, 154, 66, 99, 143, 138	0.05

The best group includes various textural features, extracted for 4 dimensions: 2, 3, 4 and 5. The greedy adding algorithm maximized the discriminant analysis criterion provided worse results. The lowest classification error rate of 0.32 was achieved on the set, consisting of 47 features. We should notice that the fracture of the images that were classified correctly in the whole space of 218 features accounts for 63%. That means that both analyzed techniques have succeeded in dimension reduction and improving classification performance.

5.2. Experiments of image modelling

To carry out the experiments of image synthesis we have presented the initial images the UC Merced Land-Use dataset in the greyscale. The results of modelling are shown in figure 3.To check the quality of synthesized images we performed the comparison of the feature vectors for the input sample and the obtained image. The vectors in clude 15 best features, selected by the greedy removing algorithm. The measure of equality $\xi(x,y)$, $x \in \mathsf{R}^K$, $y \in \mathsf{R}^K$ isexpressedby

$$\xi(x,y) = 1 - \left(\frac{1}{K} \sum_{k=1}^{K} (x_k - y_k)^2 \right)^{\frac{1}{2}}.$$

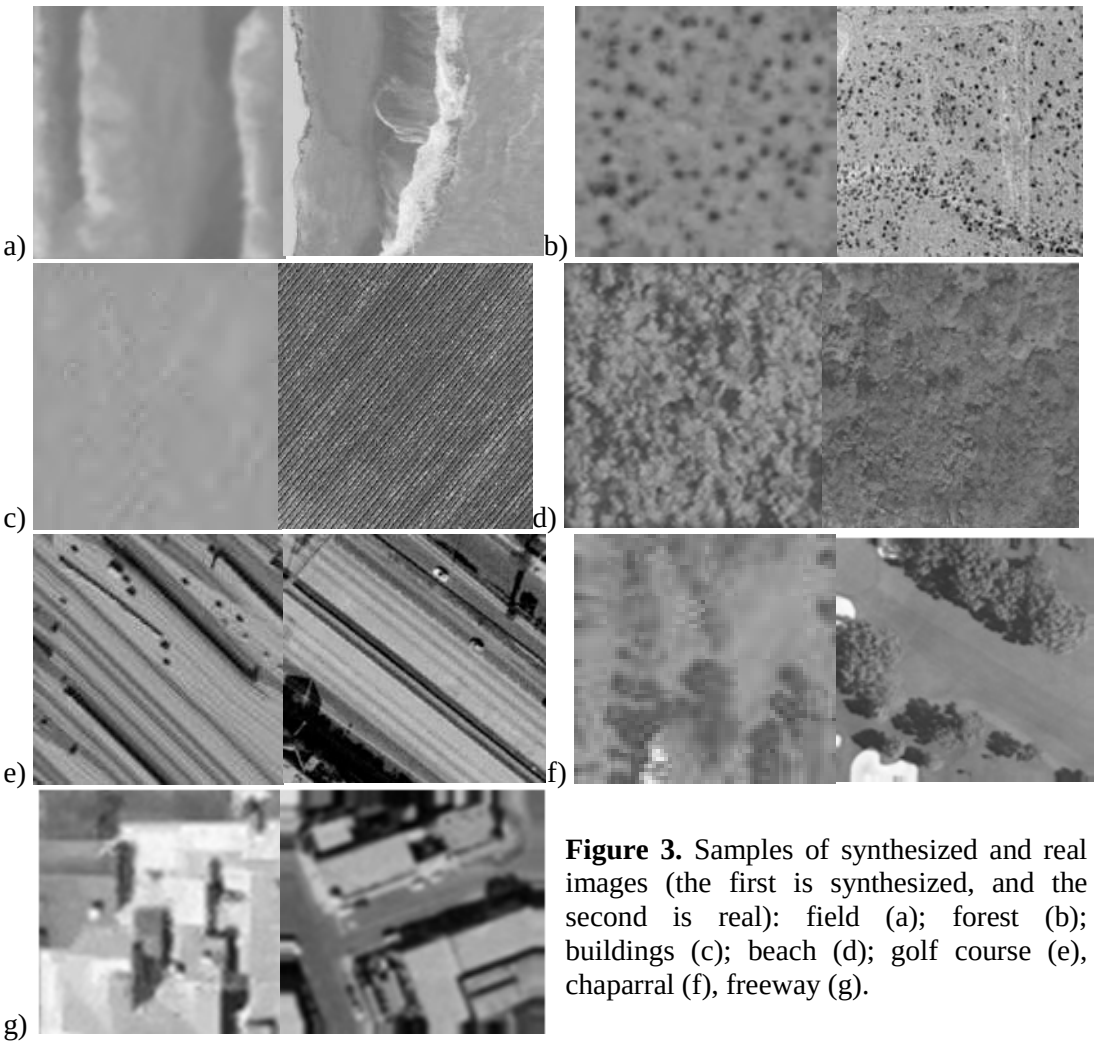


Figure 3. Samples of synthesized and real images (the first is synthesized, and the second is real): field (a); forest (b); buildings (c); beach (d); golf course (e), chaparral (f), freeway (g).

Table 4 presents the value of $\xi(x, y)$ for the images synthesized for 7 classes.

Table 4. Measure of equality for the synthesized images.

Class	$\xi(x, y)$
Beach	0.15
Chaparral	0.08
Field	0.14

Continuation of table 4

Forest	0.09
Freeway	0.09
Golf course	0.09
Buildings	0.13

The results shown in table 4 prove that the proposed method performs successfully for the images with small scale structure. For example, synthesized images belonging to the classes: chaparral and field, turned to be quite similar to the real images. However the quality of synthesized images containing large scale structure is lower. Its modelling demands large neighborhoods which leads to the increasing computational cost. To solve this problem, method, based on the multiresolution image

pyramids, is proposed in paper [4]. In that methodcomputation is saved because the large scale structures are presented more compactly by a few pixels in a certain lower resolution pyramid level.

6. Conclusion

Thus, for the task of the remote sensing images classification the subset of informative features was extracted. We proposed two greedy strategies for informative feature selection. The feature vector, selected with the greedy removing algorithm, based on building the regression model, produced the best classification performance (using the nearest-neighbor classification method) on the images from the UC Merced Land Use dataset. The minimal classification error rate made up 0.05. In comparison to that, the greedy adding algorithm maximized the discriminant analysis criterion provided worse results. The lowest classification error rate of 0.32 was achieved on the set, consisting of 47 features. We should notice that the fracture of the images that were classified correctly in the whole space of 218 features accounted for 63%.

Overall, applying the feature selection methods leads to improving the multinomial image classification performance and dimension reduction. Using only 15 (of 218 initial) descriptors allows to classify 95% of images correctly

To increase the number of images, available for analysis, we applied the algorithm of image modelling on the basis of Markov random fields. The experimental results showed that this technique can be applied for synthesis images with the low scale structure. To generate samples, containing large scale structure, the proposed algorithm should be adopted. One of the possible variants is to apply multiresolution image pyramids.

7. References

[1] Strzelecki M A, Szczypinski P, Materka A and Klepaczko A 2013 A software tool for automatic classification and segmentation of 2D/3D medical images *Nuclear Instruments and Methods in Physics Research* **702** 137-140

[2] Liu C, Wanga W, Zhao Q, Shen X and Konan M 2017 A new feature selection method based on a validity index of feature subset *Pattern Recognition. Leters.* **92** 1-8

[3] Reshma R, Sowmya V and Soman K P 2016 Dimensionality Reduction Using Band Selection Technique for Kernel Based Hyperspectral Image Classification *Proc. Computer Science* **93** 396-402

[4] Wei L Y and Levoy M 2000 Fast texture synthesis using tree-structured vector quantization *Proc. of the 27th annual conf. on Computer graphics and interactive techniques* 479-488

[5] Gaidel A V 2015 A method for adjusting directed texture features in biomedical image analysisproblems *Computer Optics* **39(2)** 287-293 DOI: 10.18287/0134-2452-2015-39-2-287-293

[6] Goncharova E and Gaidel A 2017 Feature Selection Methods for Remote Sensing Images Classification *3rd Int. conf. Information Technology and Nanotechnology* 535-540

[7] Kutikova V V and Gaidel A V 2015 Study of informative feature selection approaches for the texture image recognition problem using Laws' masks *Computer Optics* **39(5)** 744-750 DOI: 10.18287/0134-2452-2015-39-5-744-750

[8] Winkler G 2012 *Image Analysis, Random Fields and Dynamic Monte Carlo Methods* (Springer-Verlag) p 387

[9] Li S Z 2009 *Markov random field modeling in image analysis*(Springer-Verlag) p 356

Acknowledgments

This work was supported by the Federal Agency for Scientific Organizations under agreement No. 007-GZ/Ch3363/26.

Multidimensional signals interpolation based on NEDI for HGI compression

M V Gashnikov[1]

[1]Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. Adaptive interpolation of multidimensional digital signals is considered. An adaptive algorithm for digital signals interpolation is proposed, intended for hierarchical compression. The prototype of the proposed interpolator is the NEDI (New Edge-Directed Interpolation) algorithm. In this paper, the NEDI interpolation algorithm is modified for use on special hierarchical grids, which are used for hierarchical signal compression. Experimental researches of the proposed interpolator are performed with the hierarchical compression of natural digital signals. Experiments confirm that the proposed adaptive interpolator allows improving the efficiency of hierarchical compression of digital signals.

1. Introduction

Availability of digital information processing devices continues to increase. This entails an increase in the data size of processed digital signals, and this problem can not be solved by increasing capacity of storage devices. Moreover, multidimensional signals, including multi- and hyperspectral [1-3] remote sensing data, as well as results of sensing by quadrocopters and other unmanned aerial vehicles, are also becoming more accessible. This further exacerbates the problem of an excessively large size of digital signal data. The only acceptable solution at the moment is compression of digital signals [4-5].

To date, there are many [4-8] methods of compression of digital signals. The most popular of these methods is the JPEG compression method [11], based on discrete cosine transform (DCT) [9] and subsequent entropy coding [10] of transformants (DCT results). The more efficient [12] compression method JPEG-2000 [13], which uses the discrete Wavelet transform [14], is much less widely used.

These methods of the JPEG group are used very widely, due to wide variety of hardware devices in which they are embedded. However, there are a number of problems that raise requirements for the quality of compressed digital data. First of all, this is polygraphy and processing of remote sensing data. In these areas, one has to deal with digital signals, which are unique. When compressing such signals, strict quality control is necessary. In addition, such signals may also have a high bit capacity. Moreover, such signals can have substantially more than three spectral bands (hyperspectral signals often have hundreds of spectral bands). In other words, when compressing such signals, complexity can arise already at the stage of processing data of specific formats.

Fractal [15] compression methods, according to the author, currently have the largest compression ratio. However, their propagation is difficult due to specific, in most cases unacceptable signal distortions, as well as excessively high computational complexity.

Also, it should be noted an important drawback, corresponding to all the above methods of signal compression. This drawback follows from the need to transform the signal into a corresponding space of transformation coefficients. Accordingly, it is not always possible to control the error in the specified space of coefficients. For the mean-square error, such control is possible in a number of

cases due to Parseval's equality. But for more strong quality measures, for example for maximum error, the specified error control for the above compression methods is usually impossible.

In the author's opinion, using specific compression methods that do not require the transformation to spectral (or any other) auxiliary spaces is promising in specific areas that raise high demands to the quality of digital signals. In this paper, the method of hierarchical compression is chosen as such method [16, 17]. This method is based on multiple non-redundant resampling of initial array of signal samples and interpolation of signal samples based on the specified resampled arrays.

Hierarchical compression methods have a number of important advantages, such as fast multiscale access to fragments of compressed data, the ability to control the speed of formation of a compressed data stream, the possibility of increasing noise immunity and the possibility of error control (including the maximum error [18]). The task of research and further increasing the efficiency of hierarchical compression methods of digital signals is certainly topical.

An important step in hierarchical compression methods is an interpolator in which samples of more resampled signal are used to interpolate samples of less resampled signal. The most common algorithm of hierarchical interpolation is simple averaging [19-20] from the nearest signal samples of more resampled hierarchical levels of the signal. However, the averaging interpolator is not effective enough, because it is not adaptive (it performs in the same way, regardless of local signal characteristics).

One of the ways to take into account the local characteristics of a digital signal is context modeling [21-23], which has become widespread, in particular, in statistical coding [7]. In the simplest case, the context for a next encoded symbol is the previous symbol (or several previous symbols), and the context model is the estimation of conditional probability distribution of the encoded symbol. Taking into account the context, that is, using the conditional probability distribution instead of the unconditioned distribution makes it possible to increase the algorithm adaptability to variable statistical properties of the signal, which leads to an increase in the efficiency of the compression method as a whole.

In this paper, context modeling is used for the development of adaptive interpolation algorithms that are part of a compression method based on hierarchical grid interpolation (HGI). The proposed adaptive interpolators allow increasing the efficiency of the hierarchical compression method.

For the hierarchical method of signal compression, an interpolator based on the NEDI algorithm [24] using context modeling is proposed. When developing this interpolation algorithm, a set of surrounding signal samples is considered as a context for each signal sample.

2. Hierarchical compression of multidimensional signals

Hierarchical grid interpolation (HGI) [16, 25-26] is based on special hierarchical representation of an integer nonnegative multidimensional signal $\mathbf{X}=\left\{x\left(\overset{\mathrm{r}}{c}\right)\right\}$ in the form of a set of hierarchical levels $\mathbf{X}_l$:

$$\mathbf{X}=\bigcup_{l=0}^{L-1}\mathbf{X}_l\ ,\quad \mathbf{X}_{L-1}=\left\{x_{L-1}\left(\overset{\mathrm{r}}{c}\right)\right\},\quad \mathbf{X}_l=\left\{x_l\left(\overset{\mathrm{r}}{c}\right)\right\}\setminus\left\{x_{l+1}\left(\overset{\mathrm{r}}{c}\right)\right\},\ l<L-1,$$

where L is the number of hierarchical levels $\mathbf{X}_l$, $\{x_l\left(\overset{\mathrm{r}}{c}\right)\}$ is the signal resampled with step 2^l, $\overset{\mathrm{r}}{c}$ is the vector of multidimensional signal arguments.

With hierarchical compression, the hierarchical levels $\mathbf{X}_l$ are compressed sequentially, from the highest (most resampled) level $\mathbf{X}_{L-1}$ to the lower levels. The proportion of data size of the highest level $\mathbf{X}_{L-1}$ is sufficiently small already for $L\geq 4$, so the compression algorithm of this level does not matter. So, only compression algorithm of any "non highest" hierarchical level $\mathbf{X}_l$, $l<L-1$ is described.

Stage No. 1. Interpolation.

Interpolation of samples of the current signal level $\mathbf{X}_l$ is based on the samples $\{\bar{x}_k\left(m,n\right),k>l\}$ of already processed hierarchical levels $\{\overline{\mathbf{X}}_k,k>l\}$:

$$\hat{x}_l(\vec{c}) = P\left(\bigcup_{k=l+1}^{L-1} \overline{X}_k\right), \quad (1)$$

where $\hat{x}_l(\vec{c})$ is interpolating value, $P(..)$ is a function that defines a certain interpolator. Further, interpolators are considered in detail.

Stage No. 2. Calculation of difference signal.

The differences between the initial $x_l(\vec{c})$ and interpolating $\hat{x}_l(\vec{c})$ (1) values of the current level samples are calculated:

$$f_l(\vec{c}) = x_l(\vec{c}) - \hat{x}_l(\vec{c}). \quad (2)$$

Stage No. 3. Quantization.

The difference signal (2) is quantized by the quantifier with a uniform scale ([..] is the integer part of a number):

$$q_l(\vec{c}) = sign\left(f_l(\vec{c})\right)\left[\left(\left|f_l(\vec{c})\right| + \varepsilon_{\max}\right) / \left(2\varepsilon_{\max} + 1\right)\right], \quad (3)$$

allowing to control the maximum error [19] $\varepsilon_{\max}$:

$$\left|f_l(\vec{c})\right| = \left|x(\vec{c}) - \overline{x}(\vec{c})\right| \le \varepsilon_{\max}. \quad (4)$$

The quantized signal (3) is then compressed by an entropy encoder and stored in an archive.

Stage No. 4. Recovering.

The restored values of the signal samples are calculated (already during compression):

$$\overline{x}_l(\vec{c}) = \hat{x}_l(\vec{c}) + (1 + 2\varepsilon_{\max}) q_l(\vec{c}), \quad (5)$$

which are necessary for interpolation (1) of the following (more resampled) hierarchical levels $\{\mathbf{X}_k, k < l\}$ of signal.

3. Averaging interpolation for hierarchical signal compression

For reasons of computational complexity of interpolation under hierarchical compression, we usually [19-20] use trivial averaging over the nearest already restored signal samples of more resampled hierarchical levels. To simplify the algorithm description, the averaging interpolator is considered for a two-dimensional signal $\mathbf{X} = \left\{x(\vec{c})\right\} = \left\{x(m,n)\right\}$.

Let's consider two types of interpolated samples: "type I" with indices of the form $(2m+1, 2n+1)$ and "type II" with indices of the form $x_l(2m+1, 2n)$ and $x_l(2m, 2n+1)$. Let's consider the simplest way of processing samples of both samples types when using the averaging interpolator. Interpolation is performed based on the restored samples of previous (more resampled) hierarchical level:

$$\hat{x}_l(2m+1, 2n) = \left[\left(\overline{x}_{l+1}(m,n) + \overline{x}_{l+1}(m+1,n)\right)/2\right], \quad (6)$$

$$\hat{x}_l(2m+1, 2n+1) = \left[\left(1 + \overline{x}_{l+1}(m,n) + \overline{x}_{l+1}(m+1,n) + \overline{x}_{l+1}(m,n+1) + \overline{x}_{l+1}(m+1,n+1)\right)/4\right] \quad (7)$$

This averaging interpolation algorithm has low computational complexity. But this algorithm has insufficiently high efficiency, since it does not take into account any local signal characteristics.

4. Context modeling for data compression

Application of context modeling approaches [21-23] for data compression is based on the «modeling and coding» idea proposed by Rissanen and Langdon [21]. In accordance with this idea, the compression procedure consists of two steps: modeling and coding.

By modeling, we mean the construction of model of information source that generates compressed data. By coding, we mean the process of reducing data size based on the results of modeling. So, the "coder" creates a compressed stream, which is a compact form of the processed data, based on the information supplied by the "modeler".

Let the probability of the symbol s_i be $p(s_i)$. From Shannon's theorem [7] on coding the source of messages, it is known that the symbol s_i is best represented by $-\log_2 p(s_i)$ bits. Often the source structure is unknown or complex, so it is necessary to build a source model that allows you to find an estimate of the probability $\overline{p}(s_i)$ of each symbol s_i.

Estimation of the symbol probability during the simulation is performed on the basis of known (or estimated) symbols statistics and, possibly, a priori assumptions. Therefore they speak about statistical modeling. In other words, the "modeler" predicts the probability of each symbol.

At the stage of statistical coding, the symbol s_i is replaced by a code with a length of $-\log_2 p(s_i)$ bits. The more accurate the estimation of the signals probability, the more effective the codes are, and the more effective the compression as a whole.

5. Signal interpolation based on NEDI

In the initial setting, context modeling is used for entropy coding, but we use context modeling for interpolating signals. In particular, in terms of context modeling, the New Edge-Directed Interpolation (NEDI) [24] can be described.

Let's describe the interpolation algorithm NEDI. Let $\mathbf{X}=\{x(m,n)\}$ be the original signal of size *WxH*. We need to calculate a signal $Z=\{x(m,n)\}$ of size 2*W*x2*H* with twice the best resolution. If we use the NEDI, this signal is calculated as follows. Signal samples with even indices are equal to samples of the original signal:

$$z(2m,2n)=x(m,n)\,. \tag{8}$$

The signal samples with indices $(2m+1,2n+1)$ are calculated (see Figure 1) as follows:

$$z(2m+1,2n+1)=\sum_{i=0}^{1}\sum_{j=0}^{1}\alpha_{2i+j}x(m+i,n+j)\,, \tag{9}$$

where $\alpha_i, i=0..3$ are local interpolation coefficients.

So the missing signal samples are calculated as a weighted sum of the four nearest reference signal samples. In this case, it is necessary to calculate the optimal local weighting coefficients $\alpha_i, i=0..3$. These coefficients are calculated by the optimization of the quadratic error:

$$\varepsilon^2=\sum_{i,j\in D}(x(i,j)-z(i,j))^2\to\min_{\alpha_0,\alpha_1,\alpha_2,\alpha_3}, \tag{10}$$

where *D* is area of interpolation coefficients estimation. This area includes *N* signal samples.

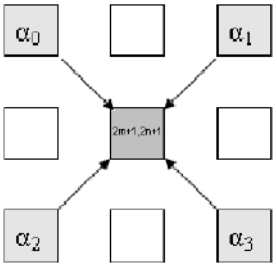


Figure 1. Location of reference signal samples for NEDI interpolation

We place the samples $x(i,j)$ from the estimation area *D* into the array $\overset{1}{Y}=[y_0,y_1,...,y_{N-1}]^T$. The elements of the array $\overset{1}{Y}$ are assigned to signal samples from the estimation area *D*, starting from the upper left corner and then according to the progressive scan of area (see Figure 2). Also consider the matrix

$$C=\begin{bmatrix} c_{0,0} & \dots & c_{0,3} \\ \dots & \dots & \dots \\ c_{N-1,0} & \dots & c_{N-1,3} \end{bmatrix}. \tag{11}$$

Each row of this matrix consists of reference signal samples, which are used to interpolation of the samples of the estimation area (see Figure 2).

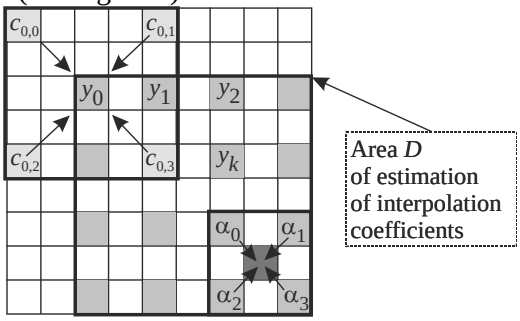


Figure 2. Estimation area D of the interpolation coefficients of the NEDI algorithm.

So criterion (10) is represented as:

$$\varepsilon^2 = \sum_{k \in D} (y_k - C_k \vec{\alpha})^2 \to \min_{\vec{\alpha}}, \tag{12}$$

where C_k is the row number k of the matrix C, $\vec{\alpha} = [\alpha_0, \alpha_1, \alpha_2, \alpha_3]^T$ is the vector of unknown interpolation coefficients.

To solve the optimization problem (12), we equate to zero the following partial derivatives:

$$\begin{cases} \frac{\partial \varepsilon^2}{\partial \alpha_0} = (y_0 - (\alpha_0 c_{0,0} + \alpha_1 c_{0,1} + \alpha_2 c_{0,2} + \alpha_3 c_{0,3})) \cdot c_{0,0} + \ldots + (y_{N-1} - (\alpha_0 c_{N-1,0} + \alpha_1 c_{N-1,1} + \alpha_2 c_{N-1,2} + \alpha_3 c_{N-1,3})) \cdot c_{N-1,0} = 0 \\ \frac{\partial \varepsilon^2}{\partial \alpha_1} = (y_0 - (\alpha_0 c_{0,0} + \alpha_1 c_{0,1} + \alpha_2 c_{0,2} + \alpha_3 c_{0,3})) \cdot c_{0,1} + \ldots + (y_{N-1} - (\alpha_0 c_{N-1,0} + \alpha_1 c_{N-1,1} + \alpha_2 c_{N-1,2} + \alpha_3 c_{N-1,3})) \cdot c_{N-1,1} = 0 \\ \frac{\partial \varepsilon^2}{\partial \alpha_2} = (y_0 - (\alpha_0 c_{0,0} + \alpha_1 c_{0,1} + \alpha_2 c_{0,2} + \alpha_3 c_{0,3})) \cdot c_{0,2} + \ldots + (y_{N-1} - (\alpha_0 c_{N-1,0} + \alpha_1 c_{N-1,1} + \alpha_2 c_{N-1,2} + \alpha_3 c_{N-1,3})) \cdot c_{N-1,2} = 0 \\ \frac{\partial \varepsilon^2}{\partial \alpha_3} = (y_0 - (\alpha_0 c_{0,0} + \alpha_1 c_{0,1} + \alpha_2 c_{0,2} + \alpha_3 c_{0,3})) \cdot c_{0,3} + \ldots + (y_{N-1} - (\alpha_0 c_{N-1,0} + \alpha_1 c_{N-1,1} + \alpha_2 c_{N-1,2} + \alpha_3 c_{N-1,3})) \cdot c_{N-1,3} = 0 \end{cases}$$

The solution of this system of equations:

$$\vec{\alpha} = (C^T C)^{-1} \cdot (C^T \vec{Y}), \tag{13}$$

This expression allows us to calculate the optimal interpolation coefficients $\alpha_i, i = 0..3$. The advantage of considered context NEDI algorithm is the adaptability to local signal characteristics, the disadvantage is high computational complexity. By context, in this case, we mean the set of surrounding reference samples that constitute the area of parameters estimation. By context modeling, in this case, we mean the local estimation of interpolation coefficients.

6. NEDI-based interpolation algorithm for hierarchical signal compression

In this paper, we propose a modification of the NEDI interpolation algorithm for non-redundant hierarchical sample grids, which are used for hierarchical signal compression. At each "non-highest" hierarchical level $\mathbf{X}_l$, $0 \le l \le L-2$, an estimation area D is chosen for interpolation of each sample of hierarchical level. This area D consists of already restored samples of previous (more resampled) hierarchical levels $\mathbf{X}_k$, $l+1 \le k \le L-1$, which are stored into array $\vec{Y}_l$. The "own" matrix C_l (see the previous section) corresponds to each element of this array.

In Figure 3, for two hierarchical levels, the array's $\vec{Y}_l$ samples are shown in light color, the matrix elements belonging to the hierarchical level $\mathbf{X}_1$ are shown in dark color (the interpolated sample belongs to the hierarchical level $\mathbf{X}_0$).

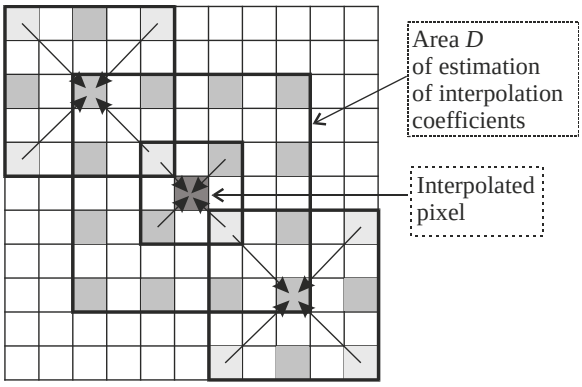


Figure 3. Estimation area *D* of the NEDI interpolation coefficients for hierarchical compression.

Estimation of local interpolation coefficients is performed in the same way as expression (13), taking into account the described specificity of estimation areas:

$$\vec{\alpha}_l = (C_l^T C_l)^{-1} \cdot (C_l^T \vec{Y}_l) . \quad (14)$$

First, the algorithm described above is applied to interpolation of samples with indices of the form (2*i*+1, 2*j*+1), then the interpolation coefficients for the remaining samples are computed (similarly, the entire situation is rotated by 45 degrees).

The advantage of the proposed interpolation algorithm is its adaptability to a context (local characteristics of the signal in a small neighborhood of current sample). This increases the accuracy of interpolation of the signal sample. The disadvantage of the algorithm is high computational complexity of calculating the interpolation coefficients.

It should be noted that it is not always possible to compute an inverse matrix to the matrix $C_l^T C_l$, since matrix $C_l^T C_l$ can be ill-conditioned. A feature of ill-conditioning is the situation when the number of conditionality of the matrix $C_l^T C_l$ exceeds a certain "conditioning threshold" *T* (parameter of proposed interpolation algorithm). In this situation, instead of NEDI interpolation, averaging interpolation (6-7) is used. The "conditionality threshold" *T* affects the efficiency of compression, because this threshold affects the interpolation algorithm. However, the proportion of signal samples for which the situation of ill-conditionality arises is quite small.

7. Experimental research of adaptive interpolation algorithm

We developed a software implementation of the proposed adaptive context interpolator based on NEDI. We built this interpolation algorithm into the hierarchical compression method. To research the effectiveness of the proposed interpolator, computational experiments were performed in natural test signals. Some of these test signals are shown in Figure 4.

As a measure of interpolator efficiency, the relative gain in the archive file size was used, that is achieved through the use of the proposed interpolator (9-14) instead of the averaging (6-7) interpolator within the framework of the hierarchical compression:

$$\Delta = \frac{S^{\text{standart}} - S^{\text{new}}}{S^{\text{standart}}} \cdot 100\% ,$$

where S^{standart}, S^{new} are archive files sizes when signal compressing by the hierarchical method using the averaging and proposed interpolators, respectively.

In each computational experiment, we selected the best (by the compression ratio) threshold values *T* of the condition number of the matrix $C^T C$ for a fixed number *L* of hierarchical levels and fixed size *N* of the estimation area. Some typical results are shown in Figure 5.

From the received experimental results it is clear, that the proposed algorithm provides the gain (up to 2%) on the archive size.

(a) (b)

Figure 4. Examples of natural test signals.

From the results of the computational experiments it can be seen that the proposed adaptive interpolator provides a noticeable (up to 8%) gain on the archive size. This gain increases when the maximum compression error and the size of the interpolation coefficient estimation area are increased.

8. Conclusion

An approach based on context modelling was used to develop an adaptive interpolator for hierarchical signal compression. Based on this approach, a context interpolator based on the NEDI algorithm is proposed. The proposed interpolator is implemented as software and is built into the hierarchical method of signal compression. Computational experiments were conducted to research the effectiveness of the proposed contextual interpolator in natural test signals. It is shown that use of the proposed context interpolator instead of the average interpolator is noticeable (up to 8%) increases the efficiency of hierarchical signal compression.

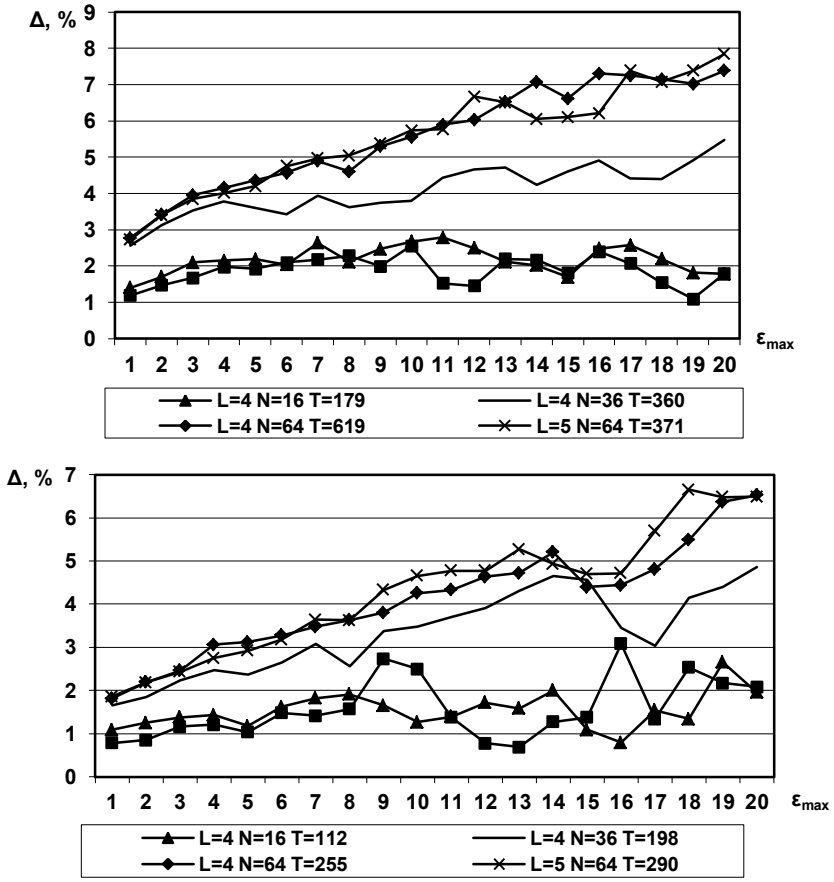


Figure 5. Gain Δ of the proposed context interpolator over the averaging interpolator as a function of the maximum error ε_{max}.

9. References

[1] Chang C 2007 *Hyperspectral data exploitation: theory and applications* (New York: Wiley-Interscience) p 440
[2] Borengasser M 2004 *Hyperspectral Remote Sensing – Principles and Applications* (CRC Press) p 128
[3] Chang C 2013 *Hyperspectral Data Processing: Algorithm Design and Analysis* (New York: Wiley Press) p 1164
[4] Vatolin D, Ratushnyak A, Smirnov M and Yukin V 2002 *Data compression methods. Archive program architecture, image and video compression* (Moscow: Dialog-MIFI Publisher) p 384
[5] Sayood K 2012 *Introduction to Data Compression (*The Morgan Kaufmann Series in Multimedia Information and Systems) p 743
[6] Woods E and Gonzalez R 2007 *Digital Image Processing* (New Jersey: Prentice Hall) p 976
[7] Salomon D 2007 *Data Compression. The Complete Reference* (Springer-Verlag) p 1118
[8] Pratt W 2007 *Digital image processing* (New York: Wiley Press) p 807
[9] Plonka G and Tasche M 2005 Fast and numerically stable algorithms for discrete cosine transforms *Linear Algebra and its Applications* **394(1)** 309-345
[10] Lidovski V 2004 *Information theory: Tutorial* (Moscow: Kompania Sputnik+) p 111
[11] Wallace G 1991 The JPEG Still Picture Compression Standard *Communications of the ACM* **34(4)** 30-44
[12] Ebrahimi F, Chamik M and Winkler S 2004 JPEG vs. JPEG2000: An Objective Comparison of Image Encoding Quality *Proc. of SPIE Applications of Digital Image Processing XXVII* **5558** 300-308
[13] Li J 2003 Image Compression: The Mathematics of JPEG-2000 *Modern Signal Processing* **46** 185-221
[14] Gupta V, Sharma A and Kumar A 2014 Enhanced Image Compression Using Wavelets *International Journal of Research in Engineering and Science (IJRES)* **2** 55-62
[15] Woon W M, Ho A T S, Yu T, Tam S C, Tan S C and Yap L T 2000 Achieving high data compression of self-similar satellite images using fractal *Proc. of IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS)* 609-611
[16] Gashnikov M V, Glumov N I and Sergeyev V V 2000 Compression Method for Real-Time Systems of Remote Sensing *Proc. 15th Int. Conf. on Pattern Recognition* **3** 232-235
[17] Gashnikov M V 2017 Minimizing the entropy of post-interpolation residuals for image compression based on hierarchical grid interpolation *Computer Optics* **41(2)** 266-275 DOI: 10.18287/2412-6179-2017-41-2-266-275
[18] Lin S 2004 *Error Control Coding: Fundamentals and Applications, second edition* (New Jersey: Prentice-Hall) p 1260
[19] Efimov V M and Kolesnikov A N 1997 Effectiveness estimation of the hierarchical and line-by-line lossless compression algorithms *Proc. of the III conf. Pattern recognition and image analisys* **1** 157-161
[20] Gashnikov M V 2016 Interpolation for hyperspectral images compression *CEUR Workshop Proc* **1638** 327-333
[21] Rissanen J J and Langdon G G 1981 Universal modeling and coding *IEEE Transactions on Information Theory* **27(1)** 2-23
[22] Cleary J G and Witten I H 1984 Data compression using adaptive coding and partial string matching *IEEE Transactions on Communications* **3(4)** 396-402
[23] Bloom C 2015 *California Institute of Technology: Solving the problems of context modeling* (Access mode: http://www. cbloom.com/papers/ppmz.zip)
[24] Li X and Orchard M T 2001 New Edge-Directed Interpolation *IEEE Trans. Image Process* **10** 1521-1527
[25] Gashnikov M V and Glumov N I 2016 Development and Investigation of a Hierarchical Compression Algorithm for Storing Hyperspectral Images *Optical Memory and Neural Networks* **25(3)** 168-179

[26] Gashnikov M V and Glumov N I 2015 Hyperspectral images repository using a hierarchical compression *Posters proc. of 23 Int. Conf. on Computer Graphics, Visualization and Computer Vision* 1-4

Acknowledgments
This paper was funded by RFBR according to the research projects 18-01-00667, 18-07-01312.

3D image combination in aviation computer vision systems

A I Novikov[1], A I Efimov[1] and D A Kolchaev[1]

[1]Ryazan State Radio Engineering University, Gagarina 59/1, Ryazan, Russia, 390005

Abstract. The relevance of 3D image formation in computer vision systems being applied in aviation is caused by the need to solve many problems using helicopter and unmanned aviation. These problems include mounting of various technical installations with the help of helicopter, rescue mission in complex terrain as well as landing helicopter on unknown sites. The work considers the possibility to use two alternative methods to form 3D images of the underlying surface from the sequence of 2D images obtained with the help of stereo pair placed on aircraft. The first method is based on applying the algorithm of flat image combination to three-dimensional case with the help of homography matrix. The mathematical apparatus developed for the case of flat homography with arbitrary number of key points is generalized to a three-dimensional case. Homography matrix takes into account all kinds of point cloud deformations (shift, rotation, scale change) and allows to obtain the final solution in one step. The second method of point cloud combination is based on applying an iterative procedure to sequentially refine the transformation of one point cloud into the other. In every iteration the key points are found using FAST algorithm. Correspondence between them is established using non-dense optical flow algorithm, optimum estimate of rotation matrix is found in accordance with Kabsch algorithm.

1. Introduction

The article presents the research results of two methods to reconstruct 3D images of underlying surface in the Earth's plane based on the sequence of flat video images of this surface obtained from aircraft (AC) with the help of stereo pair. Image processing is supposed to be performed in an onboard computer in real-time mode. The reconstruction process of underlying surface 3D images is associated with the solution of several interrelated tasks: point clouds sequence formation based on depth maps and matrix with camera parameters; determination of key points on a pair of neighbouring images and establishing the correspondence between them; combination of current and previous point clouds. Conversion of one point cloud to the other involves two transformations: performing point cloud rotation in space and shift along the vector. The transformations required can be performed either simultaneously during the process of transformation with the help of homography matrix, or sequentially, usually after several iterations.

2. The object of the research

The object of the research is the process of converting 2D images of the underlying surface into point clouds in R^3 space with further "sewing" these clouds into a single 3D image. The problem of clouds sewing has the following mathematical statement: Two point clouds $V = \{M_i(x_i, y_i, z_i)\}_{i=1}^{n}$ and $V' = \{M_i'(x_i', y_i', z_i')\}_{i=1}^{m}$ are given in R^3 space. The second point cloud is shifted and rotated in space

relative to the first cloud. In the first and second point clouds there is a certain set of $k\ (k \le \min(n,m))$ points that are images of the same points of a real scene. These points are called appropriate. The $G = \{M_j, M'_j\}_{j=1}^{k}$ set of appropriate points is the basis to determine transformation F parameters, which allows two surfaces (two point clouds) to be "sewn together". The required transformation F $(X = F(X'))$ includes rotation matrix $R = (r_{ij})_{i,j=1}^{3}$ and shift vector $t = (t_x, t_y, t_z)^T$ to be determined:

$$X = R \cdot X' + t = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \quad (1)$$

The transformation parameters F are found from the condition

$$[R|t] = \arg\min_{R,t} \|X - R \cdot X' - t\|^2 \quad (2)$$

Where $X = (x\ y\ z)^T$ and $X' = (x'\ y'\ z')^T$ are column vectors, the elements of which are the coordinates of appropriate points from point clouds being sewn.

Solving the problem of combining two point clouds requires solving two particular problems:

1) forming the set $G = \{M_j, M'_j\}_{j=1}^{k}$ of appropriate point pairs in coincident point clouds;

2) determining the transformation parameters F (rotation matrix R and shift vector t).

The most difficult one is the first task. The known methods to solve it can be divided into three groups:

1) surface feature identification of compared point sets [1];

2) usage of spin image signatures [2, 3];

3) complete point permutation in clouds being compared [4].

Correct determination of the relationships between appropriate parts of point cloud pair allows them to be combined in one step. Otherwise, it is necessary to use multi-step registration procedures, which sequentially reduce error, but require correspondingly large cost of computational time for their implementation.

3. Methods of the research

The authors have developed an algorithm based on the usage of spin image signatures [2, 3] to form the set $G = \{M_j, M'_j\}_{j=1}^{k}$ of appropriate points. A general scheme of the method proposed for a regular grid is as follows: at every point M_{i_0} of the first and second point clouds, the least squares method is used to find estimates $\hat{a}, \hat{b}$ of coefficients a, b of equation of the plane $z = ax + by$ passing through the given point. Optimum estimates $\hat{a}, \hat{b}$ are found as a result of minimizing the deviation of the required plane from some set of points $\{M_i\}_{i=1}^{t}$, $t = 4s(s+1)$ taken from the square neighbourhood of point M_{i_0} sized $(2s+1)\times(2s+1)$. These estimates $\hat{a}, \hat{b}$ are necessary for normal vector $\mathbf{n} = (\hat{a}, \hat{b}, -1)$ formation at every point. All normal vectors are normalized $\left(\mathbf{n}^0 = \frac{\mathbf{n}}{|\mathbf{n}|}\right)$.

Each point M_{i_0} from the first and second point clouds is associated with vector $\{(\gamma_j, \alpha_j, \beta_j)\}_{j=1}^{t}$, elements of which are triples of numbers $(\gamma_j, \alpha_j, \beta_j)$ - point descriptors M_{i_0}. Where:

γ_j - the corner between normal vector $\mathbf{n}_{i_0}^0$ of the current point M_{i_0} and vector $\mathbf{n}_j^0$ $j \in \{1, 2, \ldots, t\}$ of point M_j from point neighborhood M_{i_0};

β_j - algebraic projection of vector $\overline{M_{i_0}M_j}$ on vector $\mathbf{n}^0_{i_0}$, that is $\beta_j = \mathrm{Pr}_{\mathbf{n}^0_{i_0}} \overline{M_{i_0}M_j} = \left(\overline{M_{i_0}M_j}, \mathbf{n}^0_{i_0}\right)$; $\alpha_j = \sqrt{\left|\overline{M_{i_0}M_j}\right|^2 - \beta_j^2}$.

Comparison of point descriptors from the first and second point clouds is performed using the following algorithm. Let us suppose that $\left\{\left(\gamma_j^{(1)}, \alpha_j^{(1)}, \beta_j^{(1)}\right)\right\}_{j=1}^{t}$ - the set of descriptors of a fixed point $M_{i_0}^{(1)}$ from the first cloud, and $\left\{\left(\gamma_j^{(2)}, \alpha_j^{(2)}, \beta_j^{(2)}\right)\right\}_{j=1}^{t}$ - the set of descriptors of some arbitrary point $M_i^{(2)}$ from the second cloud, checked for correspondence to point $M_{i_0}^{(1)}$. For each point $M_i^{(2)}$ from the second cloud, the proximity condition for each parameter is checked first, namely, in the cycle of *j* from 1 to *r*, the following conditions are checked: $\left|\gamma_j^{(1)} - \gamma_j^{(2)}\right| < \varepsilon_1$, $\left|\alpha_j^{(1)} - \alpha_j^{(2)}\right| < \varepsilon_2$, $\left|\beta_j^{(1)} - \beta_j^{(2)}\right| < \varepsilon_3$ where $\varepsilon_1, \varepsilon_2, \varepsilon_3$ are the given numbers. If at least one of three conditions is not met for at least one value of *j* index, then transition to a new point of the second cloud occurs. Otherwise, the sums are found

$$\Sigma_\gamma\left(M_i^{(2)}\right) = \sum_{j=1}^{t}\left|\gamma_{i_0 j}^{(1)} - \gamma_{ij}^{(2)}\right|, \quad \Sigma_\alpha\left(M_i^{(2)}\right) = \sum_{j=1}^{t}\left|\alpha_{i_0 j}^{(1)} - \alpha_{ij}^{(2)}\right|, \quad \Sigma_\beta\left(M_i^{(2)}\right) = \sum_{j=1}^{t}\left|\beta_{i_0 j}^{(1)} - \beta_{ij}^{(2)}\right|, \quad i \in I. \tag{3}$$

For each sum in the composition (3) there are points $M_s^{(2)}$, $s = 1, 2, 3$, in each of which the minimum of corresponding sum from (3) is reached.

$$M_{i_1}^{(2)} = \arg\min_{M_i^{(2)}} \Sigma_\gamma\left(M_i^{(2)}\right), \quad M_{i_2}^{(2)} = \arg\min_{M_i^{(2)}} \Sigma_\alpha\left(M_i^{(2)}\right), \quad M_{i_3}^{(2)} = \arg\min_{M_i^{(2)}} \Sigma_\beta\left(M_i^{(2)}\right) \tag{4}$$

If $i_1 = i_2 = i_3 = i^*$, then point $M_{i^*}^{(2)}$ from the second cloud is identified as appropriate to point $M_{i_0}^{(1)}$ from the first cloud. Otherwise, point $M_i^{(2)}$ is recognized as inappropriate to point $M_{i_0}^{(1)}$.

The described algorithm to form pairs of corresponding points was tested on a model point cloud pair. The first point cloud V is the surface of a hyperbolic paraboloid

$$V = \left\{(x, y, z) \;\middle|\; z = \frac{x^2}{50} - \frac{y^2}{75},\ x = \overline{-10,20};\ y = \overline{-8,8}\right\}.$$

The second point cloud V' was obtained by a similar formula and with the same coordinates along Oy axis, and along Ox axis in the interval from 10 to 30.

$$V' = \left\{(x, y, z) \;\middle|\; z = \frac{x^2}{50} - \frac{y^2}{75},\ x = \overline{10,30};\ y = \overline{-8,8}\right\}.$$

The first and second point clouds have a common part (overlapping area) in amounts of 187 points: on coordinate x from 10 to 20, and on coordinate y - from -8 to 8. Both point clouds are shown in figure 1.

The points of the first cloud are marked with circles of blue, and the second cloud - with red squares. The intersection of point clouds is clearly visible in figure 1 as a horizontal strip from combined points of the first and the second clouds in the interval from 10 to 20 along Ox axis.

After point cloud formation, the second cloud underwent shear and rotation transformations. It was shifted by 25 units on coordinate z, so the shift vector t has $t = (0, 0, 25)^T$ form. Then it was rotated relative to each of three axes Oz, Oy и Ox, by angles in accordance with $3^o, 3^o, 6^o$. The rotation was carried out in the following sequence: firstly- in the plane of roll by $\gamma = \frac{\pi}{30}$ angle, then in the plane of

pitch by $\theta = \frac{\pi}{60}$ angle and in the end - in the plane of course by $\psi = \frac{\pi}{60}$ angle. As a result, a rotation matrix (5) was obtained.

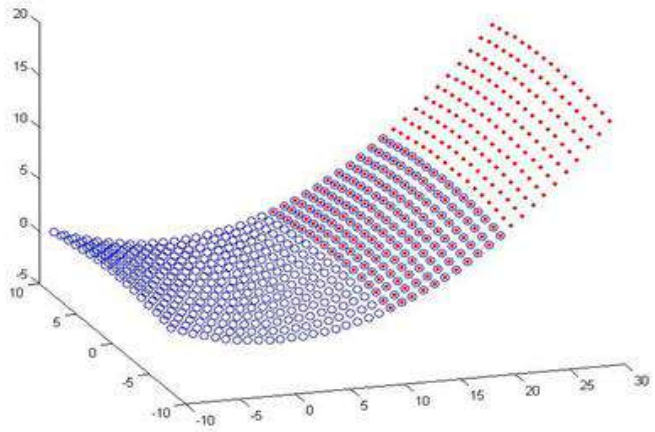


Figure 1. The first and second point clouds before conversion.

$$R(\psi,\theta,\gamma) = T(\psi)\cdot T(\theta)\cdot T(\gamma) = \begin{pmatrix} 0.9973 & -0.0466 & 0.0574 \\ 0.0523 & 0.9934 & -0.1017 \\ -0.0523 & 0.1044 & 0.9932 \end{pmatrix} \quad (5)$$

The result of the second cloud points conversion by formula (1) is shown in figure 2. Now it's necessary to perform appropriate point search in a cloud pair using algorithm (3) - (4). Algorithm results are shown in figure 3.

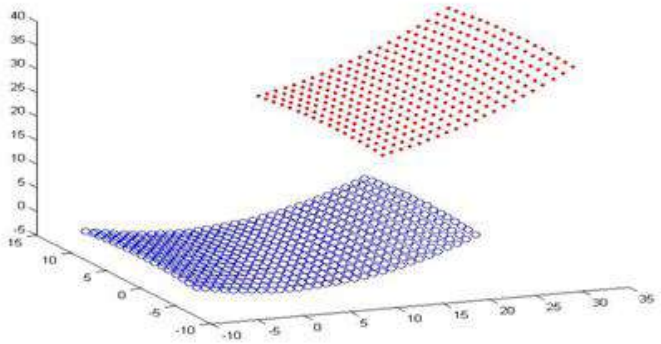


Figure 2. Point clouds after the second point transformation.

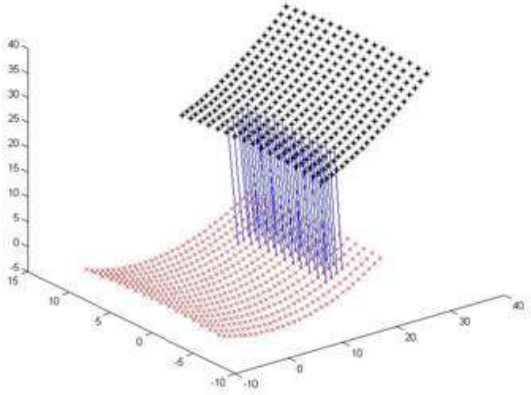


Figure 3. Correspondences found between points in a cloud pair.

The analysis of obtained correspondences between points of two surfaces proved their correctness. However, the algorithm "eliminated" some set of true correspondences, leaving only 45 pairs of

points. In this case, as it can be seen below, this number of points is enough to perform proper surface combination.

By analogy with the problem of combining two-dimensional images using homography matrix, the problem of combining (sewing) two point clouds $V = \{(x_i, y_i, z_i)_{i=1}^n\}$ and $V = \{(x_i', y_i', z_i')_{i=1}^n\}$ can be solved by homography in R^3 space. We will search a matrix of the following type:

$$H = \begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \\ h_{41} & h_{41} & h_{41} & 1 \end{pmatrix},$$

that would lead to fulfill the following matrix equalities in homogeneous coordinates

$$\begin{pmatrix} \lambda x_i \\ \lambda y_i \\ \lambda z_i \\ \lambda \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \\ h_{41} & h_{42} & h_{43} & 1 \end{pmatrix} \begin{pmatrix} x_i' \\ y_i' \\ z_i' \\ 1 \end{pmatrix} \tag{6}$$

In this matrix $H_{4\times4}$, submatrix $R = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}$ is responsible for point cloud rotation, and submatrix $T = (h_{14} \;\; h_{24} \;\; h_{34})^T$ - for point cloud shift.

Having eliminated parameter λ from the system of equations (6), we obtain a system of linear algebraic equations (SLAE) for each pair of appropriate points $M_i(x_i, y_i, z_i)$ and $M_i'(x_i', y_i', z_i')$, $i = \overline{1,k}$ from the point clouds combined

$$\begin{cases} x_i'h_{11} + y_i'h_{12} + z_i'h_{13} + h_{14} - x_i'x_ih_{41} - y_i'x_ih_{42} - z_i'x_ih_{43} = x_i, \\ x_i'h_{21} + y_i'h_{22} + z_i'h_{23} + h_{24} - x_i'y_ih_{41} - y_i'y_ih_{42} - z_i'y_ih_{43} = y_i, \\ x_i'h_{31} + y_i'h_{32} + z_i'h_{33} + h_{34} - x_i'z_ih_{41} - y_i'z_ih_{42} - z_i'z_ih_{43} = z_i. \end{cases}$$

or in matrix form $A \cdot G_H = B$. Where: A - block matrix of $3k \times 15$ order:

$$A_{3k\times15} = \begin{pmatrix} Z & O & O & W_1 \\ O & Z & O & W_2 \\ O & O & Z & W_3 \end{pmatrix}$$

With blocks in matrix form $Z_{k\times4} = (x_i, y_i, z_i, 1)_{i=1}^k$, $W_1 = (-x_ix_i', -y_ix_i', -z_ix_i')_{i=1}^k$, $W_2 = (-x_iy_i', -y_iy', -z_iy')_{i=1}^k$, $W_3 = (-x_iz_i', -y_iz_i', -z_iz_i')_{i=1}^k$; , O - zero matrix of order $k \times 4$, G_H -column vector of unknowns of size 15×1: $G_H = (h_{11}\, h_{12}\, h_{13}\, h_{14}\, h_{21}\, h_{22}\, h_{23}\, h_{24}\, h_{31}\, h_{32}\, h_{33}\, h_{34}\, h_{41}\, h_{42}\, h_{43})^T$; B - column vector of free terms $B = (x_1\; x_2\; \ldots x_k\; y_1\; y_2\; \ldots y_k\; z_1\; z_2\; \ldots z_k)^T$ of size $3k\times1$. Matrices W_1, W_2, W_3 are of $k\times3$ $(\dim W_i = k\times3, i = 1,2,3)$ order.

Due to measurement errors, matrix equality (6) may not be fulfilled for some points from point clouds. We will look for matrix H, which minimizes the distance between points of one set and pre-images of the other one. Problems to find homography matrices in transformation of one set of points to another set of points with an arbitrary number of points in them, both in two-dimensional case and in three-dimensional case, are formulated identically

$$\|AG_H - B\|^2 \to \min_{G_H}. \tag{7}$$

Where $\|(\cdot)\|$ - Euclidean norm.

The solution of task (7) with three-dimensional homography by least square method leads to normal SLAE [5]

$$(A^TA)G_H = A^TB \tag{8}$$

with a basic matrix

$$A^T A = \begin{pmatrix} Z^T Z & O_{4\times 4} & O & Z^T W_1 \\ O & Z^T Z & O & Z^T W_2 \\ O & O & Z^T Z & Z^T W_3 \\ W_1^T Z & W_2{}^T Z & W_3{}^T Z & \sum\limits_{i=1}^{3} W_i^T W_i \end{pmatrix}.$$

The right side of normal equation (8) has the following form: $A^T B = \begin{pmatrix} Z^T B_1 & Z^T B_2 & Z^T B_3 & \sum\limits_{i=1}^{3} W_i^T B_i \end{pmatrix}^T$. In these expressions $B_1 = (x_1\ x_2\ ....x_k)^T$, $B_2 = (y_1\ y_2\ ...y_k)^T$, $B_3 = (z_1\ z_2\ ...z_k)^T$ - column vectors of size $k \times 1$.

We apply the algorithm of three-dimensional homography for point cloud combination shown in fig.2. The system of linear algebraic equations that follows from matrix equation (8) consists of 45 equations with respect to 15 unknowns - elements of homography matrix in three-dimensional case - and is considered to be a redefined SLAE. Normal pseudo solution [6, 7] of this system was found as the solution of SLAE (8).

Homography matrix obtained as the result of SLAE solution (8) that ensures point cloud combination (transformation of first point cloud to second) has the following form

$$H = \begin{pmatrix} 0.997 & -0.047 & 0.057 & 2.58\cdot 10^{-10} \\ 0.052 & 0.993 & -0.102 & 7.16\cdot 10^{-12} \\ -0.052 & 0.104 & 0.993 & 25.0 \\ -2.29\cdot 10^{-12} & -2.36\cdot 10^{-15} & 3.69\cdot 10^{-12} & 1 \end{pmatrix} \quad (9)$$

It should be noted that submatrix $H_R = \begin{pmatrix} 0.997 & -0.047 & 0.057 \\ 0.052 & 0.993 & -0.102 \\ -0.052 & 0.104 & 0.993 \end{pmatrix}$ in its composition, which is the estimate of rotation matrix, coincides with a given matrix $R(\psi,\theta,\gamma)$ (5) within the accuracy chosen $(\varepsilon = 10^{-5})$. Submatrix $H_t = \begin{pmatrix} 2.58\cdot 10^{-10} & 7.16\cdot 10^{-12} & 25.0 \end{pmatrix}$ also gives almost accurate estimate of shift vector *t*. The result of point cloud sewing through the use of obtained homography matrix is almost ideal.

Let's now explore the influence of errors while determining point coordinates in point clouds on the accuracy of surface combination with homography matrix. We shall add normally distributed random component with zero mathematical expectation and given noise level to coordinate *z* in each point cloud. Let's perform two experiments. In the first $\sigma = 0.01$, and in the second $\sigma = 0.1$,

Homography matrix

$$H = \begin{pmatrix} 0.996 & -0.046 & 0.060 & 0.009 \\ 0.053 & 0.993 & -0.103 & -0.006 \\ -0.052 & 0.104 & 0.993 & 24.985 \\ -7.46\cdot 10^{-5} & 2.70\cdot 10^{-5} & 0.0001 & 1 \end{pmatrix}$$

found from noisy data at noise level $\sigma = 0.01$, insignificantly varies in comparison with homography matrix (9) obtained from undistorted data. In rotation submatrix H_R, the values of elements have changed by no more than 0.001. Elements of the fourth row of the matrix have changed to a greater degree, but not so much to significantly affect the results of point cloud transformation.

Errors of surface combination substantially increase with increasing noise level**.** Homography matrix found at $\sigma = 0{,}1$ error level in the original data has the following form

$$H = \begin{pmatrix} 0.831 & -0.043 & 0.225 & 0.850 \\ 0.047 & 0.931 & -0.092 & 0.017 \\ -0.252 & 0.099 & 1.255 & 25.136 \\ -0.007 & 1.64\cdot 10^{-5} & 0.012 & 1 \end{pmatrix}.$$

Substantial distortions appeared in submatrix H_R that is responsible for the rotation. Elements h_{13}, h_{31} values increased fourfold in comparison with (9), diagonal elements of this submatrix changed absolutely in the range from 0.062 to 0.262 (by 26%). Some elements of the fourth row of the matrix also changed. They also affect the result of point clouds transformation. Transformed point cloud distortions can be observed in figure 4 in point cloud projection on *Oxy* plane.

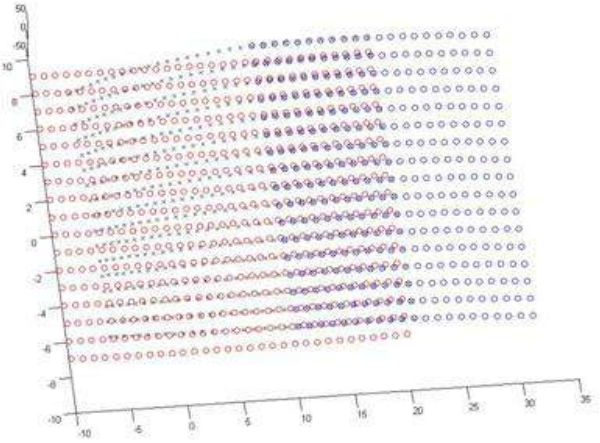


Figure 4. Point cloud combination results in points projection on *Oxy* plane at noise level $\sigma = 0.1$.

The research allows to make the following conclusion: 3D Surfaces combination based on one point cloud transformation to the other using homography matrix gives good results at precisely known point coordinates in point clouds and requires low computational costs. However, even low level of errors in point coordinates in clouds being combined can lead to significant distortions of combining results.

Unstable results of 3D image combination with the help of homography matrix force to look for other combination methods. A method intended to form 3D images of underlying surface in an onboard computer of aircraft must meet serious requirements for both speed and precision of results. These requirements are generally satisfied by sequential refinement method to transform one point cloud to the other. In the algorithm proposed, point cloud formation is made from a sequence of 2D image pairs of the underlying surface obtained from stereo pair placed on aircraft. Point cloud sewing is performed through the use of Kabsch algorithm [8]. The algorithm consists of several interrelated blocks, each of which solves its own independent task, namely:

- Block for depth maps formation from a sequence of 2D images pairs of the underlying surface from the stereo pair;
- Block for point cloud sequence formation based on depth maps and a matrix containing camera parameters;
- Block for key points determination on *i-th* and (*i +1*)*-th* images using FAST algorithm [9];
- Block to establish the correspondence between key points in a point cloud pair using Lucas-Kanade non-dense optical flow algorithm [10];
- Block to combine *i-th* and (*i +1*)*-th* point clouds with the search of rotation matrix R using Kabsch algorithm.

In the process of combining the images containing uniform surfaces (table, floor, flat areas of the earth's surface without pronounced features, etc.), the standard version of ICP algorithm gives poor results. On such surfaces, it is difficult to find key points and establish correspondences between them. Using FAST key point detector and optical flow algorithm in a stereo pair scheme allows finding key points on adjacent frames of flat images and establishing the correspondence between them in two-dimensional space. The correspondences established between points on flat images allow finding correspondences between key points of points in three-dimensional space.

The choice of FAST algorithm as a detector of key points is caused mostly by its high performance. Unlike other detectors using partial derivatives calculation, this algorithm is based on direct comparison of investigated pixel brightness and therefore has a minimum number of computational operations. FAST algorithm constructs a circle of 16 pixels inscribed in a square with the side of 7 pixels. The brightness of each of these pixels is compared with the brightness of a central pixel and is referred to one of three classes (darker, comparable, brighter ones). As a result, set P of all checked pixels is divided into three subsets, the decision tree is constructed in accordance with algorithm [9]. The decision tree built is used to identify singular points.

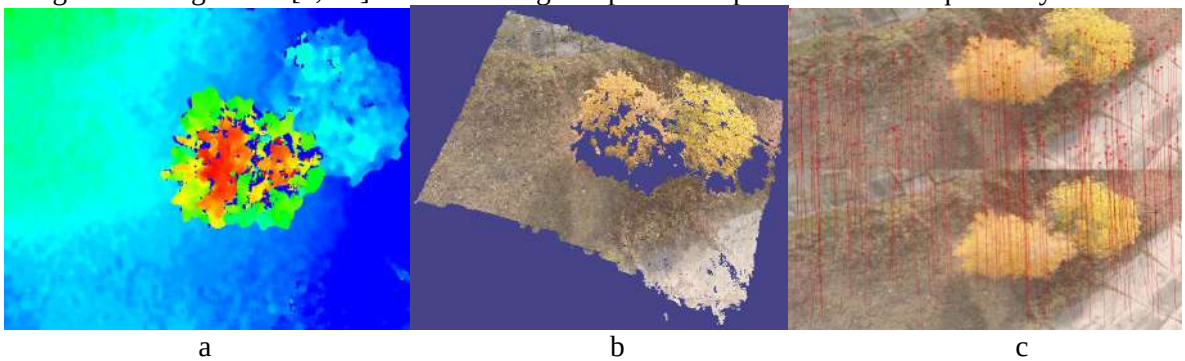
Depth map is constructed on the basis of determining displacement value between corresponding pixels of left and right images. The search of a corresponding pixel in a row is made by a window of specified size passing through image and determining objective function maximum. To improve the quality of a depth map received, cameras are calibrated, image rectification is performed [13], and image filtration and improvement for example using the method described in [14] can also be made. Figure 5a shows depth map in pseudo coloring, where values of disparities are respectively large (red) and small (blue) from red to blue.

On the second step, point clouds P_i and Q_{i+1} are formed for *i-th* and (*i+1*)*-th* frames by displaying corresponding depth maps to three-dimensional space using matrix H containing camera parameters. Figure 5b shows a three-dimensional point cloud formed at this stage.

The third step of the algorithm is designed to search multiple key point pairs. Key points on *i-th* image are found by FAST detector. For key points found, using optical flow, appropriate points on (*i* +1)*-th* image (Figure 5c) are found. In the method considered, optical flow is realized using Lucas-Kanade algorithm [10] and its enhancement in [15].

For neighboring frames of time arrival from stereo pair, ordered arrays $P_i^0 = \left\{p_{i1}^0, p_{i2}^0,\ldots,p_{ik}^0\right\}$ and key points $Q_{i+1}^0 = \left\{q_{i+1,1}^0, q_{i+1,2}^0,\ldots,q_{i+1,k}^0\right\}$ are formed. In these arrays, point $p_{ij}^0 = \left(x_j, y_j, z_j\right)$ with number j from *i-th* point cloud corresponds to point $q_{i+1,j}^0 = \left(x'_j, y'_j, z'_j\right)$ with the same number from cloud *i*+1.

The next 8 steps of the algorithm implement an iterative procedure of two point clouds combination in accordance with the logic of ICP algorithm. In this case, rotation matrix R is found using Kabsch algorithm [8, 11]. The following computational procedures are sequentially realized.



a b c

Figure 5. Step-by-step results of modified iterative point cloud combination algorithm.

P_i^0 arrays mass centers are calculated Q_{i+1}^0

$$\bar{p}_i^0 = \frac{1}{k}\sum_{j=1}^{k} p_{ij}^0, \qquad \bar{q}_{i+1}^0 = \frac{1}{k}\sum_{j=1}^{k} q_{i+1,j}^0 \tag{10}$$

The coordinates of mass centers found are subtracted from the coordinates of each point of a corresponding array

$$p_{ij}^1 = p_{ij}^0 - \bar{p}_i^0 \qquad q_{i+1,j}^1 = q_{i+1,j}^0 - \bar{q}_{i+1}^0 \qquad j = \overline{1,k},$$

or, in a matrix form

$$P_i^1 = P_i^0 - \bar{p}_i^0 \cdot E, \qquad Q_{i+1}^1 = Q_{i+1}^0 - \bar{q}_{i+1}^0 \cdot E \tag{11}$$

Then, a covariance matrix is calculated

$$H = \left(P_i^1\right)^T Q_{i+1}^1 \tag{12}$$

and then its singular decomposition is found

$$H = USV^T \tag{13}$$

Product of right and transpose left singular matrices gives required rotation matrix R

$$R = VWU^T \tag{14}$$

where $W = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & d \end{pmatrix}$, and $d = sign(\det(VU^T))$.

Shift vector is calculated

$$t = \bar{p}_i^0 - R \cdot \bar{q}_{i+1}^0 \tag{15}$$

Combination of $i+1$ point cloud (array of points Q_{i+1}^0) with i cloud is performed.

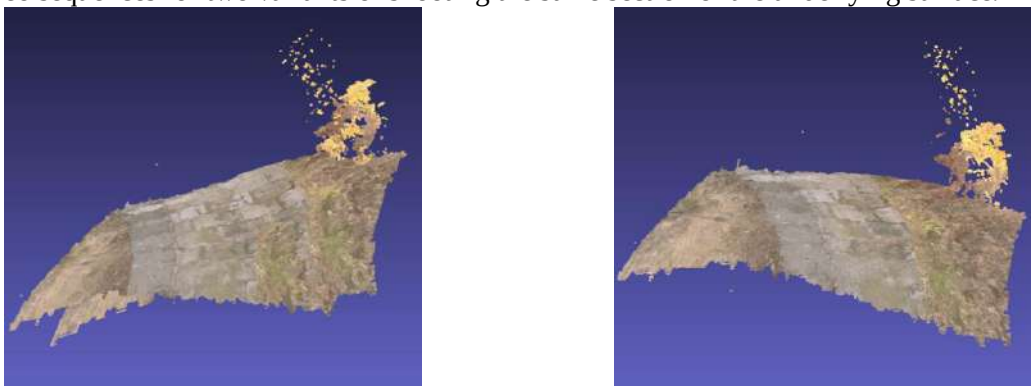
$$\hat{Q}_{i+1} = R \cdot Q_{i+1}^0 + t \tag{16}$$

On the last 11th step of the algorithm, approximation error of converted $i+1$ point cloud $\hat{Q}_{i+1}$ to *i-th* cloud P_i^0 at $r+1$ iteration is calculated

$$\delta_{i+1}^{r+1} = \left\| \hat{Q}_{i+1} - P_i^0 \right\| = \sum_{j=1}^{k} \left(\left(\hat{x}_j - x_j\right)^2 + \left(\hat{y}_j - y_j\right)^2 + \left(\hat{z}_j - z_j\right)^2 \right).$$

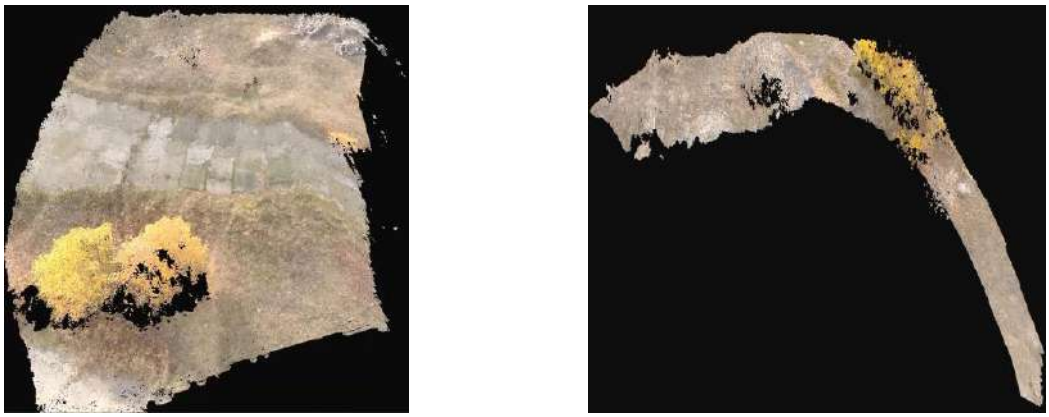
If in the equation $\left|\delta_{i+1}^{r+1} - \delta_{i+1}^{r}\right| < \varepsilon$ is fulfilled, then the decision to complete an iterative process of successive point cloud approximation $i+1$ to *i-th* cloud is made. Otherwise, calculation cycle is repeated.

Let us consider the example of applying the described algorithm for the formation and combination of point cloud sequence based on video images of the underlying surface obtained with video cameras pair placed on hexacopter suspension. Figure 6a shows two point clouds (arrays P_i^0 and Q_{i+1}^0) obtained at the output of the algorithm's second block. The result of their combination is shown in figure 6b - the result of their combination. Figure 7 shows point cloud “sewing” results obtained from two stereo sequences for two variants of shooting the same section of the underlying surface.



a b

Figure 6. Illustration of point cloud combination: a - point clouds position before combination, b - the result of combination.



a b

Figure 7. Results of point cloud sewing: a- by the 1st sequence of stereo frames, b- by the 2nd sequence.

4. Conclusion

The computational formulas are obtained and the study on model images of 3D image combination algorithm based on homography matrix usage is carried out. On the basis of this approach, the algorithm to find appropriate points in combined clouds is proposed. The algorithm is based on spin image signatures application. Homography matrix allows taking into account all kinds of possible point cloud transformations. The algorithm can operate in real time and it provides good quality of surface combination in the absence of errors in coordinate assignment of points in clouds being combined. At the same time, in the process of research this method was shown to be very sensitive to perturbations of point coordinates in point clouds being combined.

A modified version of ICP iterative algorithm, which provides the solution to the task of 3D images of underlying surface reconstruction in real time with good quality, is proposed. Main differences between the proposed variant of an iterative algorithm and its classical variant are the application of Lucas-Kanade algorithm of non-dense optical flow together with FAST algorithm to form a set of key point pairs as well as Kabsch algorithm to find rotation matrix.

5. References

[1] Stein F and Medioni G 1992 Structural Indexing: Efficient 3-D Object Recognition *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14(2)** 125-145

[2] Johnson A and Hebert M 1997 Surface Registration by Matching Oriented Points *Proc. 3 DIM* 145-152

[3] Chernikov I S and Krylovetsky A A 2009 Three-dimensional reconstruction and recognition by sets of range-finding data *Computer modeling 2009: works of international scientific and technical conference* (St. Petersburg) 284-287

[4] Chen C, Hung Y and Cheng J 1999 Ransac-Based Darces: A New Approach to Fast Automatic Registration of Partially Overlapping Range Images *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21(11)** 1229-1234

[5] Demidovich B P and Maron I A 1966 *Fundamentals of computational mathematics* (Moscow: Science) 664 p

[6] Voevodin V V and Kuznetsov Yu A 1984 *Matrices and Calculations* (Moscow: Science) 320 p

[7] Andrushevsky N M 2008 *Analysis of stability of solutions of systems of linear algebraic equations* (Moscow: MAKS Press) 71 p

[8] Kabsch W 1976 A solution for the best rotation to relate two sets of vectors *Acta Crystallographic* **32(5)** 922-923

[9] Nikolajczyk K and Schmid C 2004 Scale and affine invariant interest point detectors *International Journal of Computer Vision* **60(1)** 63-86

[10] Bouguet J-Y 2017 Pyramidal Implementation of Lucas Kanade Feature Tracker Description of algorithm *Intel Corporation Microprocessor Research Labs* (Access mode: http://robots. stanford.edu>cs223b04/algo_tracking) (11.10.2017).

[11] Stein F and Medioni G 1992 Structural Indexing: Efficient 3-D Object Recognition *IEEE Transactions of Pattern Analysis and Machine Intelligence* **14(2)** 125-145
[12] Bergström P and Edlund O 2014 Robust registration of point sets using iteratively reweighted least squares *Computational Optimization and Applications* **58(3)** 543-561 DOI: 10.1007/s10589-014-9643-2
[13] Bouguet J Y 2017 Camera Calibration Toolbox for Matlab *Computational Vision at the California Institute of Technology* (Access mode: http://www.vision.caltech.edu /bouguetj/calib_doc/) (16.10.2017)
[14] Nikonorov A V, Petrov M V, Bibikov S A, Kutikova V V, Morozov A A and Kazanskiy N L 2017 Image restoration in diffractive optical systems using deep learning and deconvolution *Computer Optics* **41(6)** 875-887 DOI: 10.18287/2412-6179-2017-41-6-875-887
[15] Lyubutin P S 2015 Development of optical flow computation algorithms for strain measurement of solids *Computer Optics* **39(1)** 94-100 DOI: https://doi.org/10.18287/0134-2452-2015-39-1-94-100

Soil characteristics influence on the winter wheat yield regression modeling using NDVI vegetation index

A Y Denisova[1] and A V Evstiforova[1]

[1]Samara National Research University, Moskovskoye shosse 34, Samara, Russia, 443086

Abstract. The general way of a yield estimation using remote sensing data is a regression modeling based on cumulative NDVI vegetation index. Usually, the other additional factors are also included in the model to enhance a yield estimation accuracy, for example, weather conditions during the vegetation season or particular crop type information. Nevertheless, such factors as soil characteristics were not properly investigated and it is still an open question whether they should be included in modeling process or not. This article reveals the regression modeling accuracy dependency on the additional soil characteristics for winter wheat yield estimation using cumulative NDVI. We analyzed several groups of factors obtained from different data sources such as remote sensing images, geoinformation systems and meteorological data. These factors were used to construct primarily yield models without soil factors. The best two models were chosen as the basic models and then they were modified to take into account the additional soil characteristics. Finally, the comparison of the basic models with the modified models including soil parameters was made. We investigated such soil parameters as soil fertility score derived from a ten-year-old soil map and the results of the current agrochemical inspections. The comparison of basic and modified models was made for the winter wheat fields located in Samara region, Russia. The experiments showed that regression models give better determination coefficient and root mean square error if the additional soil characteristics are used for modeling. Moreover, the older data obtained from soil maps and reference book of the soil fertility score delivers better prediction than the current agrochemical inspection data. Therefore, there is no need to know lots of soil parameters for successfully winter wheat yield modeling. The results of this research can be applied to forecasting winter wheat yield in Samara region.

1. Introduction

Traditional yield prediction agricultural methods require a vast amount of ground truth measurements that makes these methods very expansive in use. To provide automated yield prediction without ground truth measurements the geoinformation systems (GIS) and remote sensing data (RS data) can be used. RS data provide objective information about agricultural crop growth whereas GIS contains effective instruments for additional parameters calculation based on graphic and semantic information about the field. Therefore, the development of simultaneous GIS and RS data for agricultural purposes remains a very promising research area.

Existing GIS and RS data yield prediction approaches are highly connected with the traditional methods used in agriculture that mainly utilize linear regression modeling and simulation modeling to find the relation between the observed factor values and the yield [1]. Traditionally only one critical factor with the greatest impact on the crop is used in regression modeling. It can be air temperature, moisture availability or sun radiance. This critical factor is generally determined by the experts as the

parameter which is probably the main reason for the crop stress in the current situation. Traditional linear regression models usually give yield forecast at the beginning of the vegetation season and define the potential crop yield. Simulation modeling is more precise than regression modeling, but it is more computationally intensive and requires lots of very specific parameters to be known. The main idea behind this approach is to simulate the crop phenology by means of several equations linking bioclimatic parameters and changes in vegetation biomass. The result of simulation modeling is an assessment of dry vegetation biomass for each date of the vegetation season. The final yield is determined as a cumulative biomass estimated using simulating modeling. The RS and GIS data-based yield models works in a similar way to traditional techniques, but the main variable reflecting biomass changes is estimated by means of RS data. Usually, the vegetation indexes (NDVI, LAI and others) are used to substitute the biomass variables in traditional simulating and regression models [2], wherein the regression models use cumulative vegetation index values. In both traditional and RS and GIS data-based cases, the regression modeling has more simple mathematical form and calibration than simulating modeling. Moreover, simulation models based on RS and GIS data (EPIC [3], WOFOST [4], SAFY [2] and etc.) have a huge number of parameters that still require ground truth measurements for each particular field which cannot be obtained from the RS or GIS data. Thus, the regression models using cumulative vegetation indexes are more preferable for yield prediction than simulation models.

In our research, we apply cumulative NDVI [5] vegetation index in winter wheat yield linear regression modeling. We selected a winter wheat as the most widespread agricultural crop in Samara region. Our research aim was to study the model behavior when the additional factors reflecting soil fertility are included in the model. The motivation of our research came from the fact that the soil fertility score and agrochemical survey data are very helpful in forecasting crop yields in agriculture. However, this fact is not obvious in regression modeling based on NDVI vegetation index. To obtain the basic model without soil factors, we analyzed three groups of factors extracted from different data sources and two approaches of the cumulative NDVI estimation. Several regression models were obtained and the best models with the optimal factor set were used to investigate the influence of soil factors. As a result, the recommendations on soil factors and cumulative NDVI estimation were given.

The article is structured as follows. Section 2 describes the factor groups used to construct the basic soil independent models and the general model equation. Section 3 reveals the details of two approaches to the cumulative NDVI estimation and the process of optimal factor set selection. Section 4 contains the results of the experimental evaluation of soil fertility impact on yield regression modeling.

2. Problem statement

Generally, regression modeling is used to find the relation between the expectation of the dependent random variable y and the set of independent variables $x_1, x_2, ..., x_n$ called factors. The type of regression function is determined according to the heuristic assumptions about the relationship between the observed independent parameters and dependent variable. In the present article, we regard a linear regression model in which the dependent variable y corresponds to the winter wheat yield:

$$y = \sum_{i=1}^{n} a_i x_i + a_0 , \quad (1)$$

where $a_i, i = 0,...,n$ are the coefficients of the model.

We consider several sets of factors for the model (1). The first set (basic set of factors (BSF)) contains factors extracted from the farmer data in GIS, meteorological services and RS data. BSF includes such subgroups of factors as:

1) the remote sensing factors. These factors are calculated using NDVI time series obtained by RS data. They include cumulative and maximum NDVI values for the period of the most active vegetation as well as the vegetation period duration and the vegetation start date. The following equation explains the process of factor extraction:

$$\xi_r = \sum_{t=r}^{T+r} \xi(t), \quad (2)$$

where ξ_r is the cumulative NDVI value, $\xi(t)$ is the average NDVI value for the particular field at the date t, r is the relative date of the vegetation period beginning, T is the vegetation period duration;

2) the bioclimatic factors. These factors are obtained from meteorological services. This subgroup contains such factors as maximum and minimum temperature for the vegetation period, maximum and minimum humidity during the vegetation period and climate zone in which the field is located;

3) the data of agricultural producers join such factors as crop cultivar, seeds reproduction, farmer identifier, date of seeding and geographical coordinates of the field object centroid. The factors of these group are obtained from the GIS database.

The second set of factors (SFS) contains soil fertility factors determined by the GIS soil maps. These factors are the soil fertility score and the humus content. The soil fertility score was determined according to the reference study [6] and the soil type given from the soil map in GIS.

The third factor set (TSF) is constructed as a combination of the soil fertility parameters according to the agrochemical survey data. TSF incorporates such factors as average humus, phosphorus, sulfur and potassium contents in the field soil. The agrochemical survey data were obtained from the semantic information of the agrochemical surveys map in GIS.

We conducted our research in several stages. First of all, we selected an optimal subset of the basic factor set (OBSF). OBSF was determined as the subset of significant factors of the model (1) which delivers the highest value of the determination coefficient (R^2). We further refer the regression model estimated for the OBSF as Model 1. In other words, the Model 1 is a name of the general soil-independent regression model estimated using OBSF. Then, in order to get the soil-dependent models, we combined OBSF with two other factor sets SSF and TSF and estimated the regression models for the resulting factor sets. The regression models obtained for OBSF and SSF are further referred as Model 2, while the regression models estimated using OFBSS and TSF are referred as Model 3. Model 2 and Model 3 are the names of the received soil dependent regression models. Finally, the comparison of soil-dependent and soil-independent models was made.

We applied the least square method [7] to estimate the regression coefficients. The quality of the particular regression model was determined by means of R^2 value [7] and root mean square error ε (RMSE):

$$\varepsilon = \sqrt{\frac{1}{M}\sum_{j=1}^{M}\left(y_j - \hat{y}_j\right)^2}, \quad (3)$$

where $\hat{y}_j$ is a predicted yield value for the field j, y_j is an actual yield value for the field j and M is the number of fields.

3. Optimal factor set selection

The remote sensing subgroup of the basic factor set can be determined in different ways. We propose two approaches to estimate remote sensing factors that differ in estimating the vegetation beginning date r and the vegetation period duration T. Therefore, they lead to different cumulative NDVI values.

The first approach requires the NDVI threshold ξ_T to be given. This threshold is used to achieve the minimum $\tau_{\min}$ and maximum $\tau_{\max}$ relative dates for which the NDVI value is greater than the threshold ξ_T. Therefore, the vegetation beginning date r and the vegetation period duration T are defined as follows:

$$r = \tau_{\min}, \quad T = \tau_{\max} - \tau_{\min}. \quad (4)$$

Figure 1 illustrates the first approach to estimating remote sensing factors.

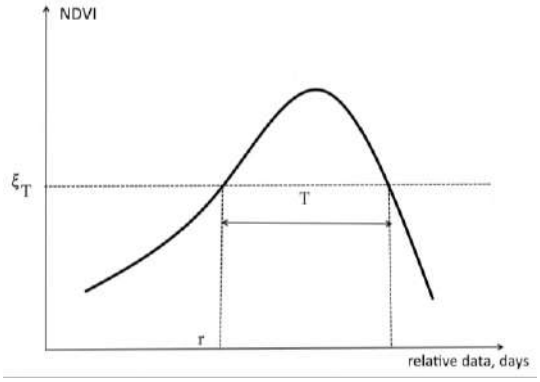


Figure 1. The first approach to estimating the remote sensing factors.

The second approach defines the vegetation period symmetrically around the maximum of NDVI value. Thus, for given vegetation period duration T, we obtain the vegetation beginning parameter in the following way

$$r = t_{\max} - T/2, \tag{5}$$

where $t_{\max} = \arg\max_t \xi(t)$. Figure 2 shows the scheme of the second approach.

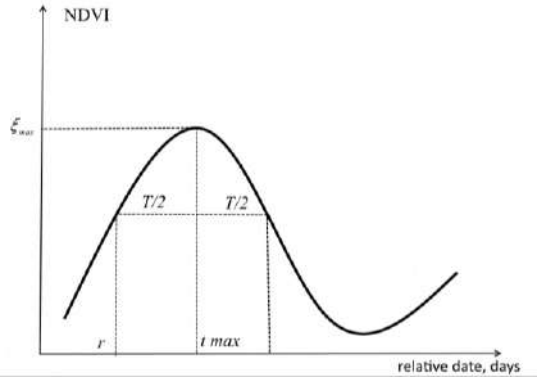


Figure 2. The second approach to estimating the remote sensing factors.

The proposed approaches determine two different subgroups of remote sensing factors. The factors corresponding to the first approach are cumulative NDVI value ξ_r, vegetation beginning date r and vegetation period duration T. The factors formed as a result of the second approach include cumulative NDVI value ξ_r, the date of maximum NDVI value $t_{\max}$ and half-width of the vegetation period $T/2$.

We tested several values of the NDVI threshold ξ_T for the first approach and several values of $T/2$ for the second approach. Nine BSFs corresponding to different remote sensing factors were formed. A complete list of the BSFs considered and their names are given in Table 1.

According to the different BSFs, nine regression models were estimated and the optimal BSF was chosen. The whole process of the optimal BSF selection is described as follows:

1. The independent categorical variables such as farmers identifier, seeds reproduction and others are previously converted in a set of binary variables taking values 0 or 1 depending on the chosen category.
2. Then all factors are normalized to get zero expectation and unit variance.
3. The normalized factors and the dependent variable values are processed by the least square method (LSM) to define the unknown regression coefficients in the expression (1).
4. For each of the BSFs, an adequacy of the constructed regression models is checked by testing the standard LSM hypotheses [7]:

Table 1. The list BSFs achieved for the different remote sensing factor extraction approaches with different parameters.

The basic factor set name	remote sensing factor extraction approach	Parameters
NDVI03	1	$\xi_T = 0.3$
NDVI04	1	$\xi_T = 0.4$
NDVI05	1	$\xi_T = 0.5$
NDVI06	1	$\xi_T = 0.6$
NDVI20	2	$T/2 = 20$
NDVI30	2	$T/2 = 30$
NDVI40	2	$T/2 = 40$
NDVI50	2	$T/2 = 50$
NDVI60	2	$T/2 = 60$

- the hypothesis about the statistical significance of the regression coefficients. This hypothesis is checked using the t-test,
- the hypothesis about the statistical significance of the regression equation. This hypothesis is checked using the F-test,
- the determination coefficient R^2 is estimated. For the proper modeling, the R^2 value has to be greater than 0.8,
- the regression residues analysis is performed. The residues have to be independent, without trend, distributed normally with zero expectation and satisfy the criteria of homoscedasticity.

5. The significant factors of regression models tested for adequacy are taken as optimal basic sets of factors if corresponding regression model has the R^2 value greater or close to 0.8.

4. Experimental research

The experimental study was conducted using the remote sensing and GIS data obtained for the season from April 1 till August 31, 2015. The farmers provided the field boundary information and semantics that includes wheat cultivar, seed reproduction, farmer identifier and crop yield in the observation period in the Agricultural GIS of Samara Region. The total number of fields used in experiments was 127. Climate zones and geographical coordinates of the field center were extracted by means of spatial queries from the vector layer in GIS. The climate zone parameter was defined as the index taking values from 1 to 3 according to the following notation [8]: 1 is northern climate zone, 2 is central climate zone and 3 is southern climate zone.

The factors based on remote sensing data were calculated using NDVI vegetation index time series. We applied the algorithm described in [9-11] and Terra and Aqua MODIS imagery with resolution 250 meters to estimate the NDVI values for each field. The remote sensing data were processed for the whole observation period from April 1 to August 31, 2015.

Climate data such as air temperature, relative humidity and precipitation were extracted from the weather archives of the following weather stations: Kinel-Cherkassy village (No. 548621) with coordinates 53°0'0"N47°0'0"E, Sernovodsk settlement (No. 496568) with coordinates 53°0'0"N92°0'0"E, Krasnoarmeyskoye village (No. 824366) with the coordinates 52°0'0"N 7°0'0"E. To estimate the climate factors we applied the measurements of the nearest to the field meteorological station among the considered ones.

As for the soil fertility factors, the soil fertility score was defined corresponding to the soil type marked on the soil map and the reference book [6]. The agrochemical survey data were provided by Agrochemical Service Station "Samarskaya".

We used 127 fields to estimate the coefficients of the Model 1 that describes the soil independent regression model. 20 fields were used to estimate the coefficients of the Model 2 that describes the soil

map-based model. And 32 fields were used to estimate the coefficients of the Model 3 that describes the model based on the agrochemical inspection data. The difference between the sample set size is explained by the various number of available data for each of the information sources used to extract the factors. The RMSE error was estimated for the same sample sets. The multiple linear regression coefficients estimation and the model adequacy verification was made using MATLAB software.

We obtained 9 models for the different BSFs listed in Table 1 and checked their adequacy. The significance level used to test adequacy was equal to $\alpha = 0.05$ for all tests. Table 2 shows the determination coefficients evaluated for each of these models. According to table 2, the significant factors of the models based on NDVI06 and NDVI02 BSFs should be taken as the optimal basic sets of factors. However, the model NDVI02 did not pass the Lilliefors test and, thus, it cannot be regarded as the statistically adequate model. The other models passed adequacy verification successfully. Therefore, we selected the model with NDVI04 BSF as an alternative to NDVI06 model.

Table 2. Determination coefficients of the models for BSFs listed in Table 1.

BSF name	Determination coefficient
NDVI03	0.7110
NDVI04	0.7162
NDVI05	0.6970
NDVI06	0.8035
NDVI20	0.7972
NDVI30	0.6999
NDVI40	0.6768
NDVI50	0.6405
NDVI60	0.7110

The further results were obtained for two OBSFs determined as the significant factors of the models based on NDVI04 BSF and NDVI06 BSF. The lists of factors included in the OBSF are given below:

1) for the model NDVI06: cumulative and maximum NDVI values during the vegetation period, the vegetation period duration and vegetation beginning date, maximum and minimum temperature during the vegetation period, maximum and minimum humidity during the vegetation period, seeds reproduction, sowing date and geographical coordinates of the field center;

2) for the model NDVI04: cumulative and maximum NDVI values during the vegetation period, the vegetation period duration and vegetation beginning date, maximum and minimum temperature during the vegetation period, maximum and minimum humidity during the vegetation period, the wheat cultivar, farmer identifier, sowing date and the geographical coordinates of the field center.

The other factors were excluded from the OBSF because they were insignificant. In further text, we will reference on OBSF according to the BSF name for which it was constructed.

To obtain soil-dependent models Model 2 and Model 3, we combined two OBSFs defined above with the SSF and TSF. Therefore, four different regression models were obtained. These models were passed the adequacy tests with the significance level $\alpha = 0.05$. The R^2 value and RMSE corresponding to the obtained soil-dependent models are given in Table 3.

Table 3. Determination coefficient R^2 and RMSE for soil-dependent models.

OBSF Name	Model 2 (OBSF+SSF)		Model 3 (OBSF+TSF)	
	R^2	RMSE	R^2	RMSE
NDVI06	0.9523	0.8866	0.8046	2.3931
NDVI04	0.9820	0.5447	0.8341	2.2051

Table 3 demonstrates that soil dependant models (Model 2 and Model 3) provide better R^2 value and RMSE than the soil-independent model (Model 1). In other words, including soil parameters

improves the crop yield forecasting quality. The best model quality corresponds to the Model 2 which applies SSF and OBSF based on NDVI04 remote sensing factors. This model has the minimum RMSE and maximum determination coefficient among the models considered.

Figures 3-8 illustrates actual and predicted yield values for the different types of regression models and OBSFs. The red color denotes the upper confidence interval for the predicted yield values. The green color denotes the lower confidence interval for the predicted yield values. The blue color denotes the yield value. The experimental results shows that the RMSE of modeling without soil parameters is in the range from 2.7 up to 3 centner per hectare and from 0.5 up to 2.4 centner per hectare with soil parameters. To summarize, we found that additional soil factors enhance the accuracy of regression modeling using NDVI data. We also recommend using soil fertility score and soil maps instead of agrochemical survey data. As for the remote sensing factors estimation, we propose to apply the first approach with the NDVI threshold equal to 0.4.

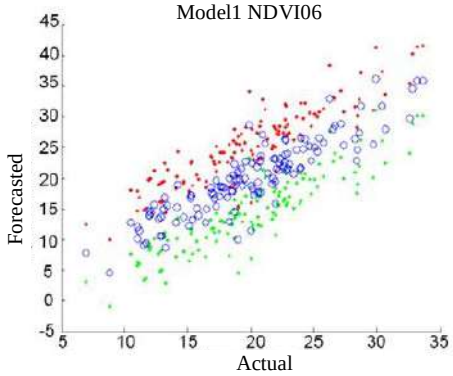


Figure 3. Diagram of actual and forecasted yield values for Model 1 with BSF NDVI06.

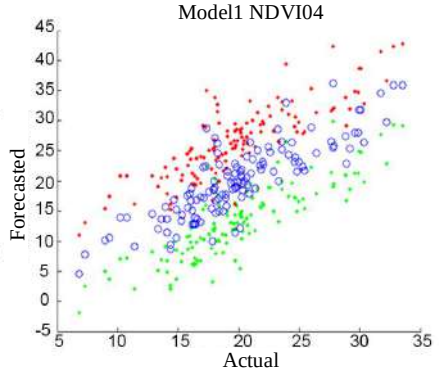


Figure 4. Diagram of actual and forecasted yield values for Model 1 with BSF NDVI04.

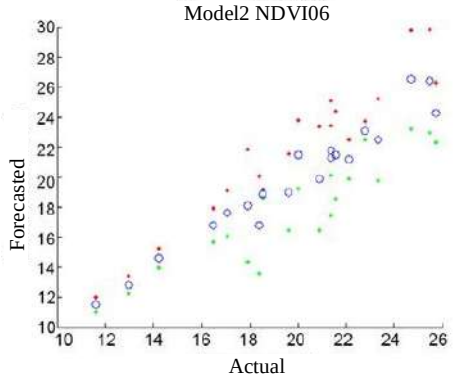


Figure 5. Diagram of actual and forecasted yield values for Model 2 with OBSF NDVI06.

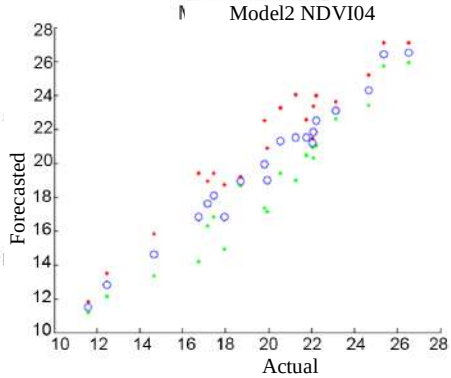


Figure 6. Diagram of actual and forecasted yield values for Model 2 with OBSF NDVI04.

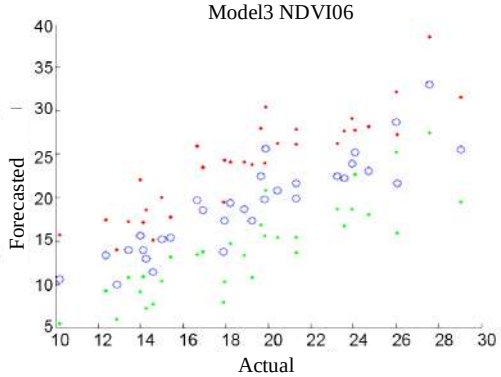


Figure 7. Diagram of actual and forecasted yield values for Model 3 with OBSF NDVI06.

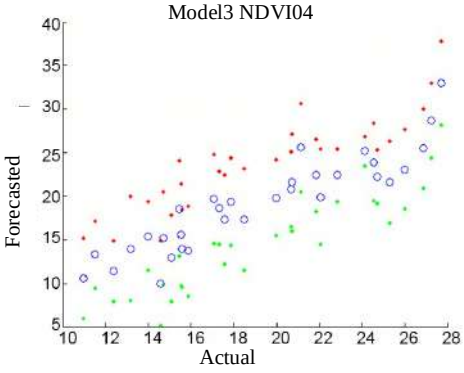


Figure 8. Diagram of actual and forecasted yield values for Model 3 with OBSF NDVI04.

5. Conclusion

The study presented in this article shows that the accuracy of winter wheat yield regression modeling using the NDVI vegetation index depends on the additional soil characteristics. We regarded several factor groups derived from the remote sensing data and the geoinformation systems to define the optimal basic factor set for the yield modeling without soil factors. Two types of additional soil factors were combined with the optimal basic factor sets, and comparative analysis of soil-dependent and soil -independent models was performed. The analysis was made for the sample fields in Samara region, Russia. As a result, we found that the soil parameters enhance the quality of winter wheat yield modeling and soil fertility score derived from a ten-year-old soil maps is provides more accurate results than the agrochemical inspection data. Therefore, only two parameters such as soil fertility score and soil type might for successfully winter wheat yield modeling instead of variety agrochemical characteristics. We also compared different approaches for the remote sensing factors computation using NDVI time series data. We concluded that the threshold 0.4 effectively determines the vegetation beginning date, vegetation period duration and cumulative NDVI in terms of the yield prediction accuracy. The results of this research can be applied to forecasting winter wheat yield in Samara region.

6. References

[1] Kayumov M K 1989 *Crop Yield Programming* (Moscow: Agropromizdat) p 368
[2] Chahbi A, Zribi M, Lili-Chabaane Z, Duchemin B, Shabou M, Mougenot B and Boulet G 2014 Estimation of the dynamics and yields of cereals in a semi-arid area using remote sensing and the SAFY growth model *International Journal of Remote Sensing* **35** 1004-1028
[3] Cabelguenne M, Debaeke P and Bouniols A 1999 EPICphase, a version of the EPIC model simulating the effects of water and nitrogen stress on biomass and yield, taking account of developmental stages: validation on maize, sunflower, sorghum, soybean and winter wheat *Agricultural Systems* **60** 175-196
[4] Ma G, Huang J, Wu W, Fan J, Zou J and Wu S 2013 Assimilation of MODIS-LAI into the WOFOST model for forecasting regional winter wheat yield *Mathematical and Computer Modelling* **58** 634-643
[5] Quarmby N A, Milnes M, Hindle T L and Silleos N 1993 The use of multi-temporal NDVI measurements from AVHRR data for crop yield estimation and prediction *International Journal of Remote Sensing* **14** 199-210
[6] Tregubov B A, Lobov G G and Holina M G 1976 *Soil fertility ranking in Kuybyshev region of Russia* (Kuybyshev: Knizhnoye isdatelstvo) p 111
[7] Kobzar A I 2006 *Applied mathematical statistics. For engineers and scientists* (Moscow: FIZMATLIT) p 816
[8] Vasin A V 2006 *Formation of highly productive multi-species agrophytocenoses of fodder crops in the Middle Volga region* (Kinel) p 513
[9] Vorobiova N S 2016 Crops identification by using satellite images and algorithm for calculating estimates *CEUR Workshop Proceedings* **1638** 419-427
[10] Vorobiova N S and Chernov A V 2016 NDVI time series modeling in the problem of crop identification by satellite images *CEUR Workshop Proceedings* **1638** 428-436
[11] Vorobiova N S, Sergeyev V V and Chernov A V 2016 Information technology of early crop identification by using satellite images *Computer Optics* **40(6)** 929-938 DOI: 10.18287/2412-6179-2016-40-6-929-938

Acknowledgements
This work was supported by the RFBR projects No. 16-29-09494 and 18-07-00748.

A public transport departure time prediction algorithm based on operation strategies and real-time monitoring data

A A Agafonov[1]

[1]Samara National Research University, Moskovskoye shosse 34, Samara, Russia, 443086

Abstract. In this paper, we consider a public transport departure time prediction problem. The problem is considered in two notations: estimation of the mean expected departure time and estimation of the time interval required to ensure a predefined probability to depart on-time. We propose departure time estimation algorithms based on a real-time monitoring data, schedule and operation strategies for the public transport management. We consider a schedule-based holding strategy and a bus bunching prevention strategy. We also propose an algorithm for departure time interval estimation. Numerical tests based on real bus routes in Samara, Russia are used for a comparative study of different departure time prediction algorithms.

1. Introduction

It is a challenge for public transportation agencies to provide reliable service because public transport often operates under conditions of high uncertainty. Variation in travel time caused by different factors of uncertainty, such as real-time traffic situation and traffic congestion, weather conditions, change of passenger flow, driver behavior [1]. The unreliability of the transport schedule reduces the efficiency of the transport infrastructure because road users have to take into account risks of arriving late while planning the route. This is especially important for the passenger transport, where the discrepancy with the schedule can accumulate and lead to the bus bunching problem.

In this paper we consider a passenger transport departure time prediction problem. We compare algorithms based on the schedule, real-traffic data and operation strategies of the public transport management.

There are many papers focused on the similar arrival time prediction problem. Proposed methodologies include

- Regression models [2] that constructed as a regression function from the set of independent variables. Non-parametric regression (NPR) models is a relatively simple method for prediction without the need to estimate parameters. K nearest neighbor (k-NN) methods are one of the most popular NPR methods. Bus travel time prediction models using k-NN was developed in [3, 4].
- Kalman filtering models [5, 6] are an efficient recursive procedure that estimates the future states of dependent variables. In [7] authors developed a path-based model and a link-based model using Kalman filter to predict bus travel times.

- Machine learning models, including artificial neural networks (ANN) and support vector machine (SVM) models. ANN has been reported to be especially useful for finding solutions for complex non-linear problems. In [8] authors proposed two ANN-based models to predict bus arrival time: the link-based ANN and the stop-based ANN. The paper [6] proposed a dynamic algorithm that integrated the ANN model and a Kalman filter-based algorithm. In [9] authors used Bayesian inference theory to combine neural networks. Their results showed that the ANN model outperformed the historical data based model and the regression model in terms of prediction accuracy. SVM is a very specific type of learning algorithm characterized by the capacity control of the decision function, the use of the kernel functions, and the sparse solution. The SVM-based models to predict bus arrival time was developed in [10–12].
- Hybrid methods [13, 14] that combine two or more models to predict arrival time precisely.

However, the described models use the real-time and historical travel time information as the basis to predict arrival time of the public transport and cannot be used directly to predict departure time of the vehicle from the starting station.

Another problem that needs to be considered in the context of this paper is the schedule design and operational management of the public transport [15]. The use of operational control strategies makes it possible to reduce the discrepancy between the scheduled time and the real time of arrival at the control points and prevent the bus bunching on the route. One of the most popular control strategies is the schedule-based holding control strategy: if a bus arrives at a stop earlier than the scheduled time, it is held until the scheduled departure time is reached [16]. Although this strategy can improve schedule adherence, it delays the operation of early buses and causes impatient among passengers on board. Another strategy is a driver schedule recovery strategy that proposes to adjust speed over segments between two consecutive control points to arrive on-time [17, 18]. However, this strategy requires that drivers highly flexible response to the traffic situation and can be limited by traffic conditions.

In this paper, we consider the problem of the public transport departure time prediction from the starting station. The problem is considered in two formulations: estimation of the mean expected departure time and minimizing the travel time interval required to ensure a predefined probability of departure on-time.

The structure of this paper is organized as follows: the second section introduces the basic notations and the problem formulation, describes the departure time prediction algorithms based on real-time traffic data and schedule. In this section, we propose an algorithm based on the operation control strategies. The third section presents the problem formulation of the departure time interval estimation and describes a base algorithm for solving this problem. In the fourth section, we present an experimental study of the algorithms. Lastly, we present our conclusions and recommendations for future work.

2. Mean expected departure time estimation

2.1. Basic notation

Introduce the following notation. Let
V be the set of public transport vehicles;
R be the set of routes;
I_r be the bus run numbers on the route $r \in R$;
DT_i be the departure time from the starting station of the bus with the run number $i \in I_r$ on the route $r \in R$;
DT_i^{sch} be the scheduled departure time from the starting station of the bus $i \in I_r$;
AT_i be the arrival time at the ending terminal of the bus $i \in I_r$;
AT_i^{sch} be the scheduled arrival time at the ending terminal of the bus $i \in I_r$.

Let Δ_{min} be the minimum recovery time at the ending terminal (time between two consecutive runs).

Denote the deviation of the real observed arrival time of the bus with the run number $i \in I_r$ from the scheduled arrival time as Δ_i^{arr}:

$$\Delta_i^{arr} = AT_i - AT_i^{sch}.$$

The waiting time between runs denote as Δ_i^{wait}:

$$\Delta_i^{wait} = DT_j^{sch} - AT_i^{sch},$$

where $(i, j), i \in I_r, j \in I_r$ is the numbers of consecutive runs of the bus $v \in V$ on the route $r \in R$.

The main problem is to predict the departure time from the starting station DT_i, i.e., the following assessment construction:

$$\hat{DT}_i, \forall i \in I_r, \forall r \in R. \tag{1}$$

In the next subsections, we describe algorithms for estimation the departure time $\hat{DT}_i$ based on the real-time monitoring information, schedule, and operation control strategies: a holding control strategy and a bus bunching prevention strategy.

2.2. Monitoring data based algorithm

The monitoring data based departure time prediction algorithm uses the real-time monitoring data taking into account the planned schedule. The main assumption of the algorithm is that the deviation of the arrival time at the ending terminal from the scheduled arrival time will be maintained when the vehicle departs from the starting station. The departure time estimation based on the monitoring data can be described as follows

$$\hat{DT}_j^M = DT_j^{sch} + \Delta_i^{arr}. \tag{2}$$

Obviously, this estimation is quite simple and does not consider any control strategies to decrease the deviation from the scheduled departure time. In the next subsections, we present departure time prediction algorithms based on schedule and operation strategies.

2.3. Schedule based algorithm

The schedule-based holding control strategy is one of the most popular strategies in the bus route schedule design problem. In this strategy, if a bus arrives early at the timing point, it will be held until the scheduled departure time is reached. If the bus is already late, it will depart from the timing point immediately after passenger pickup. We use the similar approach to adapt schedule-based holding strategy to estimate the departure time from the starting terminal.

The schedule-based algorithm with the holding strategy considers the deviation of the arrival time Δ_i^{arr} and the waiting (recovery) time Δ_i^{wait} to control the departure time in the following way:

$$\hat{DT}_j^S = \begin{cases} AT_i + \Delta_{min}, & \Delta_{wait} < 0, \\ DT_j^{sch} + \alpha\Delta_i^{arr}, & \Delta_i^{arr} < 0, \\ DT_j^{sch} + \max\left(0, \beta\Delta_i^{arr} - \gamma\Delta_i^{wait}\right), & \text{otherwise}, \end{cases} \tag{3}$$

where $\alpha \in [0, 1], \beta \in [0, 1], \gamma \in [0, 1]$ are the control strategy coefficients.

The proposed control strategy based algorithm consists of the following steps:

- if the vehicle arrived at the terminal stop later than the scheduled departure time of the next run, then shorten the waiting time to a minimum;
- if the vehicle arrived earlier than the scheduled arrival time, then reduce the deviation from the scheduled departure time by the recovery time increasing;
- otherwise, reduce the deviation from the scheduled departure time by reducing the recovery time.

2.4. Operation strategy based algorithm

The operation strategy based algorithm considers a minimum delay between departure times of the different buses on the same route to prevent bus bunching.

We assume that the minimum delay time between two consecutive runs can be different for different routes and is determined based on the route schedule.

Denote the minimum delay between the departure times as $\Delta_r^{route}, r \in R$.

Then the departure time estimation for the run number $j \in R$ of the vehicle $v \in V$ on the route $r \in R$ can be expressed as follows:

$$\hat{DT_j^{OS}} = \max(\hat{DT_j^S}, DT_k + \Delta_r^{route}), \tag{4}$$

where k is the last run number, $\hat{DT_j^S}$ is the departure time estimation obtained using the schedule-based algorithm.

Experimental study of the described departure time estimation algorithms is presented in section 4.

3. Departure time interval estimation

In the previous section, we describe the mean expected departure time estimation problem. Sometimes this formulation cannot be suitable for passengers. Often it is desirable to know not a time instant of the departure, but a departure time interval required to ensure a predefined probability of departing on-time.

Let τ be the required departure time instant, p be the required probability to depart on-time, r be the selected route.

The departure time interval estimation problem can be expressed in the following form:

$$P(\tau - \Delta^{dep} < DT_i < \tau) > p, \tag{5}$$

where DT_i is the departure time,
Δ^{dep} is the estimated departure time interval.

The base departure time Δ^{dep} estimation algorithm uses traffic statistics. Let $\Pi_r^d = \{DT_i, i \in I_r\}$ be the dataset with the departure time of the vehicles $v \in V$ on the routes $r \in R$ in the selected day d, $\Pi_r = \{\Pi_r^d\}_{d=1}^N$ be the statistics for N days.

We introduce an indicator variable of the following form:

$$p(d, \tau, \Delta) = \begin{cases} 1, & \exists DT_i \in \Pi_r^d : \tau - \Delta < DT_i < \tau; \\ 0, & \text{othewise}, \end{cases} \tag{6}$$

Then the departure time interval estimation algorithm can be expressed in the following form:

Algorithm 1: Departure time interval estimation

$\hat{p} = 0$;
$\Delta = 0$;
while $\hat{p} < p$ **do**
  $\Delta = \Delta + step$; // *step* - increment step
  $\hat{p} = \sum_{d=1}^{N} p(d, \tau, \Delta)/N$;
end
$\Delta^{dep} = \Delta$;

4. Simulation setup and results

An experimental study of the algorithms was carried out on the public transport data in Samara. To estimate the departure time, we chose the 1070 runs of different bus routes, for which the departure / arrival times to the control points according to the schedule are known.

Firstly, we estimate the deviation of the real observed arrival time to the terminal stop from the scheduled arrival time. Figure 1 shows the histogram of the arrival time deviation. The positive deviation indicates the delay in the arrival, and the negative deviation means that the vehicle arrives ahead of schedule. As can be seen from the histogram, vehicles often arrive at the ending terminal later than the scheduled time.

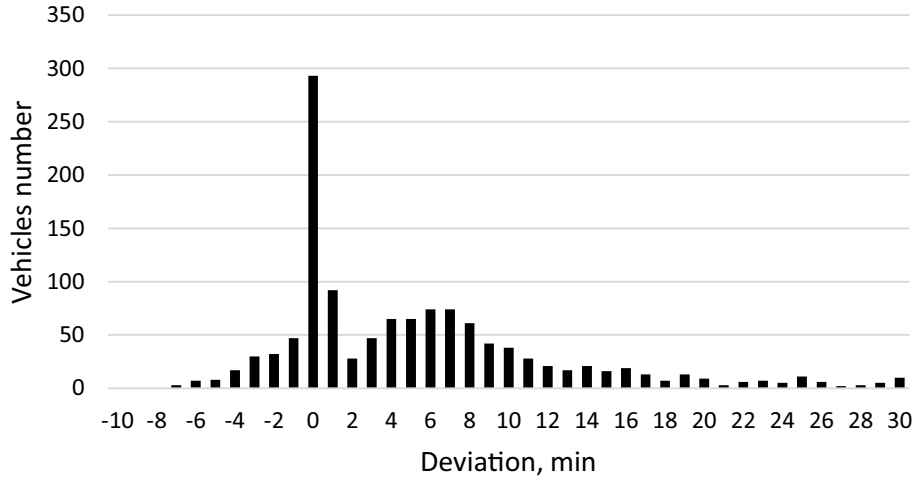


Figure 1. Arrival time deviation

Next, we compare the proposed algorithms based on monitoring data, schedule and operation strategies. Table 1 provides the mean absolute error and standard deviation of the proposed algorithms.

Table 1. Algorithms comparison

	Mean absolute error	Standard deviation
Monitoring data based algorithm	601.0	1147.35
Schedule based algorithm	249.95	385.22
Operation strategies based algorithm	260.85	377.78

The schedule-based algorithm with the holding strategy provides the best results by the selected criteria. Using bus bunching prevention control strategy do not improve the departure time estimation precision.

At the final stage of the experimental study, we estimate the deviation of the real observed departure time from the estimated departure time. Figure 2 depicts the histogram of the departure time deviation. A positive deviation means that the vehicle departed after the estimated departure time.

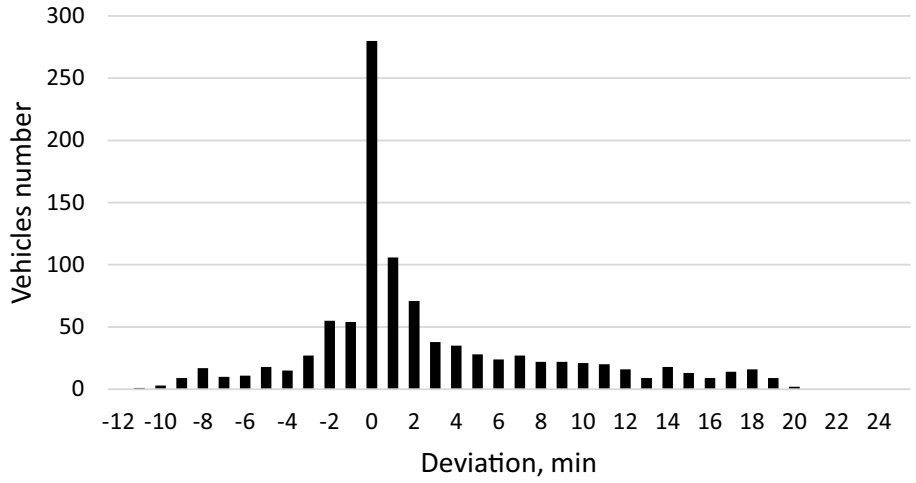


Figure 2. Estimated departure time deviation

As can be seen from the histogram, the schedule-based algorithm estimates the departure time without deviation from the observed time.

5. Conclusion

In this paper, we considered the problem of public transport departure time estimation. We compared the algorithm based on the real-time monitoring, the schedule-based algorithm with the holding operation strategy, and the operation strategy based algorithm with bus bunching prevention strategy. The schedule-based algorithm with the holding control strategy showed the best results by mean absolute error criteria. Using bus bunching prevention strategy did not decrease the error of the estimation.

The main disadvantage of the proposed algorithm is the fact that the traffic statistics are not used to estimate the departure time of the public transport.

In this paper, we considered a public transport departure time prediction problem. The problem is considered in two notations: estimation of the mean expected departure time and estimation of the time interval required to ensure a predefined probability to depart on-time. We compare algorithms based on a real-time traffic data and public transport schedule. The paper proposes an original prediction algorithm based on the operation strategies of the public transport management. Numerical tests based on real bus routes in Samara, Russia are used for a comparative study of different departure time prediction algorithms.

In addition, we consider the problem of the time interval estimation required to ensure a predefined probability to depart on-time. The base algorithm for solving this problem is proposed. The development of more complex algorithms will be conducted as our future research.

6. References

[1] Chen X, Yu L, Zhang Y and Guo J 2009 Analyzing urban bus service reliability at the stop, route, and network levels *Transportation Research Part A: Policy and Practice* **43** 722-734
[2] Agafonov A, Sergeev A and Chernov A 2012 Forecasting of the motion parameters of city transport by satellite monitoring data *Computer Optics* **36** 453-458
[3] Chang H, Park D, Lee S, Lee H and Baek S 2010 Dynamic multi-interval bus travel time prediction using bus transit data *Transportmetrica* **6** 19-38
[4] Smith B, Williams B and Keith Oswald R 2002 Comparison of parametric and nonparametric models for traffic flow forecasting *Transportation Research Part C: Emerging Technologies* **10** 303-321
[5] Vanajakshi L, Subramanian S and Sivanandan R 2009 Travel time prediction under heterogeneous traffic conditions using global positioning system data from buses *IET Intelligent Transport Systems* **3** 1-9
[6] Chen M, Liu X, Xia J and Chien S 2004 A dynamic bus-arrival time prediction model based on apc data *Computer-Aided Civil and Infrastructure Engineering* **19** 364-376
[7] Chien S J and Kuchipudi C 2003 Dynamic travel time prediction with real-time and historic data *Journal of Transportation Engineering* **129** 608-616
[8] Chien S J, Ding Y and Wei C 2002 Dynamic bus arrival time prediction with artificial neural networks *Journal of Transportation Engineering* **128** 429-438
[9] van Hinsbergen C, van Lint J and van Zuylen H 2009 Bayesian committee of neural networks to predict travel times with confidence intervals *Transportation Research Part C: Emerging Technologies* **17** 498-509
[10] Yu B, Lam W and Tam M 2011 Bus arrival time prediction at bus stop with multiple routes *Transportation Research Part C: Emerging Technologies* **19** 1157-1170
[11] Bin Y, Zhongzhen Y and Baozhen Y 2006 Bus arrival time prediction using support vector machines *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations* **10** 151-158
[12] Yu B, Yang Z and Wang J 2010 Bus travel-time prediction based on bus speed *Proceedings of the Institution of Civil Engineers: Transport* **163** 3-7
[13] Park T and Lee S 2004 A bayesian approach for estimating link travel time on urban arterial road network *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **3043** 1017-1025
[14] Agafonov A and Myasnikov V 2014 An algorithm for city transport arrival time estimation using adaptive elementary predictions composition *Computer Optics* **38**(2) 356-368
[15] Wu Y, Tang J and Luo X 2015 Comparative analysis of operation strategies in schedule design for a fixed bus route *International Transactions in Operational Research* **22** 545-562
[16] Cats O, Larijani A, Ólafsdóttir A, Burghout W, Andréasson I and Koutsopoulos H 2012 Bus-holding control strategies *Transportation Research Record* 100-108
[17] Kalaputapu R and Demetsky M J 1995 Modeling schedule deviations of buses using automatic vehicle-location data and artificial neural networks *Transportation Research Record* 44-52
[18] Yan Y, Meng Q, Wang S and Guo X 2012 Robust optimization model of schedule design for a fixed bus route *Transportation Research Part C: Emerging Technologies* **25** 113-121

Acknowledgments
This work was supported by the Russian Foundation for Basic Research (RFBR) grant 18-07-00605 A, grant 18-29-03135.

Accuracy analysis of 3D object reconstruction using RGB-D sensor

A N Ruchay[1,2], K A Dorofeev[2], A V Kober[1]

[1]Federal Research Centre of Biological Systems and Agro-technologies of the Russian Academy of Sciences, 9 Yanvarya street 29, Orenburg, Russia, 460000
[2]Department of Mathematics, Chelyabinsk State University, Bratiev Kashirinykh street 129, Chelyabinsk, Russia, 454001

Abstract. In this paper, we propose a new method for 3D object reconstruction using an RGB-D sensor. The RGB-D sensor provides RGB images as well as depth images. Since the depth and RGB color images are captured with one sensor of an RGB-D camera placed in different locations, the depth image should be related to the color image. After matching of the images (registration), point-to-point corresponding between two images is found, and they can be combined and represented in the 3D space. In order to obtain a dense 3D map of the 3D object, we design an algorithm for merging information from all used cameras. First, features extracted from color and depth images are used to localize them in a 3D scene. Next, Iterative Closest Point (ICP) algorithm is used to align all frames. As a result, a new frame is added to the dense 3D model. However, the spatial distribution and resolution of depth data affect to the performance of 3D scene reconstruction system based on ICP. The presented computer simulation results show an improvement in accuracy of 3D object reconstruction using real data.

1. Introduction

The 3D reconstruction of objects is a popular task, with applications in the field of medicine, architecture, games, agriculture, and film industry. The 3D reconstruction has many applications in object recognition, object retrieval, scene understanding, object tracking, autonomous navigation, human-computer interaction, telepresence, telesurgery, reverse engineering, virtual maintenance and visualization [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. The geometry of an object can be reconstructed from laser range scans, a set of different photos, monocular cameras and stereo cameras; each technology possesses disadvantages and limitations. RGB-D cameras like Microsoft Kinect or ASUS Xtion Pro are sensing systems equipped with an RGB camera, infrared projector, and infrared sensor. They can gather RGB information and the depth map simultaneously.

The ICP (Iterative Closest Point) is one of the most commonly used methods for pairwise alignment, where the rotation and translation to align two point clouds are determined. Many variants on the base ICP algorithm were proposed because of the limitations in the ICP algorithm: different cost functions [11] and others [12]; two-Pass ICP with color constraint to improve the error minimization process [13]; 3D representation using a set of planes to perform the registration [14]; point-to-plane matchings instead of point-to-point matchings in standard

ICP algorithm [15]; a new registration algorithm based on the Matching Signed Distance Fields [16].

In order to fill small holes and to eliminate noise, the median and binomial filters were used [17, 18, 19, 20, 21, 22]. Moreover, the use of the color information in the point correspondence process avoids false positives matches and, therefore, leads to a more reliable registration. Note that by adjusting ICP and reconstruction parameters it is possible to improve the registration and appearance of details that were invisible with just one scan due to the sensor limited precision. Finally, it was shown that with help of low precision sensors as Kinect 3D smooth surface of objects can be reconstructed [23].

A new low-cost approach to reconstruct real-time of a 3D object with Kinect sensor uses a SLAM algorithm (Simultaneous Localization and Mapping) [24]. SLAM provides an approximated solution of 3D reconstruction because the accuracy of the system often depends on a heuristic algorithm for obtaining relevant reference points.

In order to improve the accuracy and robustness of the ICP algorithm, a regularization by incorporating the spatial distances of SIFT feature pairs with dynamically adjusted weights to balance errors is proposed in [12]. A new outlier rejection method based on a dynamic thresholding and leverages the structure and sparse feature pairs from the texture of the RGB images were also suggested [12]. A new method to robustly estimates the camera motion in a dynamic environment based on RANSAC algorithm was proposed in the article [25].

We propose to utilize a graph-based SLAM algorithm [24] with loop closure detection using dense color and depth images obtained from the RGB-D camera. We show that our system is able to perform the SLAM for three-dimensional modeling in real-time.

The paper is organized as follows. Section 2 discusses related works. Section 3 describes the proposed system for object 3D reconstruction using an RGB-D sensor. In Section 4 the experimental results are discussed. Finally, Section 5 presents our conclusions.

2. Related work

An approach for realistic surface geometry reconstruction using high-frequency features of color images from Kinect is given in [26]. In order to achieve a better accuracy in the global alignment of scans, it was used a weighted ICP algorithm. An efficient 3D reconstruction approach combining a depth information based 3D model with RGB information to refine the reconstruction results when the camera fails to acquire the correct depth information is presented in [27].

The efficiency problems and real-time processing are discussed in [28]; that is, highly resource consuming of dense point cloud registration algorithm based on an ICP method and slow surfel update procedure for sufficiently large data clouds. For resolving these problems a sparse ICP algorithm a frustum culling procedure exploits the hierarchical map structure. A novel real-time algorithm for simultaneously reconstructing the geometry using a single RGB-D camera is suggested in [29]. Moreover, this method can provide global anchors by including sparse SIFT features. It allows us to potentially achieve better results for loop-closing motions.

A patch-based illumination invariant visual odometry was proposed in [30], which works well in the irregular illumination change. The planar patch selection process is employed and the illumination change model is adopted in each extracted patch in order to consider the partial light variations. Using the robust weighting function and the efficient second-order minimization (ESM) image alignment method, the proposed cost function reflecting the illumination changes is minimized. As a result, the proposed method can accurately estimate the motion of the camera regardless of the partial lighting changes. A method for visualization of occluded objects using multiple Kinect sensors at different locations is proposed in [31].

A novel robust 3D reconstruction system with an RGB-D camera was proposed in [32]. The visual and geometry features and combined SFM technique to make registration more robust

especially in depth map missing cases were used. In order to solve the drift problem, 3D information is used to detect the loop closure and to perform global refinement.

Prior-based Multi-Candidates RANSAC (PMCSAC) algorithm was proposed [32] to make the feature matching more robust and efficient in order to handle the repeated textures/structures. Missing geometry due to depth missing can be effectively completed by combining multi-view stereo and mesh deformation techniques [32].

For solving the problem of a wider range of object deformations merging a sequence of images from a single range sensor into a unified 3D model, without requiring an initial template was proposed in [33]. However, although complex scene topologies can be handled, a regular topology is restricted to be constant throughout the sequence. So, if the coarse-scale reconstruction does not correctly choose the topology, it cannot switch to a fine scale, and the computational cost is also high. A novel approach to the template-driven capture of dense detailed non-rigid deformations from video sequences to carry out simultaneously 2D dense registration and 3D shape inference is presented in [34].

In work [35] a model-based scalable 3D scene reconstruction system based on a CAD model was proposed. The accuracy of KinectFusion algorithm and analysis the noise effect of reconstruction and localization errors based on a CAD model was tested in [36].

An iterative low-cost method for 3D body registration, dealing with unconstrained movements and accuracy is suggested in [37]. A novel method for 3D object reconstruction from RGB-D data that applies sub mapping to 3D bundle adjustment is presented in [38]. A 3D shape descriptor for object recognition with RGB-D sensors is proposed in [39].

A novel volumetric multi-resolution mapping system for RGB-D images was proposed in the paper [40]. Proposed approach generates a textured triangle mesh from a signed distance function that it continuously updates as new RGB-D images arrive. For this, an octree uses as the primary data structure which allows us to represent the scene at multiple scales, and it allows us to grow the reconstruction volume dynamically.

A system for autonomous flight using RGB-D sensors was suggested [41]. It uses a combination of visual odometry techniques and mapping and is able to conduct all sensing and computation required for local position control.

A method to estimate both odometry and scene flow with RGB-D cameras is presented in the paper [42]. The main advantage of the proposed approach is that it provides accurate results with a very low runtime.

A new visual odometry system based on the alignment between consecutive frames by minimization both on the photometric and geometric error was proposed [43], where the original ICP algorithm for frame alignment and visual odometry computation was completely substituted by the proposed method. The use of the inverse depth instead of the depth to parametrize the geometric error is the main contribution of the proposed method.

A new RGB-D based method was proposed to improve the accuracy and robustness of visual odometry [44], where the set of line segments generate from maximum-clique filtered point correspondences.

In the paper [45], a robust RGB-D DVO algorithm BaMVO was proposed for use in a dynamic environment, where consecutive depth images were warped using the preobtained ego-motions to equalize the viewpoints. The background image was estimated using a nonparametric model from the differences between consecutive pairs in the warped depth images. The energy function between successive RGB-D images was represented by the form of the weighted least squares for the calculation of DVO.

A novel 3D geometry enhanced superpixel algorithm for RGB-D data is presented in [46]. The depth map is converted into 3D geometrical information, and superpixels are iteratively clustered according to a distance metric designed from the color information and 3D geometry. By introducing the geometrical information, the proposed superpixel method overcomes the

difficulty in distinguishing adjacent objects with similar colors.

3. The proposed system

This section provides the proposed system with a fusion of information from a moving RGB-D sensor for object 3D reconstruction.

3D reconstruction with our system consists of the following steps:

(i) Registration a point cloud PC_i using RGB and Depth data.
(ii) Detection and matching of global key points in RGB data of PC_i and PC_{i-1} with SURF algorithm.
(iii) Remove outliers with RANSAC.
(iv) Count transformation matrix with ICP using the associate 3D points of the inliers.
(v) Apply hybrid approach that combines ICP odometry with the RGB-D odometry.
(vi) Adding results to a general model (dense 3D map).

The ICP odometry has a problem in tracking drifting or failure for the presence of smooth surfaces [47]. Thus, we have implemented a hybrid approach that combines ICP odometry with the RGB-D odometry, which aims to minimize the photometric error between consecutive RGB-D frames [48]. The hybrid approach combines frame-to-model registration with the increased stability provided by the use of both geometric and photometric cues [49].

The hybrid odometry approach estimates transformation T_i for each frame i by minimizing the following function:

$$R(T_i) = R_{ICP}(T_i) + \lambda R_{RGBD}(T_i),$$

where $R_{ICP}(T_i)$ is based on the point-to-plane error used in ICP algorithm:

$$R_{ICP}(T_i) = \sum_{(p,q)\in K} \|(p - T_i q)^\top n_p\|^2,$$

with n_p is the surface normal at p and K is a set of corresponding point pairs found by projective data association. $R_{RGBD}(T_i)$ is the photometric error between frames $i-1$ and i given as

$$R_{RGBD}(T_i) = \sum_x \|I_i(\pi(T_i^{-1} T_{i-1} \pi^{-1}(x, D_{i-1}))) - I_{i-1}(x)\|^2,$$

where D_k is the depth image from frame k, x passes over all pixels in frame $i-1$, π is a projection operator that projects a 3D point in the camera coordinate frame to the image data, π^{-1} is an operator, that produces a 3D point in the camera coordinate frame that corresponds to a given pixel, and I_k is the intensity of x pixel in color image from frame k. The coefficient λ balances the two counted errors terms. It is calculated empirically and is identical for all reconstructions.

4. Experimental results

In this section, we present experimental results for evaluation of the performance of the proposed system with a fusion of information from an RGB-D sensor for object 3D reconstruction.

The metric of evaluation is the root mean square error (RMSE) of measurements.

$$RMSE = \sqrt{(E(ED - RD)^2)},$$

where ED is the estimated measurement by a device and RD is the real known measurement of the object.

The object to be mapped is a small box (Kinect V2 adapter box). Fig. (1) shows a real-world object and one frame from RGB-D device (Asus Xtion Pro).

Figure 1. Real world object and one frame from the RGB-D sensor.

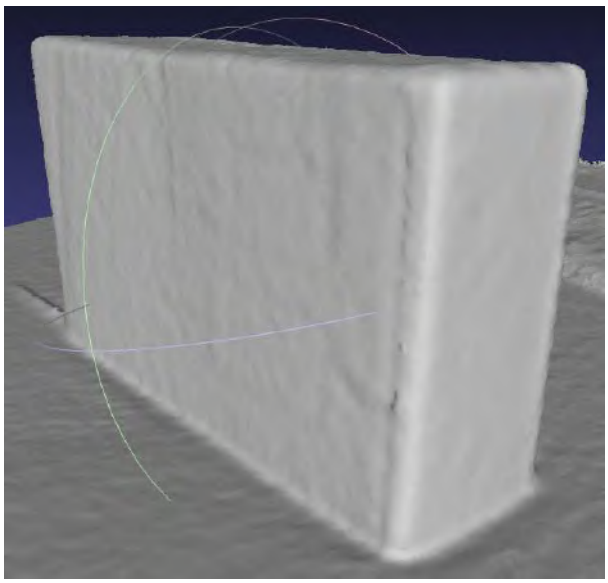


Figure 2. 3D model of the object.

Using the iterative ICP algorithm we obtain the corresponding 3D model of the object as shown in Fig. 2.

Linear distances between 2 points in the point cloud are measurements shown in Fig. 3. The geodesic distance between two points passing through surfaces and the surface area is also shown in Fig. 3.

The average values over five measurements acquired by our system. Corresponding RMSE values calculated for two sensors (Asus Xtion Pro and Kinect V2) are shown in Table 1.

The results show that Kinect V2 yields a more accurate 3D model of the object. The obtained accuracy allows us to make all measurements on the 3D model as on a real object.

5. Conclusion

We proposed a system for reconstruction of the 3D model of a real-life object with a small error in measurements. We have also developed a software to perform arbitrary measurements on

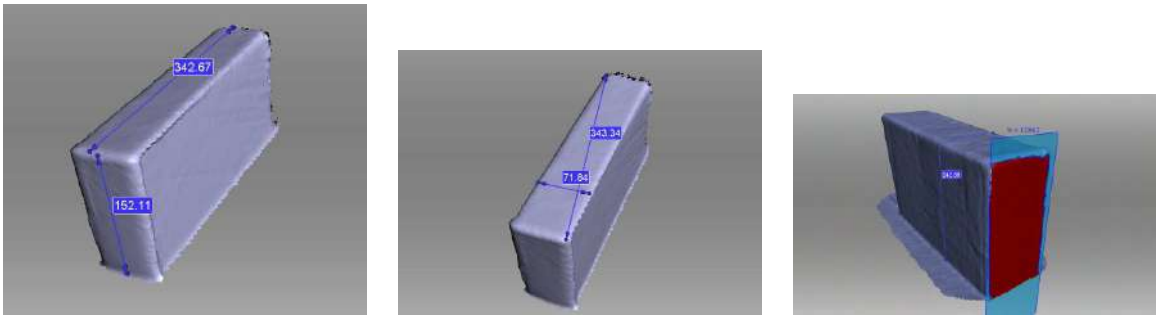


Figure 3. Measurements of the box.

Table 1. Results of measurements

	Real World	AsusXtionPro	RMSE	Kinect V2	RMSE
Length	345	342.67	2.33	344	1.00
Width	76	71.84	4.16	75.6	0.40
Height	158	152.11	5.89	155.87	2.13
Diagonal	352	343.34	8.66	348.78	3.22
Geodetic	228	242.68	14.68	234.85	6.85
Area	12008	12862	854.00	12003	5.00

the 3D model of the object. Our system is sensitive to poor illumination. The obtained results showed that it is possible to use multiple snapshots (sequence of frames) from low precision sensors such as Kinect for accurate reconstruction of the 3D model of objects. The accuracy of 3D reconstructed models is confirmed by the measured distances: linear (distance between two points), geodetic (distance over surfaces), and surface area.

In future, we improve the performance of the proposed algorithms to get more accurate 3D models of objects. We also develop more complicated measurement methods such as a section of an object, calculation of circle perimeter and section area, calculation of arbitrary geodetic distances and curvature of the lines.

6. References

[1] Echeagaray-Patron B A, Miramontes-Jaramillo D and Kober V 2015 *International Conference on Computational Science and Computational Intelligence (CSCI)* 843-844
[2] Echeagaray-Patron B A and Kober V 2015 *Proc. SPIE* **9598** 95980V
[3] Sochenkov I and Vokhmintsev A 2015 *Procedia Engineering* **129** 440-445
[4] Vokhmintsev A, Makovetskii A, Kober V, Sochenkov I and Kuznetsov V 2015 *Proc. SPIE* **9599** 959929
[5] Tihonkih D, Makovetskii A and Kuznetsov V 2016 *Proc. SPIE* **9971** 99712D
[6] Sochenkov I, Sochenkova A, Vokhmintsev A, Makovetskii A and Melnikov A 2016 *Proc. SPIE* **9971** 997124
[7] Picos K, Diaz-Ramirez V, Kober V, Montemayor A and Pantrigo J 2016 *Optical Engineering* **55** 55-55-11
[8] B Adriana and Echeagaray-Patron V K 2016 *Proc. SPIE* **9971** 9971-9976
[9] Echeagaray-Patrón B A, Kober V I, Karnaukhov V N and Kuznetsov V V 2017 *Journal of Communications Technology and Electronics* **62** 648-652
[10] Smelkina N, Kosarev R, Nikonorov A, Bairikov I, Ryabov K, Avdeev A and Kazanskiy N 2017 *Computer Optics* **41(6)** 897-904 DOI: 10.18287/2412-6179-2017-41-6-897-904
[11] Xie J, Hsu Y F, Feris R S and Sun M T 2013 *IEEE International Symposium on Circuits and Systems* 2904-2907
[12] Xie J, Hsu Y F, Feris R S and Sun M T 2015 *Journal of Visual Communication and Image Representation* **32** 194-204
[13] Rhee S M, Lee Y B and Lee H E 2014 *IEEE International Conference on Consumer Electronics* 89-90
[14] Thomas D and Sugimoto A 2013 *IEEE International Conference on Computer Vision* 2800-2807
[15] Chen Y and Medioni G 1991 *Proceedings IEEE International Conference on Robotics and Automation* **3** 2724-2729
[16] Masuda T 2002 *Computer Vision and Image Understanding* **87** 51-65
[17] Aguilar-Gonzalez P M and Kober V 2011 *Optical Engineering* **50** 50-59

[18] Aguilar-Gonzalez P M and Kober V 2012 *Optics Communications* **285** 574-583
[19] Ruchay A and Kober V 2016 *Proc. SPIE* **9971** 99712Y
[20] Ruchay A and Kober V 2017 *Proc. SPIE* **10396** 10396
[21] Ruchay A and Kober V 2017 *Proc. SPIE* **10396** 10396
[22] Ruchay A and Kober V 2018 *Analysis of Images, Social Networks and Texts* (Cham: Springer International Publishing) 280-291
[23] Takimoto R Y, de Sales Guerra Tsuzuki M, Vogelaar R, de Castro Martins T, Sato A K, Iwao Y, Gotoh T and Kagei S 2016 *Mechatronics* **35** 11-22
[24] Aguilar W G, Rodríguez G A, Álvarez L, Sandoval S, Quisaguano F and Limaico A 2017 *Real-Time 3D Modeling with a RGB-D Camera and On-Board Processing* (Cham: Springer International Publ.) 410-419
[25] Dib A and Charpillet F 2015 *International Conference on Advanced Robotics* 1-7
[26] Lee K R and Nguyen T 2016 *Mach. Vision Appl.* **27** 377-385
[27] Pan H, Guan T, Luo Y, Duan L, Tian Y, Yi L, Zhao Y and Yu J 2016 *Neurocomputing* **175** 644-651
[28] Wilkowski A, Kornuta T, Stefanczyk M and Kasprzak W 2016 *Applied Mathematics and Computer Science* **26** 99-122
[29] Guo K, Xu F, Yu T, Liu X, Dai Q and Liu Y 2017 *ACM Trans. Graph.* **36** 32:1-32:13
[30] Kim P, Lim H and Kim H J 2015 *IEEE/RSJ International Conference on Intelligent Robots and Systems* 3688-3694
[31] Nasrin T, Yi F, Das S and Moon I 2014 *Proc. SPIE* **9117** 9117-9115
[32] Wang K, Zhang G and Bao H 2014 *IEEE Transactions on Image Processing* **23** 4893-4906
[33] Dou M, Taylor J, Fuchs H, Fitzgibbon A and Izadi S 2015 *IEEE Conference on Computer Vision and Pattern Recognition* 493-501
[34] Yu R, Russell C, Campbell N D F and Agapito L 2015 *IEEE International Conference on Computer Vision* 918-926
[35] Cheng S C, Su J Y, Chen J M and Hsieh J W 2017 *Model-Based 3D Scene Reconstruction Using a Moving RGB-D Camera* (Cham: Springer International Publishing) 214-225
[36] Jiang S Y, Chang N Y C, Wu C C, Wu C H and Song K T 2014 *IEEE International Conference on Automation Science and Engineering* 1020-1025
[37] Villena-Martinez V, Fuster-Guillo A, Saval-Calvo M and Azorin-Lopez J 2017 *3D Body Registration from RGB-D Data with Unconstrained Movements and Single Sensor* (Cham: Springer International Publishing) 317-329
[38] Maier R, Sturm J and Cremers D 2014 *German Conference on Pattern Recognition (GCPR)* **8753** 54-65
[39] Liu Z, Zhao C, Wu X and Chen W 2017 *Sensors* **17** 451
[40] Steinbrucker F, Sturm J and Cremers D 2014 *IEEE International Conference on Robotics and Automation* 2021-2028
[41] Huang A S, Bachrach A, Henry P, Krainin M, Maturana D, Fox D and Roy N 2017 *Visual Odometry and Mapping for Autonomous Flight Using an RGB-D Camera* (Cham: Springer International Publishing) 235-252
[42] Jaimez M, Kerl C, Gonzalez-Jimenez J and Cremers D 2017 *IEEE International Conference on Robotics and Automation* 3992-3999
[43] Gutierrez-Gomez D, Mayol-Cuevas W and Guerrero J 2016 *Robot. Auton. Syst.* **75** 571-583
[44] Zhang Y, Hou Z, Yang J and Kong H 2016 *23rd International Conference on Pattern Recognition* 2764-2769
[45] Kim D H and Kim J H 2016 *IEEE Transactions on Robotics* **32** 1565-1573
[46] Yang J, Gan Z, Gui X, Li K and Hou C 2013 *3-D Geometry Enhanced Superpixels for RGB-D Data* (Cham: Springer International Publishing) 35-46
[47] Newcombe R A, Izadi S, Hilliges O, Molyneaux D, Kim D, Davison A J, Kohli P, Shotton J, Hodges S and Fitzgibbon A 2011 *Proceedings of the 10th IEEE International Symposium on Mixed and Augmented Reality* 127-136
[48] Kerl C, Sturm J and Cremers D 2013 *IEEE International Conference on Robotics and Automation* 3748-3754
[49] Choi S, Zhou Q Y, Miller S and Koltun V 2016 *arXiv:1602.02481*

Acknowledgments

This work was supported by the Russian Science Foundation, grant no. 17-76-20045.

Application of spatial interpolation methods for restoration of partially defined images

Y D Vybornova[1]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086

Abstract. The purpose of this paper is to analyze the practical applicability of spatial interpolation methods in the task of estimating the missing pixels of partially defined images. The paper presents three methods of spatial interpolation: inverse distance weighting, interpolation based on a triangulated irregular network, and kriging. The results of experimental research on these methods are given. Experiments show that all methods demonstrate a high quality of pixel prediction, but the choice of the most appropriate method directly depends on the input data.

1. Introduction

Broadly speaking, interpolation is used to obtain intermediate values from a discrete set of known values. In the problem of image processing, the interpolation methods are used to predict the raster cell values when having a limited number of the input data points [1].

The interpolation techniques known today can be classified into deterministic and statistical [2]. Deterministic methods are based on a function of distance or area. The main advantage of such methods is high processing speed. Statistical methods are based on a function of spatial similarity. The main advantage of these methods is the sensitivity to the multidirectional data. That is why such methods are commonly used to interpolate various kinds of surfaces (for example, to create a digital terrain model).

In this paper, three methods of spatial interpolation are considered: inverse distance weighting, interpolation based on a triangulated irregular network, and kriging. These methods are used extensively in a wide variety of applied sciences, including geology, hydrology, meteorology and oceanography. However, in relation to image processing field, these methods have not been investigated in detail yet [3].

Our research is devoted to the applicability analysis of the above-mentioned spatial interpolation methods in the task of estimating the missing pixels of partially defined images, i.e. images with a predetermined fraction of unknown pixels.

2. Methods of spatial interpolation

2.1. Inverse Distance Weighting

Inverse distance weighting (IDW) is a deterministic algorithm based on the assumption that the predicted value is more influenced by the nearby points than the points located further [4].

Interpolation is performed using the known values from a neighborhood of a given point. It is assumed that each point with a known value (hereinafter, called the reference point) has a local effect

decreasing with distance. The points located closer to the estimated position are assigned a greater weight, than those located further:

$$z(s_0) = \sum_{i=1}^{m} w_i z(s_i) = \frac{\sum_{i=1}^{m} z(s_i) d_{0i}^{-p}}{\sum_{j=1}^{m} d_{0j}^{-p}},$$

where $z(s_0)$ is an estimated value of a point in a certain location s_0, and $z(s_1), z(s_2), \ldots, z(s_n)$ – reference point values.

The weights are proportional to the inverse distance taken to the power of *p*. Consequently, an increase in distance leads to the fast decrease of weight. The weight decrease rate depends on the value of *p*. Thus, for *p=0*, the weights w_i are the same, and the predicted value is the average of all measured points. As *p* increases, weights of the distant points start to decrease rapidly. If the value of *p* is large, only the nearest neighborhood points will affect the predicted value.

To speed up the calculations, the weights of the most distant points with little effect can be taken as zero. It is common practice to limit the number of reference points, used to predict an unknown value, by specifying the search area. In this paper, the search area is represented by a circle of a variable radius *r*.

2.2. Kriging

Kriging [5] is a statistical interpolation method for prediction of unknown values using the values of the nearby points. Similar to the IDW method, weights are assigned to each point in accordance with the distance to the unknown value. However, here the estimation is performed on the basis of data correlation.

To date, research on kriging in image processing tasks is limited to several publications [6-8].

The basic kriging formula is:

$$z(s_0) = \sum_{i=1}^{n} \lambda_i z(s_i),$$

where $z(s_1), z(s_2), \ldots, z(s_n)$ are the reference points, and $z(s_0)$ is a value of the point to be evaluated. Also, it should be noted that $\sum_{i=1}^{n} \lambda_i = 1$.

The main task is to determine the weights λ_i so as to minimize the variance of the estimate, taking into account the unbiasedness requirement $E\{z(s_0) - z(s)\} = 0$.

There are several methods of kriging, which differ in the way of obtaining the weight components λ_i. In this paper, the method of ordinary kriging is considered. This type of kriging is the most common for spatial data modeling, and is considered the best as it minimizes the estimation error variance.

The estimation process begins with the construction of an empirical variogram, for all pairs of locations separated by distance *h*:

$$V(h) = \frac{average((z(s_i) - z(s_j))^2)}{2}.$$

After the spatial description is obtained, in order to make a prediction we need to determine the most appropriate variogram model (a curve that models the empirical variogram trends). In this paper, the following models are considered:

1) linear: $\gamma(h) = c_0 + c_1 t$;
2) circular: $\gamma(h) = c_0 + c_1 \frac{2}{\pi}\left(t\sqrt{1-t^2} + \arcsin t\right)$;

3) spherical: $\gamma(h) = c_0 + c_1\left(\frac{3}{2}t - \frac{1}{2}t^3\right)$;
4) exponential: $\gamma(h) = c_0 + c_1\left(1 - e^{-3t}\right)$;
5) gaussian: $\gamma(h) = c_0 + c_1\left(1 - e^{-3t^2}\right)$;
6) stable: $\gamma(h) = c_0 + c_1\left(1 - e^{-3t^w}\right), 0 < w \le 2$.

Here, $t = \frac{h}{a}$, a is a range of influence, c_0 is a nugget, $c_0 + c_1$ is a threshold.

The characteristics for model description are shown in figure 1. As can be seen from the figure, the nugget is a point at which semivariogram intersects the ordinate axis, and the threshold is the value at which the model starts to equalize. The distance to the threshold is called the range of influence. There is no spatial autocorrelation among points outside the range.

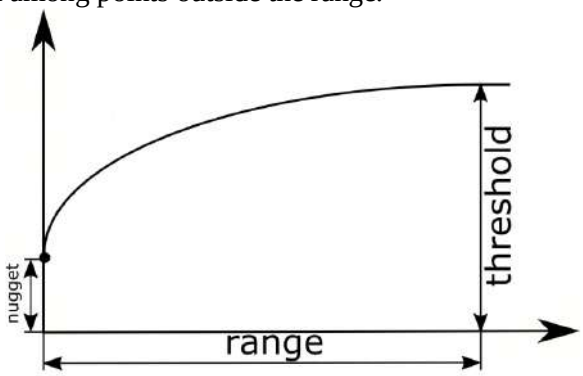


Figure 1. Range, threshold, and nugget.

The selected model allows to assign weights for the reference points and to estimate the unknown values. It should be noted that there is no universal model suitable for all input data. As a rule, the model is selected experimentally.

2.3. Triangulated Irregular Network

The interpolation method based on a triangulated irregular network (TIN) uses triangulation of data points to produce a two-dimensional function estimating the unknown values within each triangle [9].

Generally, an irregular network is obtained using the Delaunay triangulation [10]: points are connected with line segments in such a way that for any triangle obtained, all the points, except its vertices, lie outside the circumcircle of this triangle. Compared to other triangulation methods, Delaunay triangulation has several advantages:

1) the resulting triangles are close to equiangular, which makes it possible to reduce the numerical accuracy problems that can potentially arise in case of dividing the surface into long narrow triangles;
2) any point on the surface is located close to the reference point;
3) it does not depend on the point processing order.

Thus, triangulation allows to obtain the appropriate reference points for each unknown point. When the triangulated irregular network is constructed, we can accurately estimate the unknown values by applying any standard interpolation method within each triangle.

In this paper, the triangulated irregular network is constructed using Delaunay triangulation, and interpolation is performed by the following standard methods [11]:

1) linear;
2) cubic;
3) nearest neighbour;
4) natural neighbour.

3. Experimental results

The practical applicability of spatial interpolation methods, described in this paper, is demonstrated on the example of binary images.

The test images were converted to partially defined by removing 30, 50 and 70 percent of the pixels. The location of pixels to remove was selected using a pseudo-random number generator. The original and the resulting partially defined images are shown in figure 2. Here, the undefined pixels are indicated in gray.

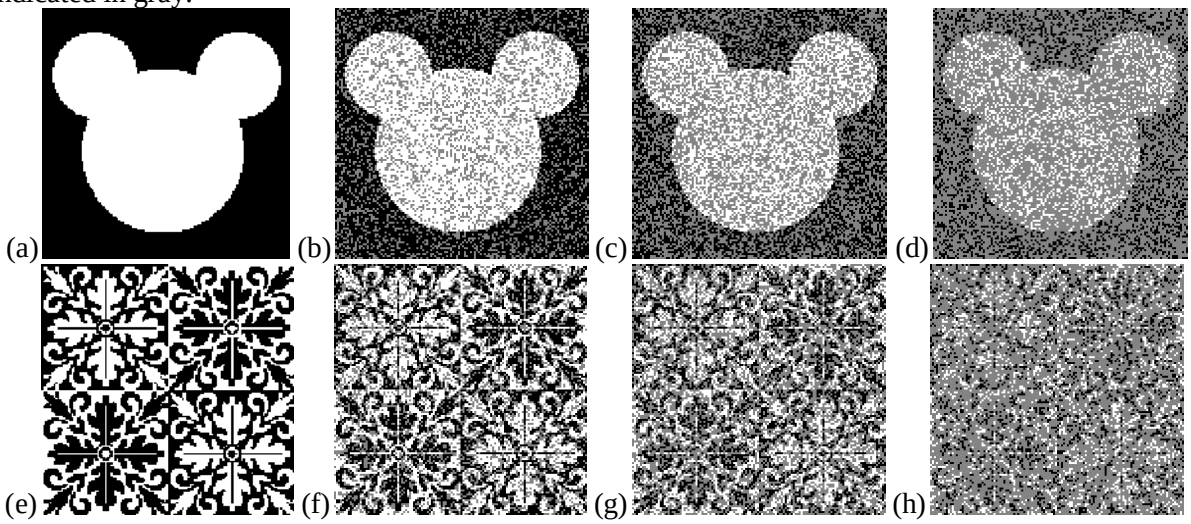


Figure 1. Test images: (a) original "Mickey" image; (b) "Mickey", 30% pixels removed; (c) "Mickey", 50% pixels removed; (d) "Mickey", 70% pixels removed; (e) original "Ornament" image; (f) "Ornament", 30% pixels removed; (g) "Ornament", 50% pixels removed; (h) "Ornament", 70% pixels removed.

To evaluate the quality of interpolation methods, we calculated the fraction of pixels matched for the original and interpolated images.

It is necessary to take into account that the image, obtained after interpolation, is halftone, while the original image is binary. To provide an accurate comparison, the resulting halftone images were converted into binary by thresholding. The comparison results for three interpolation methods are given in Table 1.

Table 1. Experimental results.

		Mickey			Ornament		
Method	Parametres	30%	50%	70%	30%	50%	70%
IDW	p=0, r=inf	0.8522	0.7756	0.6534	0.8506	0.7507	0.6489
	p=1, r=inf	0.9917	0.9854	0.9750	0.8698	0.7813	0.6819
	p=2, r=inf	0.9966	0.9954	0.9893	0.9366	0.8866	0.8120
	p=3, r=inf	0.9973	0.9957	0.9896	0.9476	0.9022	0.8313
	p=4, r=inf	0.9972	0.9952	0.9890	0.9503	0.9058	0.8340
	p=5, r=inf	0.9969	0.9950	0.9886	0.9512	0.9065	0.8345
	p=6, r=inf	0.9968	0.9951	0.9885	0.9513	0.9063	0.8344
	p=7, r=inf	0.9968	0.9949	0.9885	0.9505	0.9051	0.8334
	p=8, r=inf	0.9968	0.9947	0.9883	0.9502	0.9049	0.8337
	p=9, r=inf	0.9966	0.9944	0.9882	0.9497	0.9031	0.8328
	p=10, r=inf	0.9966	0.9943	0.9882	0.9494	0.9025	0.8327
	p=2, r=0	0.8522	0.7566	0.6534	0.8506	0.7507	0.6489
	p=2, r=1	0.9952	0.9825	0.9077	0.9485	0.8953	0.8034
	p=2, r=2	0.9974	0.9957	0.9849	0.9506	0.9053	0.8340
	p=2, r=3	0.9974	0.9957	0.9899	0.9449	0.9000	0.8279

	p=2, r=4	0.9973	0.9958	0.9900	0.9411	0.8936	0.8218
	p=2, r=5	0.9973	0.9958	0.9899	0.9394	0.8912	0.8199
	p=2, r=6	0.9972	0.9958	0.9901	0.9391	0.8910	0.8197
	p=2, r=7	0.9973	0.9958	0.9902	0.9392	0.8908	0.8196
	p=2, r=8	0.9973	0.9957	0.9902	0.9393	0.8908	0.8194
	p=2, r=9	0.9971	0.9958	0.9903	0.9391	0.8917	0.8190
	p=2, r=10	0.9971	0.9958	0.9901	0.9383	0.8909	0.8177
TIN	linear	0.9952	0.9934	0.9890	0.9402	0.8938	0.8253
	nearest	0.9955	0.9935	0.9871	0.9348	0.8888	0.8267
	natural	0.9973	0.9960	0.9912	0.9559	0.9124	0.8361
	cubic	0.9965	0.9944	0.9893	0.9496	0.9047	0.8351
Kriging	linear	0.8572	0.7522	0.6639	0.8506	0.7507	0.8033
	circular	0.8522	0.9825	0.9917	0.8506	0.8896	0.8037
	spherical	0.8522	0.9211	0.9079	0.9544	0.9092	0.8033
	exponential	0.9975	0.9958	0.9916	0.9576	0.9145	0.8385
	gaussian	0.8522	0.7566	0.6815	0.8506	0.8478	0.6489
	stable	0.8522	0.8318	0.9913	0.9598	0.9171	0.6489

From Table 1 it is evident that in the case of the "Mickey" image, where the object and background are clearly distinguishable, all methods show approximately the same result for certain parameters. For the "Ornament" image, the kriging method demonstrates the best result, while the IDW method proved to be the least accurate. However, in the case of 70% removed pixels, the quality of all interpolation methods decreases significantly for both images.

The best results for each interpolation method are represented as halftone images in Table 2.

Table 2. Interpolation results for best parameters.

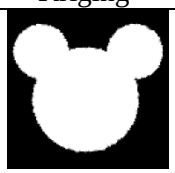
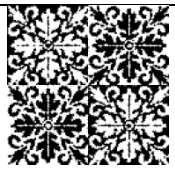
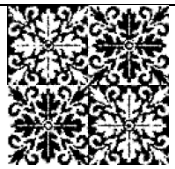
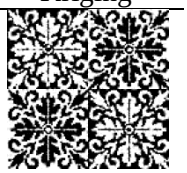
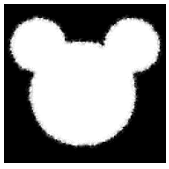
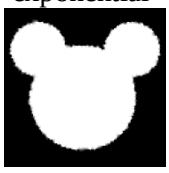
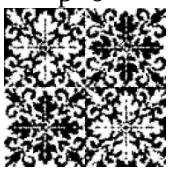
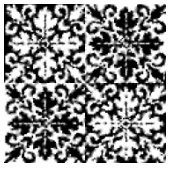
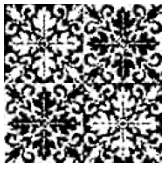
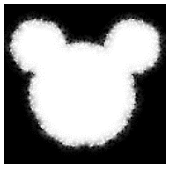
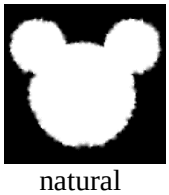
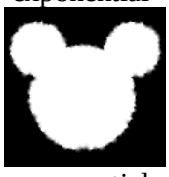
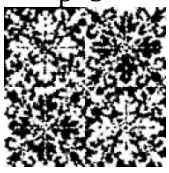
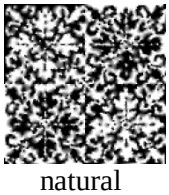
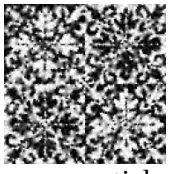
%	IDW	TIN	Kriging	IDW	TIN	Kriging
30						
	r=2	natural	exponential	p=6	natural	stable
50						
	r=4	natural	exponential	p=5	natural	stable
70						
	r=9	natural	exponential	p=5	natural	exponential

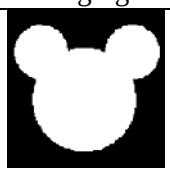
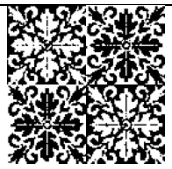
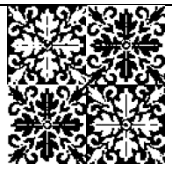
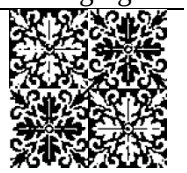
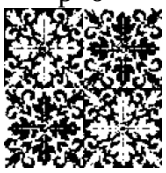
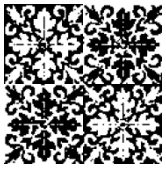
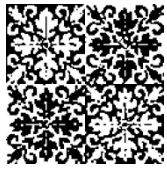
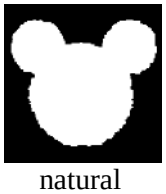
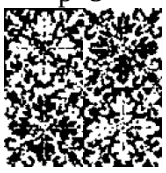
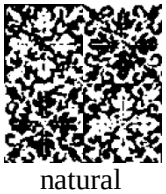
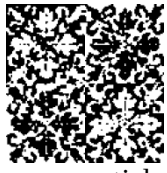
Table 3 also shows the best interpolation results obtained for each method, but the images are converted to binary using the thresholding procedure.

It can be seen from Tables 2 and 3 that in the case of 30% removed pixels, all the methods demonstrate an excellent result, and it is almost impossible to distinguish a small difference in quality. When 50% of the pixels are removed, the interpolation quality for the "Ornament" image slightly decreases. The 70% pixel removal significantly reduces the quality of all methods for both images.

On the whole, Tables 2 and 3 confirm the results of Table 1. Thus, according to the experimental results, it can be concluded that all the investigated methods can be used for interpolation of partially defined images, but it should noted that there is no single interpolation method suitable in all

situations: all methods have advantages and disadvantages. In practice, the choice of a particular interpolation method (and also the choice of parameters) should depend on the sample data and the error tolerance.

Table 3. Interpolation results after thresholding.

%	IDW	TIN	Kriging	IDW	TIN	Kriging
30						
	r=2	natural	exponential	p=6	natural	stable
50						
	r=4	natural	exponential	p=5	natural	stable
70						
	r=9	natural	exponential	p=5	natural	exponential

4. Conclusion

In this paper, three spatial interpolation methods are investigated: inverse distance weighting, interpolation based on a triangulated irregular network, and kriging. The practical applicability of these methods in the task of estimating the missing pixels of partially defined images is analyzed. The experiments show that all three methods can be used for solving the image processing problems, but the result of pixel prediction depends on the input data characteristics, such as the number of reference points. Hence, to provide the better results, the interpolation method should be selected separately for each input image.

5. References

[1] Gonzalez R C and Woods R E 2002 *Digital Image Processing* (New Jersey: Prentice Hall) p 943

[2] Fadnavis S 2014 Image interpolation techniques in digital image processing: an overview *Int. Journal of Engineering Research and Applications* **4(10)** 70-73

[3] Gashnikov M V, Glumov N I, Kuznetsov A V, Mitekin V A, Myasnikov V V and Sergeev V V 2016 Image processing, pattern recognition: Hyperspectral remote sensing data compression and protection *Computer Optics* **40(5)** 689-712 DOI: 10.18287/2412-6179-2016-40-5-689-712

[4] Watson D F and Philip G M 1985 A refinement of inverse distance weighted interpolation *Geo-Processing* **2** 315-327

[5] Jassim F A and Altaany F H 2013 Image interpolation using kriging technique for spatial data *Canadian Journal on Image Processing and Computer Vision* **4(2)** 91-96

[6] Panagiotopoulou A and Anastassopoulos V 2007 Super-resolution image reconstruction employing Kriging interpolation technique *Proc. IEEE Int. Workshop Syst., Signals and Image Processing* **1** 144-147

[7] Ruiz-Alzola J, Alberola-Lopez C and Westin C F 2005 Kriging filters for multidimensional signal processing *Signal Processing* **8** 413-439

[8] Panggabean M and Rønningen L 2011 Chroma interpolation using windowed kriging for color-image compression-by-network with guaranteed delay *Proc. IEEE EURASIP 17th International Conference on Digital Signal Processing (DSP)* **1** 1-6

[9] Mitas L and Mitasova H 1999 Spatial interpolation *Geographical Information Systems: Principles, Techniques, Management and Applications* **1** 481-492
[10] Lee D T and Schachter B J 1980 Two algorithms for constructing a Delaunay triangulation *International Journal of Computer and Information Sciences* **9(3)** 219-242
[11] Dumitru P D, Plopeanu M and Badea D 2013 Comparative study regarding the methods of interpolation *Recent Advances in Geodesy and Geomatics Engineering* **1** 45-52

The analysis of image characteristics on the base of energy features of the wavelet transform

S A Lyasheva[1], M V Medvedev[1], M P Shleymovich[1] and V V Mokshin[1]

[1]Kazan National Research Technical University -KAI, K. Marks 10, Kazan, Russia, 420111

Abstract. The article shows the relevance of creating models and methods that provide effective solutions to image processing and analyzing problems in computer vision systems. We consider models of an average level of image representation. They are constructed on the basis of their characteristic features (contours, regions and points of interest). To construct such models, we suggest using the procedure of forming energy features based on the wavelet transform. As a result, the original image will be transformed to a view where different points will have different weights. That characterize their contribution to the overall energy of the image. It is also possible to provide a selection of tuning coefficients. It allows to take into account the interrelations between the wavelet coefficients of various scales. After receiving the weight images, they can be processed to form the required characteristics. For example, to obtain contours, you can perform binarization of a weight image with a certain threshold. To get singular points, you can define a specified number of the most significant weights in different areas of the weight image. For texture analysis, you can use statistical characteristics calculated by the histogram of the scale.

1. Introduction

Computer vision software and hardware based on models and methods of image processing and analyzing are actively developing. For example, they are used in the navigation and control systems for pilotless aerial vehicles, systems for remote sensing of the Earth, access control systems for protected systems, industrial control systems, etc.

Generally, in computer vision systems, you need to provide registration, transformation, description, and image analysis. The effectiveness of these problems solutions is associated with the applicable models and methods optimized for the specific conditions of systems functioning.

The description of models and methods of image processing and analyzing is based on the paradigm of D. Marr. He proposed a three-level model of human perception of real-world objects. According to this paradigm, we single out low, medium and high levels of representation of images and their processing. [1].

At a low level, functional, probabilistic and hierarchical models are used. Functional models describe images in the form of some functions. For example, the description of the image as a function of spatial coordinates. In the case of a binary or halftone image, the values of the function are scalar. In the case of color or multispectral image, they have vector quantities. Probabilistic models describe images in the form of realizations of random processes and use probability density functions and statistical moments (mathematical expectation, variance, etc.). Hierarchical models represent images in the form of image sets of different scales. An example of a hierarchical model is the Gaussian image pyramid.

To represent images at the middle level there used the description of their characteristic features — contours, regions, points of interest. The construction of middle-level models is carried out in two stages. At the first stage, the segmentation of the image is performed. On the second stage - a description in the form of a set of attributes characterizing the selected segments.

At a high level of representation of images, models of explicit and implicit use of knowledge are applied. An example of a model of implicit use of knowledge is a model based on the use of template images. The knowledge on objects in that model is contained in their sample images. The model of explicit use of knowledge is a set of rules for interpreting images.

Note that low-level models are used to build middle-level models, and models of medium and low levels are used to build high-level models. In any case, it is necessary to determine what features of the images are used to describe them. Usually, the signs of color, texture, shape and structure are used. They allow you to describe images in terms of their color content, spatial distribution of colors or brightness, the characteristics of regions, the presence of certain objects and their relative location.

2. Models of image characteristic

As already mentioned, mid-level models of image representation are based on their characteristic features – contours, regions and points of interest.

Contour models are based on the selection and analysis of boundaries between areas in the image. There are many methods for delimiting boundaries. For example, the methods based on morphological operators and derivative operators of Roberts, Prewitt, Sobel, Laplace and others [2,3].

Region models describe the color or texture content of image areas. They are built on the basis of colour and texture features: a colour histogram [4], a colour connection vector [5], a correlogram of colours [6], colour moments [7], a dominant colour descriptor [8], statistical textural features [2], local binary patterns [9], spectral features [10], Tamura's features [1] and others.

Point models describe points of interest (point features, singular points, key points). These models are based on detectors and descriptors of singular points. The first are designed to search for special points on images, and the latter - to describe them. Currently, many detectors and descriptors are used in practice, e.g., the detector of Moravets, the Haris detector, SUSAN, SIFT, SURF, FAST, BRIEF, ORB, GLOH, FREAK, BRISK [12, 13].

Although there are many models of image characteristics, we can see new methods of their construction emerging. Some of the approaches to constructing effective models are based on applying the wavelet transform [14,15].

3. The wavelet transform of images

In general terms the wavelet transform is represented as follows:

$$Wf(\mathbf{u}, s) = \int_{-\infty}^{+\infty} f(\mathbf{x}) \frac{1}{s^{D/2}} \psi^* \left(\frac{\mathbf{x} - \mathbf{u}}{s} \right) d\mathbf{x}, \quad (1)$$

where Wf – the transform result; f – initial function; ψ^* – complex conjugation of the shifted and scaled function ψ that has zero mean value, the center at the zero point, and the unit norm; D – signal dimension; u – D-dimensional vector of the shift parameters; s – scale parameter [16].

For digital images, they often use discrete orthogonal multiple-scale wavelet transforms. The transforms are based on the representation of a discrete function f(x) describing the original signal as a sum of approximating fa(x) and detailing fd(x) components:

$$f(\mathbf{x}) = f_a(\mathbf{x}) + f_d(\mathbf{x}). \quad (2)$$

The process of transforming the function f(x) can also be represented in the iterative form, which determines it at different levels of the expansion:

$$f(\mathbf{x}) = f_a^J(\mathbf{x}), \quad (3)$$

$$f_a^j(\mathbf{x}) = f_a^{j-1}(\mathbf{x}) + f_d^{j-1}(\mathbf{x}), \quad (4)$$

where J – number of decomposition levels; $j = J, \ldots, j_0 + 1$. The result of the discrete wavelet transform has the form of a set of approximation $\{a_{j_0,l}\}$ and detail $\{d_{j,l}\}$ coefficients [17].

Wavelet transformations of single-channel images (e.g., halftone images) are usually performed in two stages: first converting for lines, then for columns (or vice versa). The results of the transformation at the *j*- level are grouped into a matrix of approximating coefficients $\left[LL_{j,m,n}\right]_{m,n=0}^{2^j-1}$ and matrixes of detailed horizontal $\left[LH_{j,m,n}\right]_{m,n=0}^{2^j-1}$, vertical $\left[HL_{j,m,n}\right]_{m,n=0}^{2^j-1}$, and diagonal $\left[HH_{j,m,n}\right]_{m,n=0}^{2^j-1}$ coefficients. For multi-channel images (for example, color images), each channel is transformed individually.

4. Energy features of images

The equation for an orthonormal wavelet transform is:

$$\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} f_{k,l}^2 = \sum_{m=0}^{2^{j_0}-1}\sum_{n=0}^{2^{j_0}-1} LL_{j_0,m,n}^2 + \\ + \sum_{j=j_0}^{J-1}\sum_{m=0}^{2^j-1}\sum_{n=0}^{2^j-1} LH_{j,m,n}^2 + \sum_{j=j_0}^{J-1}\sum_{m=0}^{2^j-1}\sum_{n=0}^{2^j-1} HL_{j,m,n}^2 + \sum_{j=j_0}^{J-1}\sum_{m=0}^{2^j-1}\sum_{n=0}^{2^j-1} HH_{j,m,n}^2, \quad (5)$$

where $f_{k,l}$ – brightness of the image point; $LL_{j,m,n}$, $LH_{j,m,n}$, $HL_{j,m,n}$, $HH_{j,m,n}$ – wavelet coefficients.

The sums on the right side of (5) show the contribution of the coefficients of different levels to the total energy of the image at different scales. This contribution makes it possible to get energy estimates for each point of the image. For example, using the orthonormal wavelet transform, when the number of rows and columns of the image is N and is divisible by two, the energy estimates can be obtained as follows:

1. Convert to level j_0;
2. Assign value:

$$w_{j_0-1,m,n}^2 = LL_{j_0,m,n}^2; \quad (6)$$

3. Calculate energy estimates:

$$w_{j,m,n}^2 = 0.25 w_{j-1,m,n}^2 + LH_{j,m/2,n/2}^2 + HL_{j,m/2,n/2}^2 + HH_{j,m/2,n/2}^2. \quad (7)$$

sequentially for $j = j_0, \ldots, J-1$, where $J = \log_2 N$, $m = 0, 1, \ldots, 2^{j+1}-1$, $n = 0, 1, \ldots, 2^{j+1}-1$

The calculated values preserve the energy equality:

$$\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} f_{k,l}^2 = \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} w_{k,l}^2, \quad (8)$$

where $w_{k,l}^2 = w_{J-1,k,l}^2$. Each of the set of values $\left\{w_{k,l}^2\right\}_{k,l=0}^{N-1}$ or $\left\{w_{k,l}\right\}_{k,l=0}^{N-1}$ can serve as the weight of the corresponding pixel, characterizing its contribution to the total energy of the image.

The energy equation (5) is valid only for orthonormal wavelet transforms. For a more general case, the described procedure for calculating energy estimates can be modified as follows:

1. Convert to level j_0;
2. Assign value:

$$w_{j_0-1,m,n}^2 = K'_{j_0-1} LL_{j_0,m,n}^2; \quad (9)$$

3. Calculate energy estimates:

$$w_{j,m,n}^2 = K'_j w_{j-1,m,n}^2 + K''_j \times \left[LH_{j,m/2,n/2}^2 + HL_{j,m/2,n/2}^2 + HH_{j,m/2,n/2}^2\right]. \quad (10)$$

sequentially for $j = j_0, \ldots, J-1$, where $J = \log_2 N$, $m = 0, 1, \ldots, 2^{j+1}-1$, $n = 0, 1, \ldots, 2^{j+1}-1$

The introduction of tuning coefficients K'_j and K''_j makes it possible to provide an optimal calculation of a set of values $\left\{w_{k,l}\right\}_{k,l=0}^{N-1}$ in accordance with the problem being solved. In this case,

these values can be considered as weights of points. Thus, you can build an image in which each point will be associated with its weight - weight image.

5. The application of the energy features model for image analysis

Basing on the energy features model, we can construct a description of the images for detecting and analyzing their characteristic features.

As a result of the described procedure, the original image is transformed so that different points have different weights. As already mentioned, the weights characterize the points' contribution to the overall energy of the image. In this case, you can select the tuning factors so that the weight points at the boundary points will be larger than for the inners. It is possible, because while transferring from one area of the image to another the boundary points are in the places of brightness difference. To estimate the magnitude of the difference, we use the expression:

$$\Delta f = [(LH^2 + HL^2 + HH^2)/3]^{1/2}, \quad (11)$$

where Δf – amount of difference in the point of the image; LH, HL, HH – detailing wavelet coefficients at the point of the image. Thus, the energy of the difference in the point is proportional to the magnitude $LH^2 + HL^2 + HH^2$. Besides, the wavelet transform allows us to evaluate the significance of points at various scales. Since the image points of different scales are interrelated, we can get integral characteristics that take into account the significance of points on all considered scales. This reasoning can be used as a basis for the procedure of detecting and analyzing contours in an image. If we consider the distribution of the image point's weights, we can get a description of the texture that characterizes its different regions. Except contours and regions, the analysis of singular points is often used. In this case, by special points we mean the points with the largest weights in the neighborhoods of the given dimensions.

To illustrate the approach to detecting and analyzing the characteristic features of the image, we took a photo of the Kazan Kremlin territory in Figure 1 and its weight models in Figure 2 and Figure 3, obtained using different sets of tuning coefficients. The values of the tuning coefficients for the weight models in Figure 2 are shown in Table 1, and the values of the tuning coefficients for the weight models in Figure 3 are given in Table 2. As the source image and weight images have dimensions of 256 × 256 pixels, the tuning coefficients in Table 1 and Table 2 are shown for eight levels of decomposition from $j_0 = 0$ to $J - 1 = 7$. Here the value of the coefficient K'_{-1} is assumed to be 1. Table 1 demonstrates that moving from level to level, the influence of the previous levels is significantly reduced, and the influence of the detailing coefficients of this level increases. The tuning coefficients given in Table 2 make it possible to reduce the influence of the previous levels weights on the current level. At the same time, they let increase significantly the influence of the detailed wavelet coefficients of this level for increasing the importance of the boundary points.

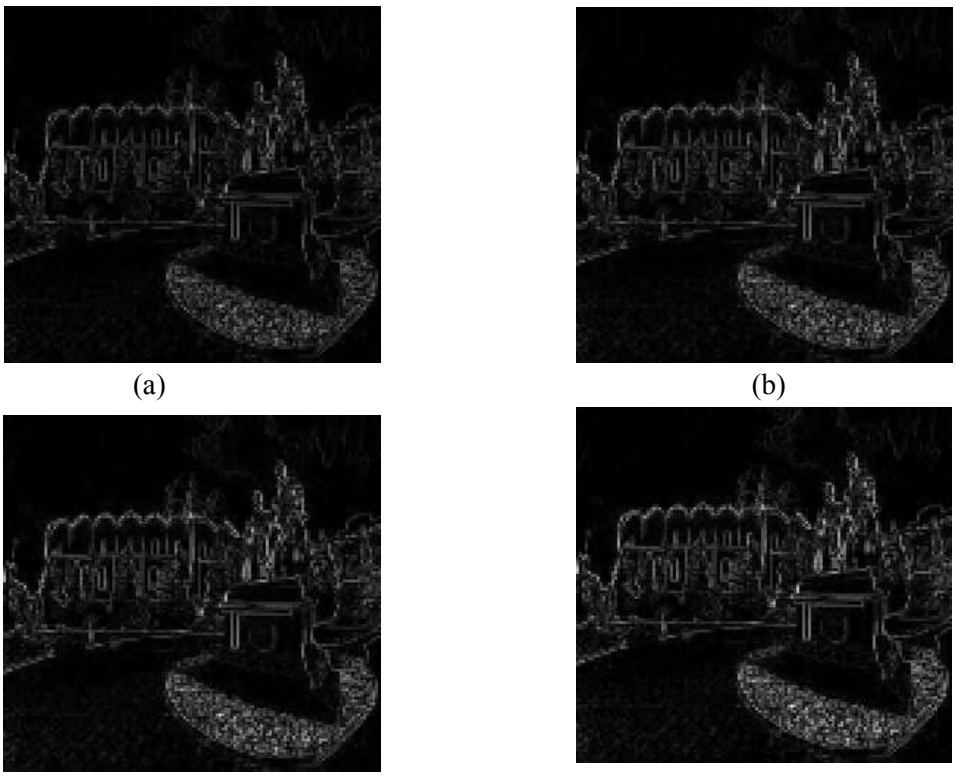
Table 1. The tuning coefficients for the weight models in the figure 2.

	Level 0		Level 1		Level 2		Level 3		Level 4		Level 5		Level 6		Level 7	
	K'_0	K''_0	K'_1	K''_1	K'_2	K''_2	K'_3	K''_3	K'_4	K''_4	K'_5	K''_5	K'_6	K''_6	K'_7	K''_7
Figure 2 (a)	1	1	0.67	1	0.44	1	0.29	1	0.19	1	0.13	1	0.09	1	0.06	1
Figure 2 (b)	1	2	0.67	2	0.44	2	0.29	2	0.19	2	0.13	2	0.09	2	0.06	2
Figure 2 (c)	1	3	0.67	3	0.44	3	0.29	3	0.19	3	0.13	3	0.09	3	0.06	3
Figure 2 (d)	1	4	0.67	4	0.44	4	0.29	4	0.19	4	0.13	4	0.09	4	0.06	4

Table 2. The tuning coefficients for the weight models in the figure 3.

	Level 0		Level 1		Level 2		Level 3		Level 4		Level 5		Level 6		Level 7	
	K'_0	K''_0	K'_1	K''_1	K'_2	K''_2	K'_3	K''_3	K'_4	K''_4	K'_5	K''_5	K'_6	K''_6	K'_7	K''_7
Figure 3 (a)	1	1	0.91	2.00	0.83	4.00	0.75	8.00	0.68	16.0	0.62	32.0	0.56	64.0	0,51	128
Figure 3 (b)	1	1	0.91	1.67	0.83	2.78	0.75	4.63	0.68	7.72	0.62	12.9	0.56	21.4	0.51	35.7
Figure 3 (c)	1	1	0.91	1.43	0.83	2.04	0.75	2.92	0.68	4.16	0.62	5.95	0.56	8.49	0.51	12.1
Figure 3 (d)	1	1	0.91	1.25	0.83	1.56	0.75	1.95	0.68	2.44	0.62	3.05	0.56	3.81	0.51	4.77

Figure 1. Original image.

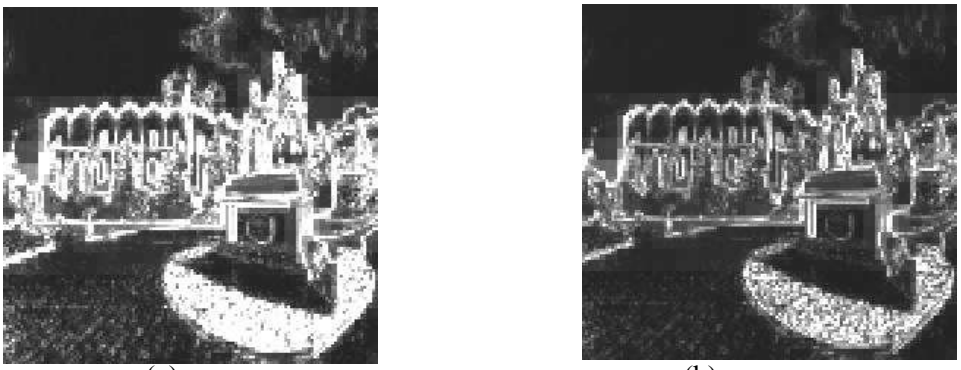


(a) (b)

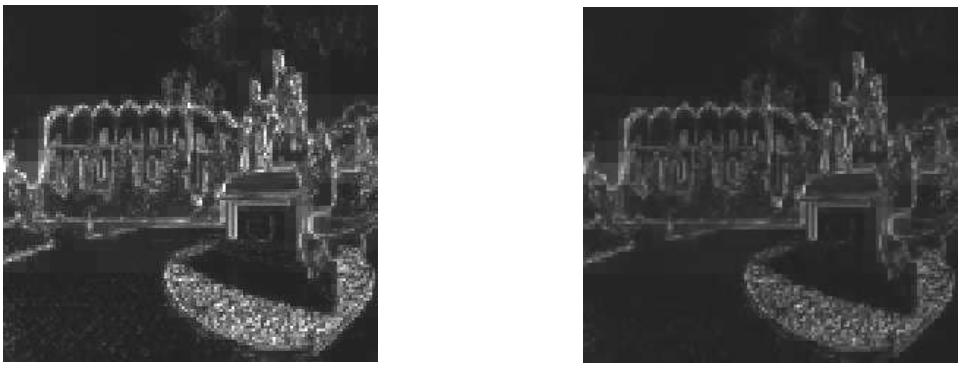
(c) (d)

Figure 2. Weight images without approximating wavelet coefficients.

After receiving the weight images, we can process them to form characteristics describing contours, texture or singular points in the original images. For example, to get contours, you can binarize a weight image with a specified or calculated threshold value. Figure 4 shows the results of binarization of images in Figure 2 using the same threshold. To get singular points, you can determine the specified number of the most significant weights in different areas of the weight image.

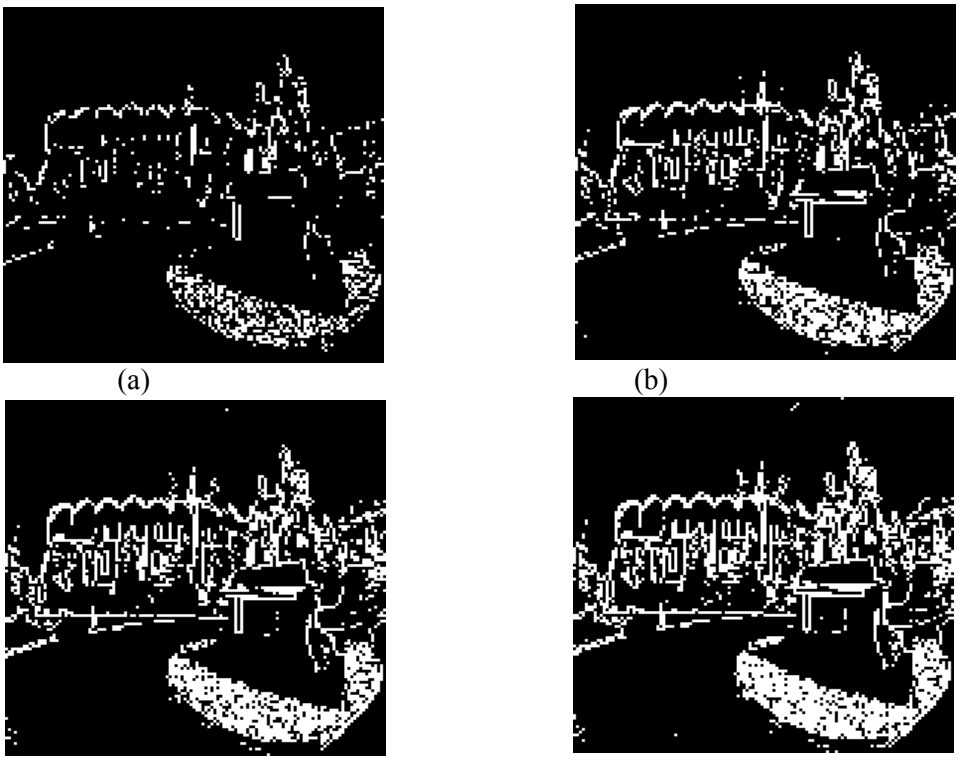


(a) (b)



(c) (d)

Figure 3. Weight images with approximating wavelet coefficients.



(a) (b)

(c) (d)

Figure 4. Binary images.

In our case, the singular points will also coincide with certain points of the binary image. For texture analysis, you can use such statistical characteristics as mathematical expectation, variance, and asymmetry coefficient, coefficient of excess, smoothness, homogeneity and entropy. They can be calculated from the weight values histogram. Note that compared to the formation in the processing of original images, a higher noise resistance is provided in the formation of texture characteristics based on the energy characteristics model. [18, 19].

6. Conclusion

The described approach to detecting and analysing the characteristic features of images can serve as a basis for constructing systems of objects detection and recognition in various systems based on methods and means of computer vision, including onboard systems of pilotless aerial vehicles – to detect and recognise the objects on images, process control systems, intelligent transport systems, etc. [20–22].

7. References

[1] Potapov A S 2007 *Pattern Recognition and Machine Perception: A General Approach Based on the Principle of Minimum Description Length* (St. Petersburg: Politekhnika) p 548
[2] Gonzalez R C and Woods R E 2007 *Digital Image Processing* (Prentice Hall) p 976
[3] Belim S V and Kutlunin P E 2015 Boundary extraction in images using a clustering algorithm *Computer Optics* **39(1)** 119-124 DOI: 10.18287/0134-2452-2015-39-1-119-124
[4] Long F, Zhang H and Feng D 2003 Fundamentals of content-based image retrieval *Multimedia Information Retrieval and Management – Technological Fundamentals and Applications* 1-26
[5] Pass G and Zabih R 1996 Histogram refinement for content-based image retrieval *Proceedings 3rd IEEE Workshop on Applications of Computer Vision* 96-102
[6] Huang J, Kumar S R, Mitra M, Zhu W-J and Zabih R 1999 Spatial Color Indexing and Applications *International Journal of Computer Vision* **35(3)** 245-268
[7] Stricker M and Orengo M 1995 Similarity of Color Images *Proceedings of the SPIE Conference* **2420** 381-392
[8] Deng Y, Manjunath B S, Kenney Ch, Moore M S and Shin H 2001 An efficient color representation for image retrieval *IEEE Transactions on image processing* **10(1)** 140-147
[9] Pietikinen M, Hadid A, Zhao G and Ahonen T 2011 *Computer Vision Using Local Binary Patterns* (Springer) p 224
[10] Tuceryan M and Jain A K 1998 Texture analysis *The Handbook of Pattern Recognition and Computer Vision* 207-248
[11] Tamura H, Mori S and Yamawaki T 1978 Texture features corresponding to visual perception *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-8(6)** 460-473
[12] Awad A I and Hassaballah M 2016 *Image Feature Detectors and Descriptors* (Springer) p 438
[13] Fan B, Wang Z and Wu F 2015 *Local Image Descriptor: Modern Approaches* (Springer) p 108
[14] Tang Y Y 2009 *Wavelet Theory and Its Application to Pattern Recognition* (World Scientific Publishing Company) p 482
[15] Spitsyn V G, Bolotova Y A, Phan N H and Bui T T 2016 *Using a Haar wavelet transform, principal component analysis and neural networks for ocr in the presence of impulse noise Computer Optics* **40(2)** 249-257 DOI: 10.18287/2412-6179-2016-40-2-249-257
[16] Addison P S 2002 *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance* (Institute of Physics Publishing) p 368
[17] Mallat S 2009 *A Wavelet Tour of Signal Processing: The sparse way* (Academic Press) p 832
[18] Shleymovich M P and Lyasheva S A 2016 The energy characteristics of images based on the wavelet transform *The II International Conference on Information Technology and Nanotechnology ITNT-2016* 334-341
[19] Lyasheva S A and Shleymovich M P 2016 Image analysis based on the wavelet transform energy features *New Technologies, Materials And Equipment of the Russian Aerospace Industry* 147-153
[20] Lyasheva S A, Medvedev M V and Shleimovich M P 2014 Terrain object recognition in unmanned aerial vehicle control system *Russian Aeronautics* **57(3)** 303-306
[21] Shleymovich M P, Medvedev M V and Lyasheva S A 2017 Image analysis in unmanned aerial vehicle on-board system for objects detection and recognition with the help of energy characteristics based on wavelet transform *Proceedings of SPIE* **10342**
[22] Shleymovich M P, Medvedev M V and Lyasheva S A 2016 Object detection in the images in industrial process control systems based on salient points of wavelet transform analysis *Proceedings 2nd International Conference on Industrial Engineering, Applications and Manufacturing* 7911633

Acknowledgments
The research was carried out within the framework of the state task No 2.1724.2017/4.6.

Hippocampus detection and calculation of its characteristics in magnetic resonance imaging of the brain

V N Gridin[1], M I Truphanov[1] and V I Solodovnikov[1]

[1]Center of Information Technologies in Engineering RAS, Marshal Biryuzov Str. 7a, Odintsovo, Moscow region, Russia, 143000

Abstract. Hippocampus is the most informative object of the brain for the purpose of detection of Alzheimer's disease signs. This paper presents the approach which was developed for detection of hippocampus and subsequent calculation of its parameters while analyzing the series of images obtained in a sagittal projection by means of a magnetic resonance tomography. This paper introduces the algorithms for detecting the keyframes in the entire series which contain the hippocampus and for identifying the hippocampus among other brain structures. The problems of measuring the volumatric parameters of the hippocampus and calculating its characteristics are considered. These characteristics serve as the basis for instrumental calculation of the signs which characterize the possible presence of Alzheimer's disease.

1. Introduction

The increased life expectancy of the world's population results in the increase of the occurrence and prevalence of Alzheimer's disease. According to the World Health Organization, the number of people suffering from Alzheimer's disease will double every twenty years and will reach 115.4 million by 2050. Regarding this issue, it is critical to solve the task of identifying the disease at the earliest stage before the cognitive impairments affect the daily activity of a person, while there is still the potential possibility to slow down the progression of thedisease.

Unfortunately, there is no harmless, inexpensive, and most importantly, non-invasive methodto diagnose Alzheimer's disease with a high degree of certainty. Currently,confirmation of the final diagnosis includes a histopathological analysis of brain tissue or the cerebrospinal fluid study in order to determine the formation of beta-amyloid, the characteristic of Alzheimer's disease, for which the patient is required to undergo a spinal cord puncture. Also, there is a method of positron emission tomography of amyloid based on the use of radioactive material, which is a technologically complex and expensive process and implies the effect of radioactive radiation on the examined person. In this case, particular importance is acquired by the tasks related to development of the automated methods, algorithms, software and hardware to detect abnormalities in the structure and activity of the brain which are based on the visual analysis of optical cross-sections of the head in various planes obtained with the help of a magnetic resonance tomography (MRI).

The extensive range of works is devoted to the problems of the construction of the automated system for MRI images processing aimed to analyze the morphological features, which are inherent to Alzheimer's disease. Most of these works note that, first of all, the changes affect the temporal region of the brain and especially hippocampus. Thus, the primary and essential tasks are the detection of

hippocampus and calculation of its volumetric parameters for the purpose of the automated diagnosis of Alzheimer's disease [1, 2].

Currently, various software products (both commercial and freeware) are known that are used to automate the measurement process and calculate the characteristics of brain structures, in particular, the hippocampus and adjacent areas [3,4]. Their common drawback is the mandatory use of some manual operations which complicates the analysis process and leads to the errors in measurement of the key characteristics. The consequence is the potential errors in the diagnosis.

Thus, in [5] area measurements of the hippocampal region of rats brain were performed by means of analysis of the images with the help of a specialized software tool, focusing on the "Rat Brain" atlas, Paxinos G. and Watson C. The authors have used this atlas to find pre-known positions of the hippocampus in the general structure the brain. Obviously, an approach based on the use of atlas can be applied only for reading the estimated coordinates of the hippocampal region. It requires considerable refinement for the purpose of detection and measurement of the linear and volumetric parameters of the hippocampus.

In [6], the following digital image processing operations were applied providing the solution of a part of the problem in automatic mode: subtraction of the background to minimize the influence of the background component; the transformation of the image into monochrome; increasing of contrast; binarization; noise suppression. However, it is indicated that these operations were performed in manual mode using specialized software. The results of using this approach present a certain interest in terms of determining the certain parameters of individual steps during automated visual data processing.

The key features of the hippocampus structure are specified in [7, 8]. The authors in [8] carried out the research of the hippocampus images using specialized programs in an automated mode in the "FreeSurfer" software environment, as well as in semi-automatic and manual modes in the "Display" package. This software allows a user to perform operations on allocation of the brain structures and their volumetric evaluation. Standard operations of digital image processing (contour highlighting, binarization, etc.) were used to obtain the results. However, human participation was still necessary for the detection of the hippocampus. These approaches are not applicable for the automatic evaluation of the volumetric and dimensional parameters of hippocampus and adjacent brain structures.

In a number of works, the analysis of statistical characteristics and texture analysis as well as artificial neural networks are used to highlight brain structures (including hippocampus) [9, 10, 11].

In particular, in [12] the authors showed the possibility of using texture image characteristics together with an artificial neural network for analysis of the hippocampus region. However, a large training sample is needed to use this approach, which complicates the practical application of this approach. Another difficulty includes the automatic normalization which is necessity for orientation and spatial characteristics of the three-dimensional regions formed during the analysis of textures, which also reduces the practical value of this approach.

The issue of tomographic images analysis automation for other internal organs also becomes more and more relevant. So in the article [13] the simplest technology of automatic recognition of emphysema of lungs by sets of two-dimensional diagnostic images of computed tomography is considered. In [14], a method for segmentation of organs of the retroperitoneal space on tomographic images based on the level function was proposed.

In this connection, it can be concluded that further research is needed to improve the methodology and mathematical apparatus for automation of magnetic resonance imaging analysis and the calculation of hippocampus characteristics as a key informative brain structure, which would be an essential step in the Alzheimer's disease diagnosis.

2. The key stages in processing the series of MRI images

The developed approach for the visual MRI data processing aims to create tools for constructing automatic subsystems of analysis of brain structures images, in particular, the hippocampus for the Alzheimer's disease diagnosis. The following main stages could be picked out:

- selection of series slices in the sagittal projection, presumably containing the hippocampus;

- subsequent iterative processingin order to confirm /not confirm the detection of the hippocampus in each frame;
- clarifying the location of the hippocampus and calculating its characteristics by analyzing a series of neighboring frames or by repeating the selection of frames sequence presumably containing the hippocampus;
- transmission of the received parameter vector to the decision support module to get information about the possible presence of the Alzheimer's disease, the healthy state or the presence of brain changes not caused by Alzheimer's disease.

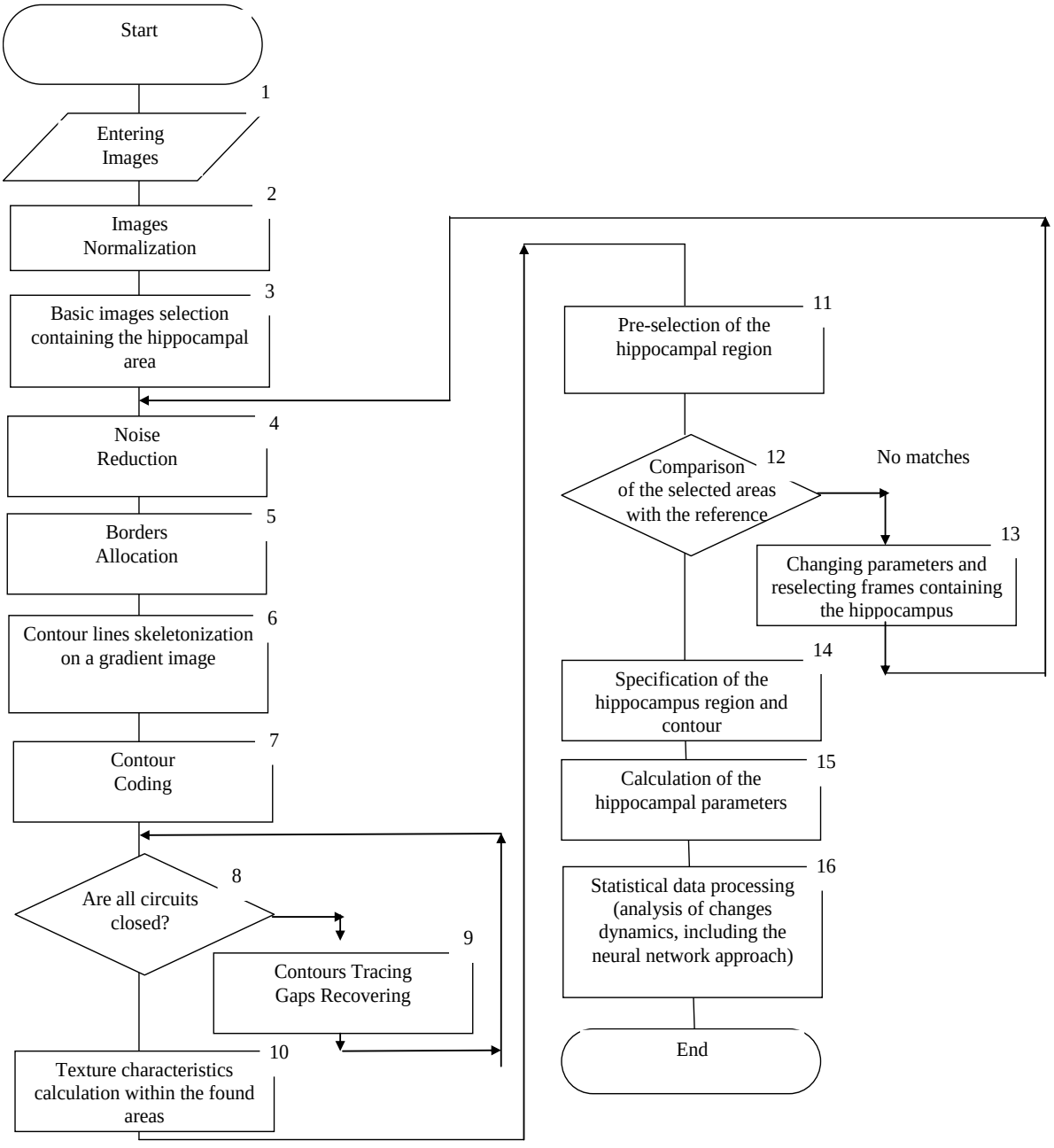


Figure 1. A generalized algorithm for hippocampus detecting and measuring of its parameters for the Alzheimer's disease signs detection.

Let’s consider the first group of operations aimed at the slice search and selection in the MRI sequence containing the hippocampus. Note that here and further we consider the location of the hippocampus in the left and right hemispheres of the brain. The main idea is to sequentially scan

images in the saggital projection and search for a structure that is supposedly appropriate to the hippocampus shape and location. The location is clarified in relation to the location of the eye, which is very informative since it is characterized by a unique close to round shape and as having a closed contour. Also, the calculations of the brain eccentricity and the neck vessels eccentricity are used to determine the orientation in space and the boundaries of the brain as the closed objects which have the largest area. The brain boundaries are calculated by detecting several curves of identical shape and length, located close to each other. These curves are the boundaries of the cranium and intracranial fluid.

The generalized algorithm of the proposed approach is shown in Figure 1.

3. The key stages for the series of MRI images processing

The hippocampus could be detected in a series of consecutive frames, on which there is an elongated body framed by a liquid, characterized by a closed contour of a pre-known shape and with pre-computed textural characteristics.

Figure 2 presents images, which illustrate the process of primary detection of the hippocampus in a series of sagittal MRI images.

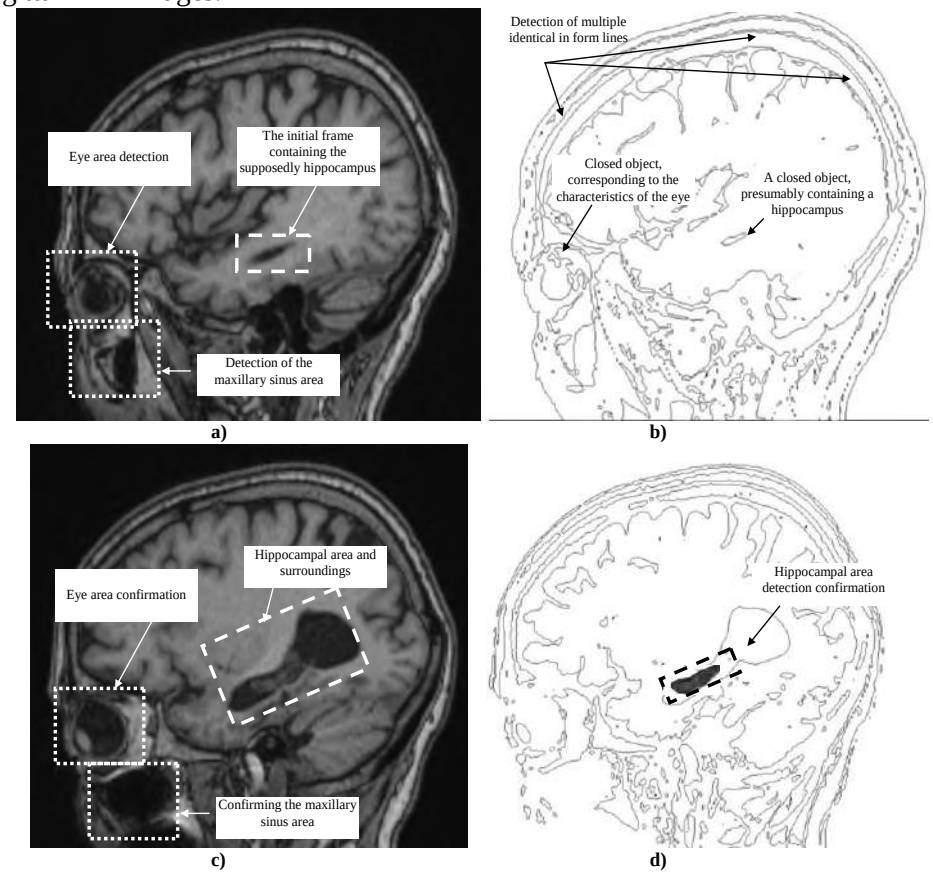


Figure 2. The process of search of the hippocampus: a, b) - the initial frames presumably containing the hippocampus; c, d) the confirmed frames containing the hippocampus in the middle of the sequence.

Thus, the search in the general sequence of sagittal images is directly related to the images detection on which the eyes and maxillary sinuses could be detected, as the most informative and precisely localized objects. Then, the skull boundary and the brain boundary are determined on the found sections to establish the relative position of the hippocampus and to calculate its potential location. Next, by analyzing a series of successive frames, it is iteratively determined those frames on which there is a region corresponding to a generalized description of the brightness and spatial characteristics of the hippocampus. The confirmation of the detection is made at the final stage by

constructing a closed contour for boundaries description. In addition, the comparison with the reference values and parameters of the adjacent hippocampus regions takes place.

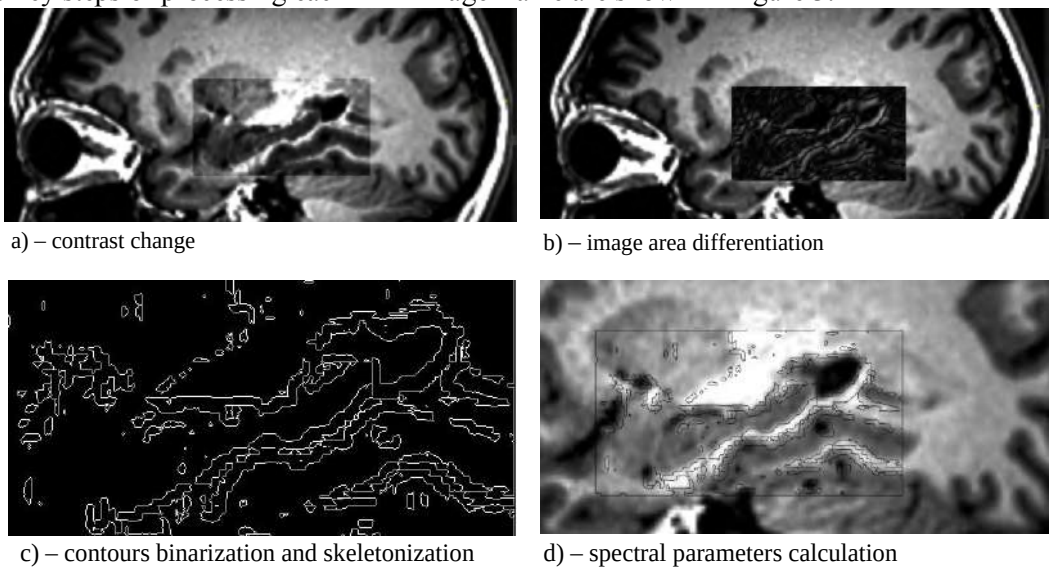
If the hippocampus detection was not confirmed, the initial frame of the sequence is selected again and the search process has to be repeated with the changed parameter values.

4. Determination of the hippocampus boundaries and position

Let's describe the process of precise determination of the hippocampus boundaries and position in more details. This process consists of the following main steps:

- pre-processing - reducing the random noise level by a Gaussian filter and adjusting the contrast to the reference parameters [15];
- locating the objects boundaries in the image by means of a differential operator;
- binarizationof the obtained boundaries by the threshold operator with adaptive threshold calculation based on the Otsu method;
- skeletonizationof the obtained contours and restoration of discontinuities in contours with the help of the initial halftone image analysis;
- additional confirmation of the detected hippocampus on the basis of generalized spectral characteristics of local image areas;
- inspecting the series of frames to determine the size of the hippocampal regions relative to the total brain size, and calculating the dimensions of the fluid, which adjacent to the hippocampus.

The key steps of processing each MRI image frame are shown in Figure 3.



a) – contrast change

b) – image area differentiation

c) – contours binarization and skeletonization

d) – spectral parameters calculation

Figure 3. The process of the hippocampus detection confirmation and its characteristics determination.

5. Noise elimination

Noise elimination consists of smoothing and eliminating of sudden jumps of brightness associated with the physical process of magnetic resonance imaging and the intrinsic noise of the tract of tomograph.

The filtering process is based on the calculation of the new luminance value g with a help of the Gaussian filter for each discrete image point according to the formula:

$$g(x,y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{d^2}{2\sigma^2}},$$

where σ is the filter parameter, $d = \sqrt{(x - x_c)^2 + (y - y_c)^2}$ is the distance of the [x,y] pixel from the center pixel of the neighborhood $[x_c, y_c]$, which determines the radius of the filter.

6. Selection of contours

The gradient vector of the image brightness function is calculated to extract the contours at each point. This gradient vector is described by the module $g(x, y)$ and the direction. The set of gradient vectors at each point of the filtered image f '(x,y) is a gradient image [16]

$$Gr = \langle g(x, y), v(x, y) \rangle, x = \overline{1, K}, y = \overline{1, N},$$

containing information about the differences in the brightness function, i.e. about contour lines (K, N are the image dimensions).

Calculation of the hippocampus and adjacent areas parameters is performed on a series of adjacent frames and the contour image shown in Figure 4.

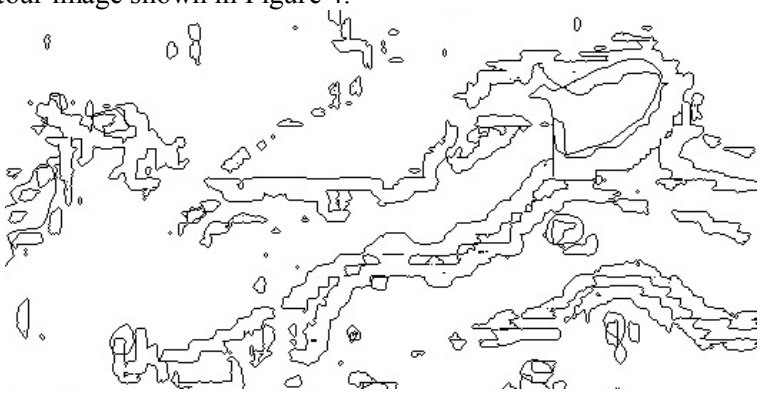


Figure 4. Contour image which is used to calculate hippocampal parameters.

It is worth noting that the obtained absolute values characterizing the state of the hippocampus, in the future, have to be transformed into relative values taking into account the total volume of the brain and the patient's gray matter. Also, all changes in the form of the hippocampus itself and of the surrounding areas could play an important role in the diagnosis. These changes, in turn, could be revealed by the resulting contour images.

7. Conclusion

The analysis of scientific, patent sources and reference information revealed that there is a lack of a developed algorithmic and software-hardware solution for automatic analysis of MRI data in order to detect abnormalities in the structure and activity of the brain inherent to Alzheimer's disease at an early stage. The task of the automatically hippocampus detection on MRI images and selection of the most informative images in a general sequence is the key task while constructing both automatic and automated means of calculating signs, which characteristic the Alzheimer's disease. This paper suggests an approach for automatic detection and measurement of the spatial characteristics of the hippocampus, as well as the allocation of nearby areas. This information is used as a characteristic space for making a decision about the possible presence of the disease. The practical significance of the being developed approach lies in the subsequent construction of a specialized domestic software product that allows automatic and automated analysis of the brain MRI images in the interests of timely detection of Alzheimer's disease and instrumental evaluation of its progress dynamics.

8. References

[1] Yahno N N, Zaharov V V and Lokshina A B 2011 *Dementia: hand-in for doctors* (Moscow: MED press-inform) p 272

[2] Lobzin V Yu, Kiselyov V N, Fokin V A, Emelin A Yu, Vorobyov S V, Lupanov I A, Sokolov A V and Efimcev A Yu 2013 Magnetic resonance morphometry usage in the diagnosis of Alzheimer's disease and vascular cognitive disorders *Bulletin of the Russian Military Medical Academy* **3(43)** 1-7

[3] Magonov E P, Kataeva G V and Trofimova T N 2015 Modern methods of automatic calculation of intracranial space volume with MRI-morphometry of the brain *Journal of the Bulletin of*

Novgorod State University. Yaroslav the Wise **2(85)** 98-104
[4] Chow N, Hwang K, Hurtz S, Green A, Somme J, Thompson P, Elashoff D, Jack C, Weiner M and Apostolova L 2015 Alzheimer's Disease Neuroimaging Initiative. Comparing 3T and 1.5T MRI for mapping hippocampal atrophy in the Alzheimer's Disease Neuroimaging Initiative *AJNR Am J Neuroradiology* **36(4)** 653-660
[5] Sidorova M V and Vakolyuk I A 2017 Determination of the hippocampus structural changes in the early period of modeling of post-traumatic stress disorder in rats *International Journal of Applied and Fundamental Research* **7(1)** 106-110
[6] Tishkina A O 2009 The automatic quantitative analysis method of brain sections microphotographs *Neurochemistry* **26(4)** 341-346
[7] McLean J 2012 *The investigation of hippocampal and hippocampal subfield volumetry, morphology and metabolites using 3T MRI* (Thesis for the degree of Ph.D. University Glasgow) p 354
[8] Ananeva N I, Ezhova R V, Galsman I E, Davlethanova M A, Rostovceva T M, Stulov I K, Vasserman L I, Shmeleva L M, Chujkova A V, Sorokina A V and Ivanov M V 2015 Hippocampus: radial anatomy, structural variants *Radiation diagnostics and therapy* **1(6)** 39-44
[9] Patil M M and Yardi A R 2011 Diagnosis of Alzheimer's Disease from 3D MR Images with Statistical Features of Hippocampus *CIIT, CCIS* **250** 744-749
[10] Kodama N, Kawase Y and Okamoto K 2016 Application of Texture Analysis to Differentiation of Dementia with Lewy Bodies from Alzheimer's Disease on Magnetic Resonance Images *IFMBE Proceedings* **14(3)** 1444-1446
[11] Rabeh A, Benzarti F and Amiri H 2016 Diagnosis of Alzheimer Disease from MRI Images of the Brain Throughout Time *Proceedings of the 16th International Conference on Hybrid Intelligent Systems, Advances in Intelligent Systems and Computing* **552** 219-226 DOI: 10.1007/978-3-319-52941-7_22
[12] Jing Z, Yu C, Jiang G, Liu W and Tong L 2012 3D texture analysis on MRI images of Alzheimer's disease *Brain Imaging and Behavior* 61-69
[13] Smelkina N A, Kolsanov A V, Chaplygin S S, Zelter P M and Khramov A G 2017 Pulmonary Emphysema Recognition by CT Scan *Computer Optics* **41(5)** 726-731 DOI: 10.18287/2412-6179-2017-41-5-726-731
[14] Eruslanov R V, Orehova M N and Dubrovin V N 2015 Retroperitoneal Space Organ Segmentation From CT Images Based on the Level Set Function *Computer Optics* **39(4)** 592-599 DOI: 10.18287/0134-2452-2015-39-4-592-599
[15] Gridin V N, Titov V S and Trufanov M I 2009 *Adaptive vision systems* (SPb.: Science) p 441
[16] Furman Ya A 2002 *Introduction to contour analysis and its applications to signal and image processing* (Moscow: Fizmatlit) p 592

Acknowledgments
The study was carried out at the expense of a grant from the Russian Science Foundation (project 17-11-01288).

Comparison of a low-frequency Butterworth filter with a symmetric SE-filter

K S Medvedeva[1]

[1]Saratov State University, Astrakhanskaya Street 83, Saratov, Russia, 410012

Abstract. The article compares two filters: a Butterworth filter and an asymmetric SE-filter.The experimental studydetermines their advantages and disadvantages.Also,experiments based on the peak signal-to-noise ratio(PSNR) metricshowvisual evaluation. The results of experimentsshow that a symmetric filter better restores images usinga small set of continuous function parametersthatare distorted by a low-frequency Gaussian filter.

1. Introduction

Due to the imperfection of forming and recording systems, images recorded by systemsare distorted (fuzzy) copiesof the original images. The main causes of distortions that resultindegradation of clarity includethe limited resolution of the forming system, refocusing, the presence of a distorting medium (for example, the atmosphere), and movement of the camera on the object being registered. Eliminating or reducing distortion for clarity is the task of image recovery.

Automatic control systems, measuring equipment, signal processing systems, and various filters with different characteristics are used to filter signalsin telecommunications. Depending on the frequency band associated with the bandwidth and the suppression band, there are odd, band, high-frequency, and low-frequency filters. Also, all-pass filters have a constant amplitude-frequency response in the required frequency range, and their phase-frequency response is a given frequency function [1].

The simplest way to restoreimage clarity is to process the observed image in the spatial frequency domain with an inverse filter [2].The drawbacks of this filter are the occurrence of edge effects, which take the form of an oscillating hindrance of high power that completely masksthe reconstructed image. Boundary effects can be observed even in noise-free images. But the main problem ofusing an inverse filter is that the corresponding transfer function can have poles close to zero.

This paper considers two frequency filters: a Butterworth filter and a symmetric SE-filter [3]. By means of experiments, comparative analysis will be carried out to identify their advantages and disadvantages. In paper [3], the technology of FIR filter synthesis is proposed, in which identification of the parameters of a continuous function approximating the impulse response is carried out instead of determining the impulse response samples. In this case, the parametric family of approximating functions, depending on a small number of parameters, is given, taking into account the desired frequency characteristics of FIR filters intended for distortion correction.

The Butterworth filter is chosen for several reasons. The amplitude response of this filter is smooth at the frequencies of the passband and almost decreased to zero at the frequencies of the suppression band. The Butterworth filter is the only filter that retains the amplitude-frequency response for higher

orders (except for the steep drop in the characterof the suppression band), whereasmany other types of filters (Chebyshev filter, Bessel filter, elliptical filter) have different forms of frequency response in different orders. In [3], a symmetric filter with a Wiener filter was compared. This paperpresents the results of an experimental comparison ofa symmetric filter and a low-frequency Butterworth filter.

2. Statement of the problem

The transfer function of a low-frequency Butterworth filter (SFNF) of order n with a cutoff frequency at a distance D_0 from the point of origin is written by the formula [4]:

$$H(u,v)=\frac{1}{1+[D(u,v)/D_0]^{2n}}. \quad (1)$$

where the distance $D(u,v)$ is given by the formula

$$D(u,v)=\left[(u-M/2)^2+(v-N/2)^2\right]^{1/2}. \quad (2)$$

The transfer function of the CFFF does not have a discontinuity, which establishes an exact boundary between the transmitted and cut frequencies. For filters with a smooth transfer function, it is common practice to determine the location of truncating frequencies as a set of points, where the values of the function $H(u,v)$ become less than some of its largest value. In the case of the Butterworth function, $H(u,v)=0,5$ (less than 50% of the largest value equal to 1) for $D(u,v)=D_0$.

The Butterworth filter is characterized by the order of filtering, a parameter that determines the steepness of the filter's transfer function. For small filter order values, the transfer function has a smooth shape, close in form to the transfer characteristic of the Gaussian filter; at high values, the filter is characterized by a steep transfer function and approaches the characteristics of an ideal filter in shape.

The advantages of low-frequency Butterworth filters include a lesser manifestation of unwanted blurring effects and the appearance of false circuits compared with an ideal low-frequency filter. As the order of the low-frequency Butterworth filter increases, the effects of blurring also increases.

An experimental comparison of the Butterworth filter with a symmetric SE-filter is interesting. In this paper, a FIR filter with a zero phase was constructed, that is with a real frequency response [2, 5]. Hence the reference region will bea $N\times N$ square with side N and center in the point with coordinates $k_1=0,\ k_2=0,\ N$ is odd and $N_1=N_2=N$.

The filter has a reference region D:

$$\{D(n_1,n_2):0\le n_1<N_1,0\le n_2<N_2\}. \quad (3)$$

The count of the image $y(n_1,n_2)$, reconstructed with the FIR filter, can be determined from the readings $x(n_1,n_2)\in D$ of the distorted image of the relation:

$$y(n_1,n_2)=\sum_{k_1=0}^{N_1-1}\sum_{k_2=0}^{N_2-1}h(k_1,k_2)x(n_1-k_1,n_2-k_2), \quad (4)$$

where $h(k_1,k_2)$ is a two-dimensional impulse response.

To describe the one-dimensional frequency response in the radial direction, Iuse the QEFR filter model [9]:

$$S(\omega)=\begin{cases}a\omega^2,\ npu\ \ |\omega|\le\omega^*,\\ e^{-c|\omega|},\ npu\ \ |\omega|>\omega^*,\end{cases} \quad (5)$$

$$S(\omega^*)=\alpha\omega^{*2}=e^{-c\omega^*} \quad (6)$$

The approximation function of a symmetric SE filter is:

$$h(r) = \frac{e^{-c\omega^*}}{\pi}\left[\frac{\sin\omega^* r}{r} + \frac{2\cos\omega^* r}{r^2\omega} - \frac{2\sin\omega^* r}{r^3\omega^{*2}} + \frac{(c\cos(\omega^* r) - r\sin(\omega^* r))}{c^2 + r^2}\right] \quad (7)$$

As (7) shows, the approximation of the impulse response $h(r)$ is reduced by tuning two parameters: ω^* and c. Parameter ω^* is the abscissa of the greatest value of the frequency response of the FIR filter. Parameter *c* characterizes the degree of presence of high frequencies in the reconstructed image.

The advantages of this technology for constructing FIR filters include the possibility of achieving a high quality of image reconstruction by tuning a small number of parameters of a continuous function approximating the impulse response corresponding to a given frequency response.

3. Results of experiments

In the Butterworth filter, also to the distance D_0 outside of which all high-frequency components of the Fourier image, which is set from 0 to 256, are removed, the filter order from 0 to 20 is additionally specified.

The filter parameters were adjusted using the "Lena" test image (Figure 1).

Figure 1. Test drawing "Lena".

Next, the original image was distorted by simulatinga low-pass Gaussian filter for several variants of specifying σ. The degree of blurring was specified for $\sigma = 3$ and $\sigma = 5$. Figures 2(a) and 2(b) show a test image of "Lena" after introducing distortions with given parameters.

In this paper, the PSNR (peak-to-noise ratio) metric is used, defined by the formula: $PSNR = 10\log_{10}\left(\frac{MAX_I^{\,2}}{MSE}\right) = 20\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right)$, where MSE (mean square error) is the root-mean-square error and MAX_I is the greatest value received by the pixel of the image.

The following results were obtained in [3]. Metric PSNR evaluated the results on the test image "Lena". The greatest values of PSNR were 30.41 for $\sigma = 3$ and 27.18 for $\sigma = 5$. For the next experiment, the test image was distorted with a blur intensity of $\sigma = 3$. Additionally, the noise was "superimposed" in the range of 20 to 50 dB. Furthermore, depending on the values of the reference region, comparisons were made ofthe PSNR metric. With the size of the reference region $N = 5$, the best result was 29.92. With the increase in the reference region, the metric values decreased. With $N = 13$, they were already 26.66.

Table 1 shows the results of the experiment with a Butterworth filter. The image is distorted by a low-frequency Gaussian filter. Two variants with different degrees of blurring ($\sigma = 3$ and $\sigma = 5$) are used. In the second column, the size of the reference region N is recorded, in the third, the distance D_0, behind all components of high-frequency Fourier images, which can be specified in the range from 0 to 256. Column n indicates the order of the filter, which is in the range from 0 to 20. N is the size of the reference area. The last column shows the results of the PSNR metric.

Table 1 shows the results of an experiment with a Butterworth filter and a blur level of $\sigma = 3$ and $\sigma = 5$. For $\sigma = 3$, the greatest result is achieved with a filter with a cutoff frequency of 170 and a

filter order of 10. For the same cutoff frequency, but with different filter orders, the PSNR metric is changed. With the blur level $\sigma = 5$, the best result is achieved with a cutoff frequency of 170 and a filter order of 5.

Table 1. PSNP with noise in the image "Lena" at $\sigma = 3$ and 5 filtered by a Butterworth filter.

	N	D_0	n	*PSNR*		*N*	D_0	n	*PSNR*
σ =3	3	256	20	25.02	σ =5	3	256	20	26.29
	5	230	15	26.16		5	230	15	24.88
	3	200	10	26.14		3	200	10	24.94
	5	170	5	26.22		5	170	5	**26.45**
	3	170	10	**28.38**		3	170	10	26.44
	5	170	15	28.25		5	170	15	26.43
	3	150	1	26.21		3	150	1	26.18
	5	100	20	26.19		5	100	20	25.71
	3	50	10	25.99		3	50	10	26.01
	5	20	5	25.55		5	20	5	24.64

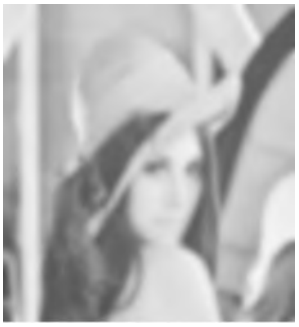
a)  b) 

Figure 2. Lena image: distorted with blur (a) $\sigma = 3$, (b) $\sigma = 5$.

Figure 3 shows the filtered images with the largest PSNR. Figure 3(a) shows the results of a symmetric filter with a blur level of $\sigma = 3$. Figure 3(b) shows the results of blurring with $\sigma = 5$. Figure 4 shows the results of the experiment with a Butterworth filter.

a)  b) 

Figure 3. Filtered image of "Lena": radially symmetric SE-filter with blurs(a) $\sigma = 3$, (b) $\sigma = 5$.

The problem of using the Butterworth filter is that onemust choose the cutoff frequency and the filter order by oneself. Choosing non-optimal values can lead to significant distortions in the filtered image. For example, if onetakes too small a cutoff frequency, most of the minor details of interest will be eliminated. If high values are selected, the processed image will be different from the test image, and the noise level will become insignificant. As the order of the low-frequency Butterworth filter increases, undesirable effects of blurring increase. Also, at very low cutoff frequencies, the so-called "ringing" effect or the Gibbs effect can appear on the filtered images. It is expressed in the appearance of false contours around real ones.

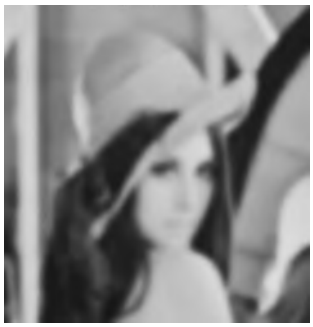
a)  b) 

Figure 4. Filtered image of "Lena" by the Butterworth filter with blur (a) $\sigma = 3$, (b) $\sigma = 5$.

4. Conclusion

As a result of these experiments, the following conclusions can be drawn. A radial-symmetric filter has a high quality of distorted imagesrecovery due to the change of a small number of continuous function parameters. Selecting the optimal filtering parameters is difficult when the Butterworth filter is used. It is necessary to conduct a large number of experiments to find the optimal values of the cutoff frequency and the order of the filtration.

Also, a symmetric filter has advantages in quality evaluationby the PSNR metric. The greatest values of this filter were 30.41 for $\sigma = 3$ and 27.18 for $\sigma = 5$. The Butterworth filter shows the following results: with a distortion ratio $\sigma = 3$, the PSNR metric was equal to 28.38; with $\sigma = 5$, the greatest result was equal to 26.45.

A visual assessment of the quality of images processed with a radial-symmetric filter and a Butterworth filter determined thatthe first one coped better with removing the Gaussian blur. Thus, it was found that the use of a radially symmetric SE-filter can provide a quality (by PSNR) superior to the result obtained by using of the Butterworth filter. It is also worth noting that edge effects and blurring, in contrast to the Butterworth filter, are very small in this case.

5. References

[1] Sorokin G A 2015 Low-pass filters *Computer technologies, management, radio electronics* **15** 100-107
[2] Pratt W 1982 *Digital Image Processing* (Moscow: The World*)* p 480
[3] Fursov V A 2016 Construction of FIR filters in a given parametric class of frequency characteristics for correction of defocusing *Computer Optics* **40(6)** 878-886 DOI: 10.18287/ 2412-6179-2016-40-6-878-886
[4] Gonzalez R and Woods R 2005 *Digital Image Processing* (Moscow: Technosphere) p1072
[5] Shapiro L and Stockman J 2006 *Computer Vision* (Moscow: BINOM) p 752
[6] Ziatdinov S I 2015 Pulse characteristic of Butterworth complex b and pass filter *Izvestiya Vysshikh Uchebnykh Zavedeniy* **5(8)** 653-658
[7] Fursov V A, Gavrilov A V, Goshin Ye V and Pugachev K G 2017 Conforming identification of the fundamental matrix in the image matching problem *Computer Optics* **41(4)** 559-563 DOI: 10.18287/2412-6179-2017-41-4-559-563
[8] Glumov N I and Myasnikov E V 2007 Method of selection of informative signs on digital images *Computer Optics* **31** 73-76

Acknowledgement

The article was supported by Doctor of Technical Sciences, Professor, Head of the Department of Supercomputers and General Informatics at the Samara University Fursov Vladimir Alekseevich.

Edge detection of objects on the satellite images

E E Kurbatova[1]

[1]Vyatka State University, Moskovskaya 36, Kirov, Russia, 610006

Abstract. Image segmentation is an important stage in image processing. The approach for the satellite images segmentation based on objects edge detectionis proposed. The approach usesthe random Markov fields as a mathematical model of an image. It is proposed to use the methods of contour and texture segmentation on different color components of a satellite image. The contour segmentation detects objects with different colors. It is applied to component with color information. The transition probability in two-dimensional Markov chains is used as a texture feature. The texture segmentation is applied to component with brightness information. The simulation results of the proposed approach in different color models, such as RGB, HSV, Lab, are presented. The accuracy of detecting contours was estimated using the set of test images on the base of five criteria. The use of a combination of color and texture characteristics of regions, made it possible to improve the accuracy of objects edge detection.

1. Introduction

Remote sensing data are widely used in different applications, including agriculture, forestry and water management, monitoring of the environment and emergencies, urban planning, cartography, etc. The thematic processing is one of the ways to process the satellite images in such systems. It includes detection, decoding and objects recognition stages. Using the thematic decoding of the satellite images, it is possible to allocate different classes of objects, such as forests, fields, rivers, urban zones, etc. [1] The obtained decoding results can be used for calculating the characteristics of objects and for tracing their changes over time. It is needed to apply the complex approach, which consist of several continues stages, for decoding of the satellite images. Different image processing methods can be used on different stages.

In general, image decoding consists of such stages:

- image acquisition;
- image enhancement (filtering, contrast enhancement, increase of resolution, etc.);
- object detection (edge analysis, segmentation into homogeneous regions);
- objects classification (sorting the allocated objects to finite number of classes).

Each stage of such process uses data obtained on the previous stage. Therefore, the quality of each stage affects the accuracy of the recognition results. Usually at each next stage the more complex algorithms are used. They require more time for processing and have the less degree of their automation. Therefore, it is preferable to use algorithms, which have a small number of parameters for setting, require small computational resources and minimum operator participation. At the same time they must provide high quality processing. It is especially actually for the algorithms applying at the first stages.

This work is about the object detection stage, the main method of which is segmentation. Different features can be used for image segmentation. Among them are object brightness, color, texture, shape, etc. In general, all segmentation methods can be divided into two classes: methods of contour analysis,

and texture methods. Contour methods are based on the objects edge detection for some feature [2-4]. Texture methods are methods, which find homogeneous regions. Such regions are characterized that a texture feature is unchanged or changesa little within the region. At the same time, it varies significantly in different regions. As a texture feature can be used different statistical, structural, morphological and spectral image characteristics [9-12]. But often it is not enough to use only one characteristic for object detection. Therefore, a combination of different features and algorithms are often used in the modern approaches for image segmentation [13-17].

In this work, the approach for image segmentation of satellite images is proposed. It is based on objects edges detection using the color and texture information. It increases the accuracy of objects edge detection. The approach uses the mathematical model based on Markov random fields for image description.

2. Image segmentation method

In the previous work [18], the mathematical model based on Markov random fields has been proposed for image description. Based on this model in some works [19,20] the contour and texture segmentation methods have been developed. They provide high efficiency and have low computational complexity. In this work it is proposed to use these methods jointly.

2.1. Mathematical model of an image

According to the used model, *g*-bits digital halftone images (DHI) are represented by the set of *g* bit binary images (BBI). Each BBI is the superposition of two one-dimensional Markov random chains with two equiprobablestatesM_1 and M_2 and matrices of transition probabilities in the horizontal and vertical directions:

$$ {}^1\Pi = \left\| \begin{matrix} {}^1\pi_{11} & {}^1\pi_{12} \\ {}^1\pi_{21} & {}^1\pi_{22} \end{matrix} \right\|, \; {}^2\Pi = \left\| \begin{matrix} {}^2\pi_{11} & {}^2\pi_{12} \\ {}^2\pi_{21} & {}^2\pi_{22} \end{matrix} \right\| \tag{1} $$

Figure 1 shows the *l*th BBI divided into areas $F_i(i = \overline{1,4})$ that are Markov chains of different dimensions. F_2 and F_3regions are one-dimensional Markov chains. F_4 region is two-dimensional Markov chain. As shown in figure 2, the neighborhood of the image element in this region consists of three elements.

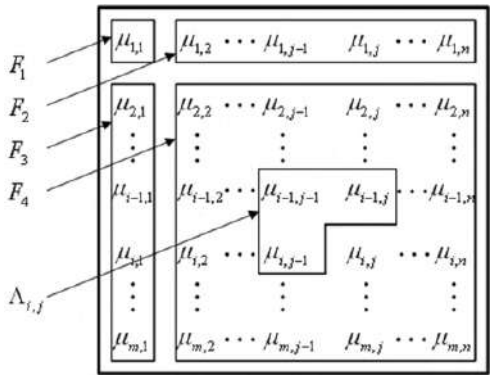
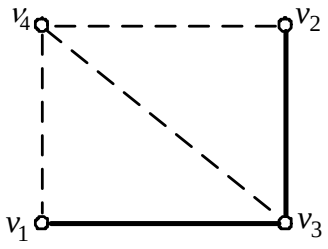


Figure 1. The areas of binary Markov random field.



$v_1 = \mu_{i,j-1};\ v_2 = \mu_{i-1,j};$
$v_3 = \mu_{i,j};\ v_4 = \mu_{i-1,j-1};$

Figure 2. The fragment of F_4region.

The entropy approach was applied for calculating the probabilities of the binary elements states. Thus, the amount of information in the elementv_3 relative to the states of the neighboring elements$v_1, v_2$is calculated by equation (2) [19].

$$ I(v_3|v_1, v_2) = -log\frac{w(v_3|v_1)w(v_3|v_2)}{w(v_3|v_2, v_1)}, \tag{2} $$

where $w(v_3|v_1)$, $w(v_3|v_2)$ are one-dimensional densities of transition probability of the neighboring elements, $w(v_3|v_2, v_1)$ is the density of transition probability in two-dimensional Markov chain.

The transition probability density in the binary two-dimensional Markov chain can be expressed by equation (3),where $\delta(\cdot)$is the delta function.

$$w(v_3|v_2,v_1)=\sum_{i,j,q=1}^{2}\pi\left(v_3=M_i|v_1=M_j,v_2=M_q\right)\times\delta\left(v_1-M_j\right)\times\delta\left(v_2-M_q\right) \quad (3)$$

Taking into account the equation (3), the transition probability matrix for various combinations of the neighboring elements states has the form (4).

$$\Pi=\begin{Vmatrix}\pi_{iii} & \pi_{iij}\\ \pi_{iji} & \pi_{ijj}\\ \pi_{jii} & \pi_{jij}\\ \pi_{jji} & \pi_{jjj}\end{Vmatrix}=\begin{Vmatrix}\alpha_1 & \alpha_1'\\ \alpha_2 & \alpha_2'\\ \alpha_3 & \alpha_3'\\ \alpha_4 & \alpha_4'\end{Vmatrix};\ i,j=\overline{1,2};\ i\neq j \quad (4)$$

The elements of this matrix are related with the elements of the ${}^1\Pi$, ${}^2\Pi$matrices by the relations (5).

$$\begin{aligned}\alpha_1=\pi_{iii}=\pi(v_3=M_1|v_1=M_1;v_2=M_1)=\ {}^1\pi_{ii}\cdot\ {}^2\pi_{ii}/\ {}^3\pi_{ii};\ \alpha_4=1-\alpha_1;\\ \alpha_2=\pi_{iji}=\pi(v_3=M_1|v_1=M_1;v_2=M_2)=\ {}^1\pi_{ii}\cdot\ {}^2\pi_{ij}/\ {}^3\pi_{ij};\ \alpha_3=1-\alpha_2,\end{aligned} \quad (5)$$

where ${}^3\pi_{ij},i,j=\overline{1,2},i\neq j$ are the elements of transition probability matrix ${}^3\Pi=\ {}^1\Pi\times\ {}^2\Pi$.

2.2. Texture segmentation method

This method [19,20] is based on two-dimensional mathematical model of an image. In general, texture is a region where some statistical properties are constant or change slowly. The estimate of transition probability in two-dimensional Markov chain is used as a texture feature. It is calculated using the sliding window method.

For the first line of window the estimate of transition probability ${}^1\hat{\pi}_{ii}$ for horizontal is calculated as[18]

$${}^1\hat{\pi}_{ii}=1-\frac{2p_1}{\hat{\chi}^{(l)}}, \quad (6)$$

where $\hat{\chi}^{(l)}$ is the estimate of the average sequence length of the identical BBI elements; p_1 is the initial probability ($p_1=0{,}5$).

From the second line, the estimate of transition probability for vertical ${}^2\hat{\pi}_{ii}$ and estimate $\hat{\pi}_{iii}$ of transition probability in two-dimensional Markov chain are calculated by the matrix (4).

All the obtained estimates are averaged within the window to produce a mean estimate of transition probability $\tilde{\pi}_{iii}$:

$$\tilde{\pi}_{iii}^{(r,k)}=\frac{1}{m\cdot n}\sum_{r=1}^{m}\sum_{k=1}^{n}\hat{\pi}_{iii}^{(r,k)}, \quad (7)$$

where *m*, *n* are height and width of the sliding window.

This mean value is used as a texture feature for the central element of the window.

A window of fixed size is moved from left to the right and top to bottom on *l*th BBI to get texture feature for each image element.

Then image element is marked by comparing the calculated texture feature with the threshold.

As a result, each image element has the label corresponding to a certain texture. The threshold can be selected on the basis of the analysis of texture feature histogram. If there are several textures on the image, it is needed to select several thresholds.

In the case of color image processing, each color component can be represented as the DHI. All color components are processed separately. The threshold is selected for each component. The segmentation results obtained on different color components are combined into a single color image. On this image different colors correspond to the regions of different textures.

2.3. Contour segmentation method

To detect objects edges the amount of information between the elementv_3 and the various combinations of the neighboring elements is calculated. It is determined with the matrix (4) and the equation (2) for each element of *l* th BBI.

The amount of information in the elementν_3 will be minimal, if the neighboring elementsν_1, ν_2 have the same states with theν_3 [18].

On the edge of other brightness region one or two neighboring elements have different states with ν_3.In this case the amount of information in the element ν_3 is increased.If the amount of information in the element$\nu_3$is greater than to h the pixel belongs to the contour. The element $\nu_3$belongs to homogeneous region in the other case.

The threshold h is calculated for each BBI taking into account the minimal amount of information and the amount of information, when one of the neighboring elements has a different state.

$$h = 0{,}5 * \left(I(\nu_3 = M_i|\nu_1 = M_i, \nu_2 = M_i) + I\left(\nu_3 = M_i\middle|\nu_1 = M_i, \nu_2 = M_j\right)\right). \quad (8)$$

It is supposed that the transition probability matrices are a priori known.

In the case of color image the contours are detected on each color component. Then the contour maps of each component are combined to asingle contour image.

2.4. Objects edge detection on satellite images

Most often, satellite images are multispectral (multicomponents). They are displayed as the color images, which have three channels. The three channels may be three multispectral bands of the same scene. Various types of color images can be prepared based on the different band combinations. True-color images use visible red, green and blue bands. False-color images use the combination of near infrared, red and green bands. Pseudocolor images contain medium and near infrared and green bands.[21,22].

Color is an important characteristic of objects that often simplifies there segmentation and recognition. There are several ways to specify colors. The *RGB* color model is the simplest and the most nature. In this case a color image consists of three components (red, green and blue), described by their corresponding intensities. This model has large color coverage, but it is poorly suited for processing tasks, because the color and the brightness information are encoded in the same three channels. In *Lab* and *HSV* color models color and brightness information are separate into different components. So they are much more convenient for processing.

In *Lab* color model, the *a* and *b* components encode color. The first component *a* determines the color position between green and magenta, the second component *b* – its position between blue and yellow. The third component *L*is independent of color information and encodes brightness only.

The HSV color model uses only one channel to describe color. The image contains of three components. They are the hue *H*, the saturation *S* and the value (or brightness) *V* components. The hue *H* component is the color position; the saturation *S* component is the amount of gray in the color. The value *V* is the brightness or intensity of the color.

The main idea of the proposed approach is that different segmentation methods are appliedto different components of a color image. The texture segmentation method, which was described in subsection 2.2, is used on the component with brightness information. The contour segmentation method, described in the previous subsection, is applied to the component with color information. As a result of texture segmentation the regions of different textures are marked by different labels. To get the contour map the second stage after texture segmentation was added. On the second stage the contour segmentation is applied to the marked image output by the texture segmentation. Then the contours detected on different components are combined into single contour image. Thus, the proposed approach takes into account the color and texture information for image segmentation. This improves the accuracy of the objects edges detection.

3. Simulation results

To estimate the performance of the proposed approach, we have simulated it on images in *RGB*, *Lab* and *HSV* color models. The software used was Matlab. The experiments were designed to analyse the role of texture and color characteristics played on image segmentation. In the first experiment we used *RGB* color model and the contour segmentation method discussed earlier in subsection 2.3. In this case the image was divided into three components (*R*, *G* and *B* components), and each component was processed by the contour segmentation method. In the second experiment we applied the texture

segmentation method to each component of *RGB* image. To detect the edges, the contour segmentation method was applied to the segmented regions output by the texture segmentation. Then, we simulated the combination of texture and contour segmentations on the images in *Lab* color model. The component *L*(lightness) was proposed by the texture segmentation method with the next contour detection. The color components *a* and *b* were processed by the contour segmentation method. In the final experiment the images in *HSV* color model were processed. The texture segmentation was applied to V component (brightness), and the contour segmentation was applied to *H* component (hue). The *S* component (saturation) was not used.

The qualitative performance was evaluated by comparing the segmenting results with a benchmark. For lack of benchmarks for the real satellite images, we tested the proposed approach on the images from the Berkeley Segmentation Dataset [23]. It contains the human-annotated ground truth segmentations corresponding to each test image.

The quantitative comparison for performance is based on five measure metrics. They are *FOM*(Figure of Merit) [24], *RMS* (root mean squared error) [24], *P*(precision), *R*(recall), *F*-measure [25]. Because the result of the proposed approach is contour image, we used metrics based on contour representation for quantitative evaluation of segmentation accuracy.

The *FOM* (figure of merit) is an empirical distance between the image with contours from the segmentation results*g* and the corresponding ground truth*f*. It shows how similar the ground truth and the segmentation result are. The *FOM* is defined as:

$$FOM = (\max\{card(f), card(g)\})^{-1} \cdot \sum_{i=1}^{card(g)} \left(1 + d_i^2\right)^{-1}, \quad (9)$$

where *card(f)* is the number of contour elements in the image *f*, *card(g)* is the number of the contour elements in the image *g*, d_i is the distance between *i*th pixel in *f* and the nearest pixel to it in *g*.

The *RMS* is the root mean squared error. It shows how different the groundtruth and segmented image are. The *RMS* is defined as

$$RMS = \left(\frac{1}{w \cdot h} \cdot \sum_{i=1}^{w \cdot h} (f_i - g_i)^2 \right)^{1/2} \quad (10)$$

where *w* and *h* are width and height of the image, f_i and g_i are the intensities of *i*th pixel in the ground truth and segmented contour image.

The *P* (precision) is the relation between the correctly detected contour elements and all elements detected as contours on the image *g*. The *R* (recall) is the relation between the correctly detected contour element and all elements detected as contours on the ground truth image *f*. They are calculated by the equations:

$$P = \frac{TP}{card(g)}; R = \frac{TP}{card(f)}, \quad (11)$$

where *TP* is the number of true positives decisions of the algorithm, i.e. the number of image elements, which are contours both on the segmented image and the ground truth.

F-measure is a widely used metric to evaluate segmentation results that combines precision and recall. It is the weighted harmonic mean of precision and recall. The *F*-measure is calculated by the equation (12).

$$F = 2 \cdot \frac{P \cdot R}{R + P}. \quad (12)$$

The *FOM*, *P*, *R* and *F*-measure are the higher the better segmentation results. The *RMS* is the lower the higher segmentation accuracy.

Table 1 gives the values of these qualitative metrics for segmentation results in different color models (the bold letters means best value). The values are averaged over all processed test images.

Figure 3 presents the segmentation results on the true-color satellite image (figure 3a) using different algorithms and color models. We use the highest BBI for segmentation, because there are the more significant region details on it.

Table 1. Quality criteria for test images.

Segmentation method	Quality criteria				
	FOM	RMS	R	P	F
Contour segmentation based on RGB color model	**0.144**	0.437	**0.772**	0.164	0.249
Texture segmentation based on RGB color model	0.137	0.322	0.519	0.197	0.274
Segmentation based on Lab color model	0.127	**0.268**	0.384	**0.256**	**0.287**
Segmentation based on HSV color model	**0.140**	**0.304**	**0.543**	**0.247**	**0.322**

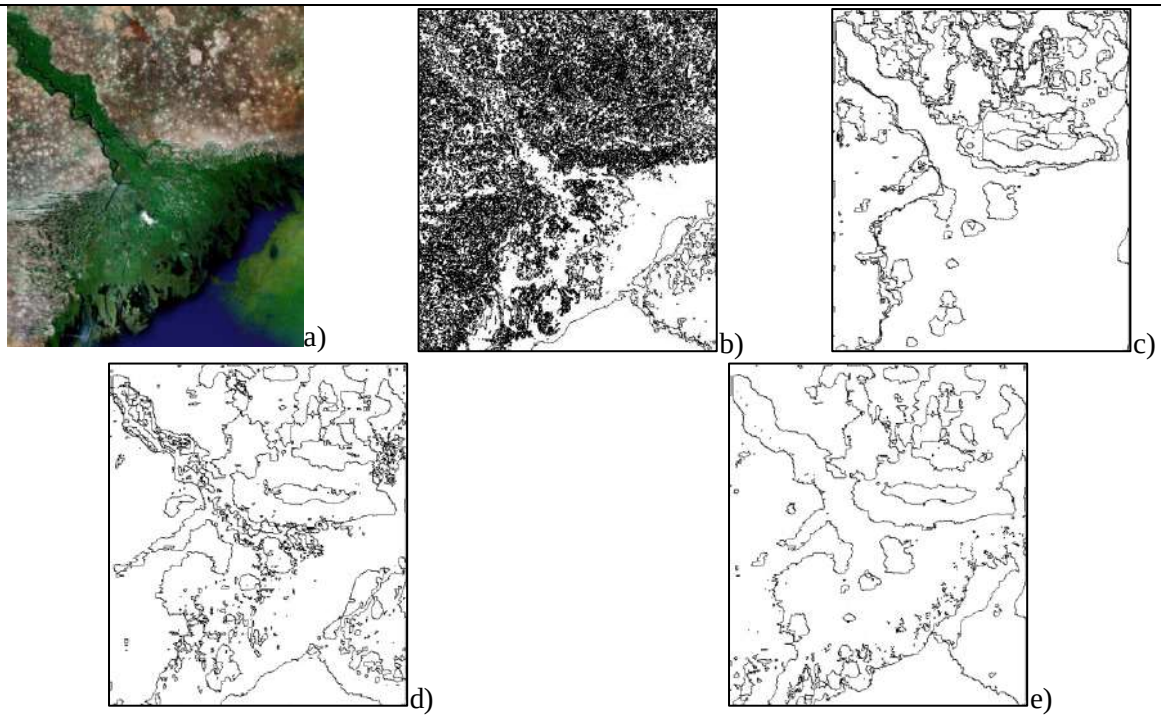


Figure 3. Segmentation results.

Figure 3b shows the results of contour segmentation using the *RGB* color model. The edges are detected by the method based on two-dimensional Markov chains on each color component (*R*, *G*, *B*).Then all contours are combined into one resulting image (figure 3b). Figure 3c shows the result of the second experiment. Here the texture segmentation method is applied to each color component of *RGB* color image. It is assumed that the initial image contains only two different textures. On the segmented image the regions of the first texture are marked as "1", and the regions of the another texture - as "0". As a result of texture segmentation binary image were obtained. The contour segmentation method is applied to the texture segmentation results. Thus, the edges of texture regions are detected. The contour images of three components are combined into one resulting image (figure 3c).

Figure 3d illustrated image segmentation result in *Lab* color model. Here only the final resulting contour image is shown. It is a combination of contour segmentation results of *a* and *b* components and contours of the texture regions detected on the *L* component. Figure 3e shows the resulting contour image obtained on satellite image in *HSV* color model. It is a combination of contours detected on the *H* component and contours of the texture regions detected on the *V* component.

4. Conclusion

From the simulation results, a conclusion can be draw that the contour segmentation method allows detecting edges of regions with different colors on the image. But it gives unsatisfactory results for texture regions. Such texture regions are often observed on satellite images. They do not have pronounced edges in terms of brightness or color. This leads to the significant over-segmentation. This case is shown on figure 3b. The texture segmentation allows detecting the edges of texture regions

more clearly. But in the same time, some edges between objects of different colors can be lost what is illustrated on figure 3c.The sea region and the flat part of coast are differ significantly in color. But in the same time, they have close values of transition probability of the elements in two-dimensional Markov chain. Consequently, the algorithm is detected them as one region.

The segmentation based on *Lab* and *HSV* color models provides similar results. Herewith, the edges of objects of different colors and the edges of different texture regions are detected more precisely. The use of *HSV* color model has another advantage that it is enough to process only two components. This allows to reduce the computational time significantly.

The results of simulation on the test images confirm these conclusions. In table 1, segmentation based on HSV color model gets one of the best values in the most quality criteria. We can also find that the contour segmentation based on *RGB* color model has the best values in *FOM* and *P*. The reason for this is that such segmentation belongs to over-segmentation which gives more details. Therefore, there are a lot of coincidences between the ground truth and segmented contours. But in addition in this case there are also a lot of false detected contour pixels. As a result such segmentation has the worst values in *RMS* and *R*. So the results of contour segmentation based on *RGB* color model are unsatisfactory.

Thus, the proposed approach consists of the joint use of contour and texture segmentation on the different image components. It takes into account color and texture characteristics of objects for segmentation. Due to this the segmentation results are more accurate. It is recommended to use HSV color model, because it shown the best results.

5. References

[1] Vorobiova N S, Sergeyev V V and Chernov A V 2016 Information technology of early crop identification by using satellite images *Computer Optics* **40(6)** 929-938 DOI: 10.18287/2412-6179-2016-40-6-929-938
[2] Gonzalez R C and Woods R E 2008 *Digital image processing* (New York: Prentice Hall) p 954
[3] Verma S and Chugh A 2016 An increased modularity based contour detection *International Journal of Computer Applications* **135(12)** 41-44
[4] Swami D and Chaurasia B J 2017 Super-pixel and Neighborhood based contour detection *Comp. & Math. Sci.* **8(6)** 226-234
[5] Borne F and Viennois G 2017 Texture-based classification for characterizing regions on remote sensing images *Journal of Applied Remote Sensing* **11(3)**
[6] Hemalatha S and Anouncia S M 2017 Unsupervised segmentation of remote sensing images using FD based texture analysis model and ISODATA *International Journal of Ambient Computing and Intelligence* **8(3)** 58-75
[7] Prudente V, Da Silva B, Johann J, Mercante E and Oldoni L 2017 Comparative assessment between per-pixel and object-oriented for mapping land cover and use *Journal of the Brazilian Association of Agricultural Engineering* **37(5)** 1015-1027
[8] Abbas A W, Minallh N, Ahmad N, Abid S A R, Khan M A A 2016 K-Means and ISODATA Clustering Algorithms for Landcover Classification Using Remote Sensing *Sindh Univ. Res. Jour. (Sci. Ser.)* **48(2)** 315-318
[9] Baya A E, Larese M G and Namias R 2017 Clustering stability for automated color image segmentation *Expert Systems with Applications* **86** 258-273
[10] Li M, Zhang S, Zhang B, Li S and Wu C 2014 A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information *European Journal of Remote Sensing* **47** 389-411
[11] Haralick R M 1979 Statistical and structural approaches to texture *Proceedings of the IEEE* **67(5)** 786-804
[12] Hemalatha S and Anouncia S M 2016 A computational model for texture analysis in images with fractional differential filter for texture detection *International Journal of Ambient Computing and Intelligence* **7(2)** 93-113
[13] Zhang J, Gao Y W and Feng S W 2015 Image segmentation with texture clustering based JSEG *International Conference on Machine Learning and Cybernetics (ICMLC)* DOI:

10.1109/ICMLC.2015.7340623
[14] Hu Y, Li Z, Li P, Ding Y and Liu Y 2017 Accurate and fast building detection using binary bag-of-features *ISPRS Hannover Workshop: HRIGI 17 – CMRT 17 – ISA 17 – EuroCOW 17* **XLII-1/W1** 613-617
[15] Liu L X, Fan S M, Ning X D and Liao L J 2017 An efficient level set model with self-similarity for texture segmentation *Neurocomputing* **266** 150-164
[16] El Merabet Y, Meurie C, Ruichek Y, Sbihi A and Touahni R 2015 Building roof segmentation from aerial images using a line-and region-based watershed segmentation technique *Sensors* **15(2)** 3172-3203
[17] Myasnikov E V 2017 Hyperspectral image segmentation using dimensionality reduction and classical segmentation approaches *Computer Optics* **41(4)** 564-572 DOI: 10.18287/2412-6179-2017-41-4-564-572
[18] Petrov E P, Trubin I S, Medvedeva E V and Smolskiy S M 2013 Mathematical Models of Video-Sequences of Digital Half-Tone Images *Integrated models for information communication systems and net-works : design and development* (IGI Global) 207-241
[19] Medvedeva E V and Kurbatova E E 2015 Image segmentation based on two-dimensional Markov chains *Computer Vision in Control Systems-2. Innovations in practice* (Springer International Publishing Switzerland) 277-295
[20] Kurbatova E E, Medvedeva E V and Okulova A A 2015 Method of isolating texture areas in images *Pattern Recognition and Image Analysis* **25(1)** 47-52
[21] Burnett C and Blaschke T 2003 A multi-scale segmentation/object relationship modelling methodology for landscape analysis *Ecological Modelling* **168(3)** 233-249
[22] Krautsou S L 2008 *Processing of remote sensing images (methods analysis)* (Minsk: UIIP NAS Belarus) p 256
[23] Berkeley Segmentation Dataset (Accsess mode: http://www.eecs.berkeley.edu/Research/ Projects/CS/vision/grouping/segbench) (01.11.2017)
[24] Zhang Y 2006 *Advances in Image And Video Segmentation* (USA: IRM Press) p 473
[25] Martin D, Fowlkes C and Malik J 2004 Learning to detect natural image boundaries using local brightness, color and texture cues *IEEE Trans. on Pattern analysis and Machine Intelligence* **26** 530-549

Comparative analysis of stochastic optimization algorithms for image registration

S Voronov[1], **I Voronov**[1] **and R Kovalenko**[1]

[1]Ulyanovsk State Technical University, Severniy Venets 32, Ulyanovsk, Russia, 432027

Abstract. This work is devoted to comparative experimental analysis of different stochastic optimization algorithms for image registration in spatial domain: stochastic gradient descent, Momentum, Nesterov momentum, Adagrad, RMSprop, Adam. Correlation coefficient is considered as the objective function. Experiments are performed on synthetic data generated via wave model with different noise-to-signal ratio.

1. Introduction

Digital image registration is a process by which the most accurate match is determined between two images, which may have been taken at the same or different times, by the same or different sensors, from the same or different viewpoints. The registration process determines the optimal transformation, which will align the two images. This has applications in many fields as diverse as medical image analysis, pattern matching and computer vision for robotics, as well as remotely sensed data processing. In all of these domains, image registration can be used to find changes in images taken at different times, or for object recognition and tracking.

Spatial domain methods operate directly on pixels, and the problem of the estimation of registration parameters $\overline{\alpha}$ becomes the problem of searching for the extreme point of a multi-dimensional objective function $\mathrm{J}(\mathbf{Z}, \overline{\alpha})$. The objective function measures the similarity between two images $\mathbf{Z}^{(1)} = \{z_{\bar{j}}^{(1)}\}$ and $\mathbf{Z}^{(2)} = \{z_{\bar{j}}^{(2)}\}$, where $\bar{j} \in \Omega$ are nodes of grid mesh Ω on which the images are defined. There is a wide variety of similarity measures that can be used as objective functions [1]. The decision of which objective function to choose is usually based on the specifics of images, deformation properties and conditions. Recently, objective functions from the theory of information are becoming more popular. Among these functions the most interesting is mutual information. It has been found to be especially robust for multimodal image registration and registration of images with great non-linear intensity distortion [2]. However, mutual information has some drawbacks. One of them is relatively high computational complexity.

The choice of optimization search technique depends on the type of problem under consideration. Traditional nonlinear programming methods, such as the constrained conjugate gradient, or the standard back propagation in neural network applications, are well suited to deterministic optimization problems with exact knowledge of the gradient of the objective function. Optimization algorithms have been developed for a stochastic setting where randomness is introduced either in the noisy measurements of the objective function and its gradient, or in the computation of the gradient approximation. Stochastic gradient ascend (descend) is one of the most powerful technique of this class [3]. It is an iterative algorithm, where registration parameters can be found as follows [4]:

$$\hat{\overline{\alpha}}_t = \hat{\overline{\alpha}}_{t-1} - \mathbf{\Lambda}_t \overline{\beta}_t (\mathrm{J}(Z_t, \overline{\alpha}_{t-1})),$$

where $\overline{\beta}$ – gradient estimation vector of the objective function J obtained using not each pixel in the images but a sample Z_t taken randomly on each iteration, $\mathbf{\Lambda}_t$ – positive-definite gain (learning rate) matrix: $\mathbf{\Lambda}_t = \|\lambda_{it}\|$, $\lambda_{it} > 0$, $i = \overline{1,m}$; m – the number of registration parameters.

The main disadvantage of this optimization algorithm is presence of a large number of local extreme points of the objective function due to the use of small samples and relatively short working range in terms of registration parameters to be estimated. To overcome these problems the number of sample elements can be increased. However, this leads to significant increase in computational efforts. Another significant problem is choosing the hyperparameter $\mathbf{\Lambda}_t$ as it largely affects not only the convergence rate but also the estimation accuracy. Thus, the problem of optimization of stochastic gradient algorithm for image registration is an important, especially for real-time processing systems. To overcome the mentioned problems some modifications of the "classical" stochastic gradient descent have been proposed. This paper is devoted to comparative experimental analysis of these modifications: Momentum, Nesterov momentum, Adagrad, RMSprop, Adam. These algorithms are very effective, especially in training artificial neural networks [5].

2. Stochastic optimization algorithms

Let us consider the most popular modifications of stochastic gradient descent optimization algorithm which can be used for solving image registration problem.

2.1. Momentum

The idea behind Momentum optimization is quite simple [6]: if one imagine a bowling ball rolling down a gentle slope on a smooth surface: it will start out slowly, but it will quickly pick up momentum until it eventually reaches terminal velocity (if there is some friction or air resistance). In contrast, regular gradient descent will simply take small regular steps down the slope, so it will take much more time to reach the bottom. Recall that gradient descent simply updates the parameter estimates $\hat{\overline{\alpha}}$ by directly subtracting the gradient of the cost function with regards to the parameters $\mathrm{J}(\mathbf{Z}, \overline{\alpha})$ multiplied by the learning rate $\lambda > 0$. It does not care about what the earlier gradients were. If the local gradient is tiny, it goes very slowly. Momentum optimization cares a great deal about what previous gradients were: at each iteration, it adds the local gradient to the momentum vector $\overline{\mathrm{m}}$ multiplied by the learning rate λ, and it updates the weights by simply subtracting this momentum vector. In other words, the gradient is used as an acceleration, not as a speed. To simulate some sort of friction mechanism and prevent the momentum from growing too large, the algorithm introduces a new hyperparameter h, simply called the momentum, which must be set between 0 (high friction) and 1 (no friction). A typical momentum value is 0.9. Thus, the equation for parameter estimate updates can be written as follows:

$$\hat{\overline{\alpha}}_t = \hat{\overline{\alpha}}_{t-1} - \overline{\mathrm{m}}_t,$$

where $\overline{\mathrm{m}}_t = h\overline{\mathrm{m}}_{t-1} + \mathbf{\Lambda}_t \overline{\beta}_t(\mathrm{J}(Z_t, \overline{\alpha}_{t-1}))$.

One can easily verify that if the gradient remains constant, the terminal velocity (i.e. the maximum size of the weight updates) is equal to that gradient multiplied by the learning rate λ multiplied by $\frac{1}{1-h}$. For example, if $h = 0.9$, then the terminal velocity is equal to 10 times the gradient times the learning rate, so Momentum optimization ends up going 10 times faster than "classical" stochastic gradient descent. This allows Momentum optimization to escape from plateaus much faster.

2.2. Nesterov momentum

In [7] the author proposes one small variant to Momentum optimization which is almost always faster than vanilla Momentum optimization. The idea behind Nesterov momentum optimization consists in measuring the gradient of the cost function not at the local position but slightly ahead in the direction

of the momentum. Hence, the only difference from vanilla Momentum optimization is that the gradient is measured on t-th iteration at the point $\hat{\bar{\alpha}}_{t-1} + h\bar{\mathrm{m}}_t$ rather than at $\hat{\bar{\alpha}}_{t-1}$:

$$\hat{\bar{\alpha}}_t = \hat{\bar{\alpha}}_{t-1} - \left(h\bar{\mathrm{m}}_{t-1} + \mathbf{\Lambda}_t \bar{\beta}_t \left(\mathrm{J}(Z_t, \hat{\bar{\alpha}}_{t-1} + h\bar{\mathrm{m}}_t\right)\right).$$

This small tweak works because in general the momentum vector will be pointing in the right direction (i.e., toward the optimum), thus it will be slightly more accurate to use the gradient measured a bit farther in that direction rather than using the gradient at the original position. Figure 1 shows this effect. Here $\lambda\beta_1$ represents the gradient of the cost function measured at the starting point $\hat{\bar{\alpha}}_{t-1}$, and $\lambda\beta_2$ represents the gradient at the point located at $\hat{\bar{\alpha}}_{t-1} + h\bar{\mathrm{m}}_t$. As one can see, the Nesterov update ends up faster optimizers slightly closer to the optimum. After a while, these small improvements add up and the procedure ends up being significantly faster than regular Momentum optimization. Moreover, we should note that when the momentum pushes the weights across a valley, $\lambda\beta_1$ continues to push further across the valley, while $\lambda\beta_2$ pushes back toward the bottom of the valley. This helps reduce oscillations and thus converges faster.

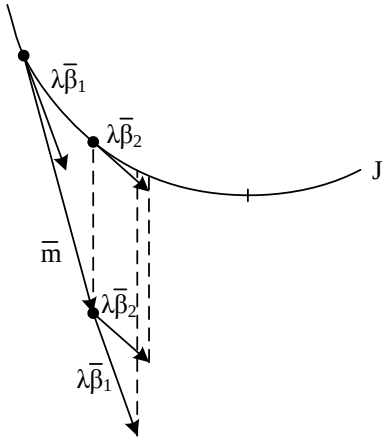


Figure 1. Difference between Momentum and Nesterov momentum optimization.

2.3. Adagrad

If we consider the elongated bowl problem again: gradient descent starts by quickly going down the steepest slope, then slowly goes down the bottom of the valley. However, it would be better if the algorithm could detect this early on and correct its direction to point a bit more toward the global optimum.

The Adagrad algorithm [8] achieves this by scaling down the gradient vector along the steepest dimensions:

$$\hat{\bar{\alpha}}_t = \hat{\bar{\alpha}}_{t-1} - \mathbf{\Lambda}_t \bar{\beta}_t \left(\mathrm{J}\left(Z_t, \hat{\bar{\alpha}}_{t-1}\right)\right) // \left(\bar{\mathrm{c}}_t + \varepsilon\right)^{1/2},$$

where $\bar{\mathrm{c}}_t = \bar{\mathrm{c}}_{t-1} + \bar{\beta}_t \left(\mathrm{J}\left(Z_t, \hat{\bar{\alpha}}_{t-1}\right)\right)^2$.

The first step of this algorithm on each iteration is accumulating the square of the gradients into the vector $\bar{\mathrm{c}}_t$. If the cost function is steep along the i-th dimension, then c_{it} will get larger and larger at each iteration. The second step is almost identical to "classical" stochastic gradient descent, but with one big difference: the gradient vector is scaled down by a factor of $\left(\bar{\mathrm{c}}_t + \varepsilon\right)^{1/2}$ (the // symbol represents the element-wise division, and ε is a smoothing term to avoid division by zero, typically set to 10^{-8}). In short, this algorithm decays the learning rate, but it does so faster for steep dimensions than for dimensions with gentler slopes. This is called an adaptive learning rate. It helps point the resulting updates more directly toward the global optimum. One additional benefit is that it requires much less tuning of the learning rate hyperparameter.

Adagrad often performs well for simple quadratic problems, but unfortunately it often stops too early when training neural networks. The learning rate gets scaled down so much that the algorithm ends up stopping entirely before reaching the global optimum.

2.4. RMSprop

Although Adagrad slows down a bit too fast and ends up never converging to the global optimum, the RMSProp algorithm [9] fixes this by accumulating only the gradients from the most recent iterations (as opposed to all the gradients since the beginning of training). It does so by using exponential decay in the first step:

$$\overline{\mathrm{c}}^{\mathrm{r}}{}_t = d\overline{\mathrm{c}}^{\mathrm{r}}{}_{t-1} + (1-d)\overline{\beta}_t\left(\mathrm{J}\left(Z_t, \hat{\overline{\alpha}}_{t-1}\right)\right)^2$$

where d is the decay rate and it is typically set to 0.9.

Except on very simple problems, this optimizer almost always performs much better than Adagrad. It also generally performs better than Momentum optimization and Nesterov momentum. In fact, it was the preferred optimization algorithm of many researchers until Adam optimization came around.

2.5. Adam

Adam [10] which stands for adaptive moment estimation, combines the ideas of Momentum optimization and RMSProp: just like Momentum optimization it keeps track of an exponentially decaying average of past gradients, and just like RMSProp it keeps track of an exponentially decaying average of past squared gradients:

$$\hat{\overline{\alpha}}_t = \hat{\overline{\alpha}}_{t-1} - \mathbf{\Lambda}_t \overline{\mathrm{m}}^{\mathrm{A}}{}_t \,//\left(\overline{\mathrm{c}}^{\mathrm{A}}{}_t + \varepsilon\right)^{1/2},$$

$$\overline{\mathrm{m}}^{\mathrm{A}}{}_t = \frac{d_1\overline{\mathrm{m}}^{\mathrm{A}}{}_{t-1} + (1-d_1)\overline{\beta}_t\left(\mathrm{J}\left(Z_t, \hat{\overline{\alpha}}_{t-1}\right)\right)}{1-d_1},$$

$$\overline{\mathrm{c}}^{\mathrm{A}}{}_t = \frac{d_2\overline{\mathrm{c}}^{\mathrm{A}}{}_{t-1} + (1-d_2)\overline{\beta}_t\left(\mathrm{J}\left(Z_t, \hat{\overline{\alpha}}_{t-1}\right)\right)^2}{1-d_2}.$$

One can notice the similarity of Adam update rule to both Momentum optimization and RMSProp. The only difference is that it computes an exponentially decaying average rather than an exponentially decaying sum for $\overline{\mathrm{m}}^{\mathrm{A}}{}_t$ and $\overline{\mathrm{c}}^{\mathrm{A}}{}_t$, but these are actually equivalent except for a constant factor as the decaying average is just $(1-d_1)$ and $(1-d_2)$ times the decaying sum respectively. The momentum decay hyperparameter d_1 is typically initialized to 0.9, while the scaling decay hyperparameter d_2 is often initialized to 0.999. As earlier, the smoothing term $\varepsilon\epsilon$ is usually initialized to a tiny number such as 10^{-8}. In fact, since Adam is an adaptive learning rate algorithm like both Adagrad and RMSProp, it requires less tuning of the learning rate hyperparameter $\mathbf{\Lambda}_t$.

3. Experiments and analysis

3.1. Synthetic data

For efficiency analysis of differentit optimization algorithms it is reasonable to use simulated images whose intensity probability distribution function and correlation function can be priori defined during their synthesis. In conducted experiments simulated images based on wave model [11, 12] with intensity probability distribution function and correlation function close to Gaussian and with different correlation radius were used. In addition, an unbiased Gaussian noise was used in simulations. Figure 2 shows an example of such synthesized image.

In order to measure the performance of the optimization algorithms we tested them on images with different noise-to-signal ratio and with different μ – the number of points in the sample using for estimation of the gradient of the chosen objective function. Correlation coefficient [1] is chosen as an objective function to be optimized. Similarity model is considered as the deformation model to be estimated. For all of the below results the deformation parameters are the following: horizontal shift –

20 pixels to the right, vertical shift – 15 pixels upwards, clockwise rotation – 17 degrees, scale factor – 0.9. In each experiment, optimal hyperparameters were chosen experimentally as different algorithms better perform with different hyperparameters and their choice is out of the scope of this article.

The number of iterations before convergence of mismatch Euclidean distance [4] E which is an integral measure of registration parameters' convergence is used as the performance criterion. Moreover, all the results are averaged by 50 realizations to make them more consistent and reproducible.

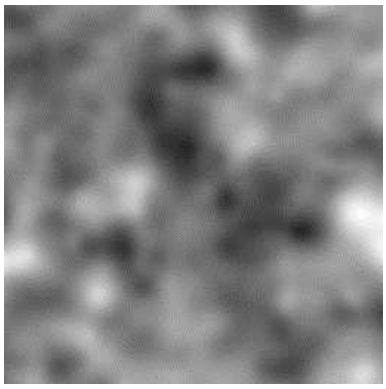
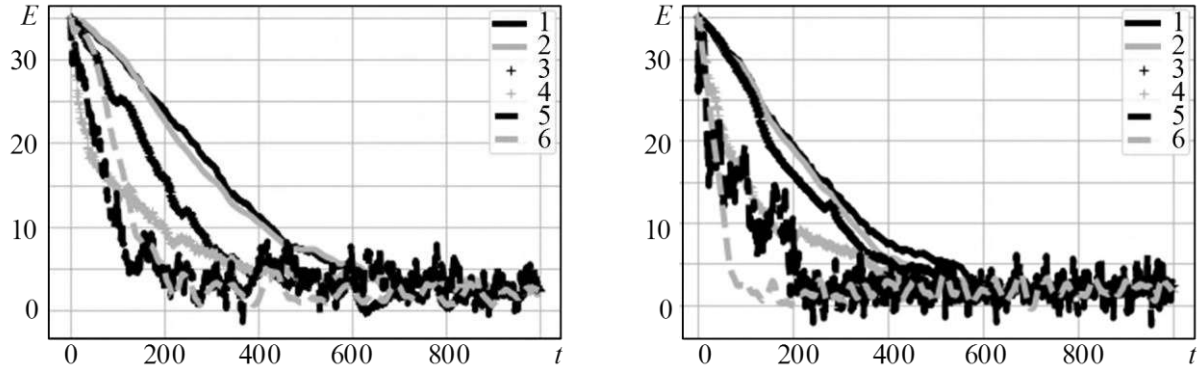


Figure 2. The image synthesized using wave model.

3.1.1. Different sample size

Figure 3 shows the convergence of mismatch Euclidean distance for the algorithms with different sample size μ. Hereafter, curve 1 corresponds to "classical" stochastic gradient descent, 2 – stochastic gradient descent with Momentum, 3 – Nesterov momentum (plus markers), 4 – Adagrad (plus markers), 5 – RMSprop (dashed line), 6 – Adam (dashed line).



a) $\mu = 5$ b) $\mu = 20$

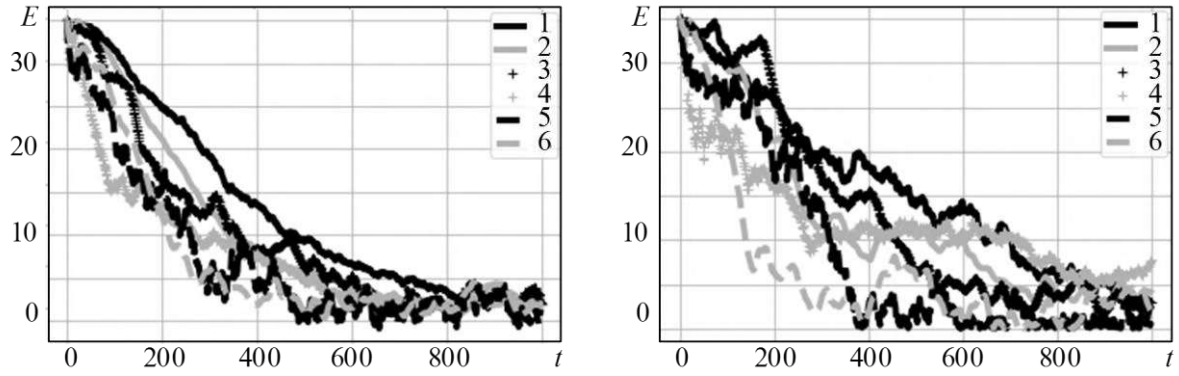
Figure 3. Mismatch Euclidean distance on the number of iterations for different sample size.

One can see that in both cases "classical" stochastic gradient descent shows the worst result as it starts to converge only after 800 iterations for $\mu = 5$ and after 700 iterations for $\mu = 20$. Momentum optimization algorithm performs almost identically but slightly better for $\mu = 20$. The best results in both cases are provided by Adam and RMSprop optimizations as they start to converge after 300 iterations for $\mu = 5$ and after 200 iterations for $\mu = 20$. Moreover, it is obvious that Adam algorithm in both situations has less variance then RMSprop, thus we can conclude that it is more stable and hence preferable. Adagrad and Nesterov momentum algorithms show close results in terms of number of iterations before convergence (500 iterations for $\mu = 5$ and after 450 iterations for $\mu = 20$), but in the beginning Adagrad has much faster convergence rate and with some optimization (e.g. increasing or dropping learning rates after a number of iterations) it possibly can outperform Nesterov momentum.

Also, we can conclude that all of the algorithms have better convergence rate with bigger sample size. It is reasonable from theoretical point of view because the objective function gradient estimates become less noisy.

3.1.2. Images with different signal-to-noise ratio

Let us test the algorithms in case of noisy images with different signal-to-noise ratio q. In this experiment we were using the sample size $\mu = 5$ for each algorithm and q. Figure 4 shows the convergence of algorithms for $q = 50$ and $q = 2$.



a) $q = 50$ b) $q = 2$

Figure 4. Mismatch Euclidean distance on the number of iterations for different signal-to-noise ratio.

As in the previous experiment, "classical" stochastic gradient descent shows the slowest convergence rate in both set-ups. Adam and RMSprop algorithms have the best result. For $q = 50$ their curves are almost identical but when the noise increases Adam has much faster convergence rate in the beginning, hence it can potentially show better result. In addition, we can notice that with very intense noise ($q = 2$) Adagrad converges with bigger error in comparison with other algorithms and it performs almost the same as "classical" stochastic gradient descent. In order to reduce the error we can choose smaller learning rates. However, with smaller rate sometimes it was not able to converge at all.

Additionally, it is clear that noise affects not only the convergence rate but also the variance of estimates for all of the algorithms as the curves become less smooth.

3.2. Real data

Satellite images were used for the comparative analysis on real data. Figure 5 shows an example of the images taken in different weather conditions.

Figure 6 shows an example of mismatch Euclidean distance convergence of the algorithms for images shown in figure 5. For this experiment μ was set to 25 and the results were averaged by 100 realizations as them became more noisy in comparison with synthesized images, especially for the algorithms with adaptive learning rates.

One can easily notice that the results on real data are almost identical to the results on synthesized images. Again, Adam and RMSprop algorithms are the fastest in terms of convergence rate. However here we can see the algorithms with adaptive learning rates have much noisier curves that the others. It can be explained by the fact that when dealing with real images we have noisier gradient estimation, thus in these algorithms the learning rate estimation becomes less stable.

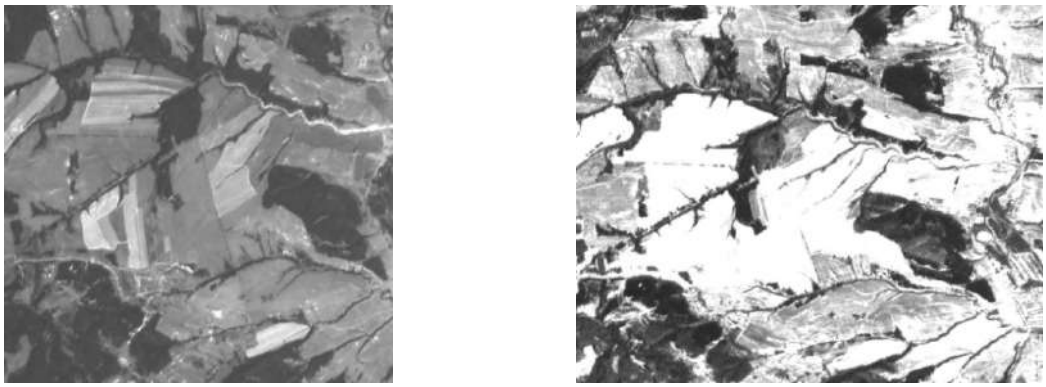


Figure 5. An example of real satellite images which were used for the comparative analysis.

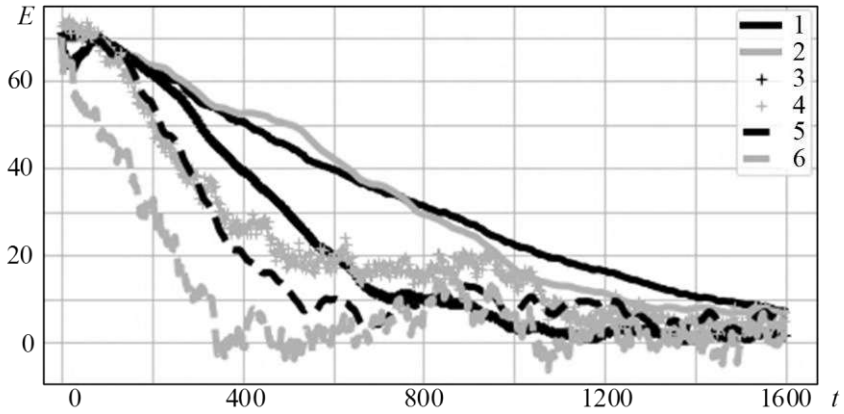


Figure 6. Mismatch Euclidean distance on the number of iterations for real images.

4. Conclusion

The comparative analysis of different optimization algorithms for solving image registration problem in spatial domain shows that in each case "classical" stochastic gradient descent shows the worst result in terms of the convergence rate of registration parameters' estimates. Momentum optimization algorithm just slightly outperforms it. Adagrad and Nesterov momentum algorithms show close results in terms of number of iterations before convergence but except for the situation with intense noise Adagrad in the beginning has much faster convergence rate and with some optimization (e.g. increasing or dropping learning rates after a number of iterations) it possibly can outperform Nesterov momentum The best results are provided by Adam and RMSprop optimizations Moreover, it is obvious that Adam algorithm is almost always preferable as it has less variance then RMSprop.

Furthermore, we can conclude that all of the algorithms have better convergence rate with bigger sample size. It is reasonable from theoretical point of view because the objective function gradient estimates become less noisy. Additionally, it is clear that noise affects not only the convergence rate but also the variance of estimates for all of the algorithms as the curves become less smooth.

Experiments on real satellite images show mostly identical results.

5. References

[1] Goshtasby A A 2012 *Image registration. Principles, tools and methods* (Springer London Dordrecht Heidelberg New York) p 441

[2] Tashlinskii A G and Voronov S V 2013 *The 11th Int. conf. Pattern Recognition and Image Analysis: New Information Technologies* **1** 326-329

[3] Papamakarios G 2014 *Comparison of Modern Stochastic Optimization Algorithms* (University of Edinburgh) p 13

[4] Tashlinskii A G, Safina G L amd Voronov S V 2012 Pseudogradient optimization of objective function in estimation of geometric interframe image deformations *Pattern Recognition and Image Analysis* **22** 386

[5] Goodfellow I, Bengio Y and Courville A 2016 *Deep Learning* (MIT Press) p 800
[6] Polyak B T 1964 Some methods of speeding up the convergence of iteration methods *Computational Mathematics and Mathematical Physics* **4** 1
[7] Nesterov Y A 1983 A method of solving a convex programming problem with convergence rate $O(1/k^2)$ *Soviet Mathematics Doklady* **27** 372
[8] Duchi J, Hazan E and Singer Y 2011 Adaptive subgradient methods for online learning and stochastic optimization *Journal of Machine Learning Research* **12** 2121
[9] University of Toronto csc321 course lecture 6 slides (Access mode: https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf) (18.11.2017)
[10] Kingma D and Ba J 2015 *Adam: A method for stochastic optimization* (Published as a conference paper at ICLR)
[11] Krasheninnikov V R 2003 *Foundations of Image Processing Theory* (Ulyanovsk: UlSTU) p 150
[12] Smelkina N A, Kosarev R N, Nikonorov A V, Bairikov I M, Ryabov K N, Avdeev A V and Kazanskiy N L 2017 Reconstruction of anatomical structures using statistical shape modeling *Computer Optics* **41(6)** 897-904 DOI: 10.18287/2412-6179-2017-41-6-897-904

Acknowledgments
This work was supported by the Russian Foundation for Basic Research, projects no. 16-41-732084 and 16-31-00468.

Wave skeletonization algorithm for raster images

M A Kudrina[1] and V S Mishenev[1]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086

Abstract. This article covers a wave skeletonization algorithm for raster images. It contains diagrams of algorithms for primary graph building and separating wave generation into wavelets. It is proposed to store the graph's structure in nested form to save memory and make it easier to modify. Deming orthogonal regression is proposed for primary graph optimization as part of successive approximations method.

1. Introduction

There is a number of tasks which require building raster image skeletons, e.g. handwritten text recognition, fingerprint comparison, processing medical or cartographical images, technical drawings [1-5], etc.

Various skeletonization algorithms exist, including Zhang-Suen algorithm [4], template method, wave algorithm [1, 6, 7], etc. The wave algorithm is covered in this article.

Essence of the wave algorithm for raster image skeleton calculation has already been described and illustrated in detail particularly in [1, 4, 6]. However, when implementing programs with this method, a number of questions, subtleties and nuances may arise, which are the subject of this article. The article contains diagrams of algorithms for primary graph building and separating wave generation into wavelets. Scientific novelty of this publication is that the graph describing object's skeleton is stored in a nested structure instead of connectivity and incident matrices which saves memory and makes it easier to change graph's structure, and also that Deming orthogonal regression is used for primary graph optimization as part of successive approximations method.

Image skeleton refers to a set of points equidistant from image borders. Several preparatory steps are usually taken before object's skeleton is built. These steps include image preprocessing (eliminating minor noise, etc.) and binarization. Binarization refers to the process of converting color and halftone images into dual-color, i.e. binary, images. Binarization is done using thresholding, point transformation, folding and border enhancement methods, extracting low-frequency and high-frequency image components, etc.

For our purpose, wave skeletization algorithm can be divided in two stages: primary graph building, including wave initiation and tracking covered distance, separation and attenuation points, followed by graph optimization where excessive information is dropped. Workflow for these stages is illustrated in figure 1.

At the first stage, spherical wave is initiated within an object. Wave generation is the wave front. Points that belong to certain wave generations are marked by different colors in the image. Central pixels of each odd generation included in the primary graph as its nodes are tracked.

However, primary graph usually contains a lot of excessive information, so redundant point are eliminated at the optimization stage.

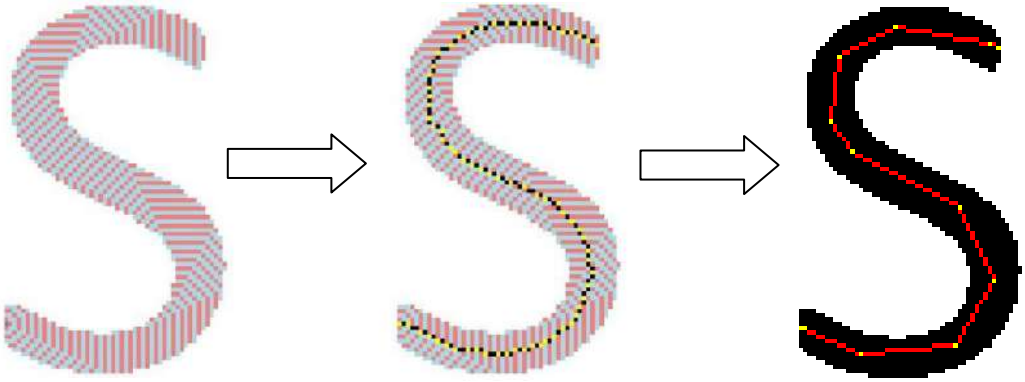


Figure 1. Key skeletonization stages - primary graph building and graph optimization.

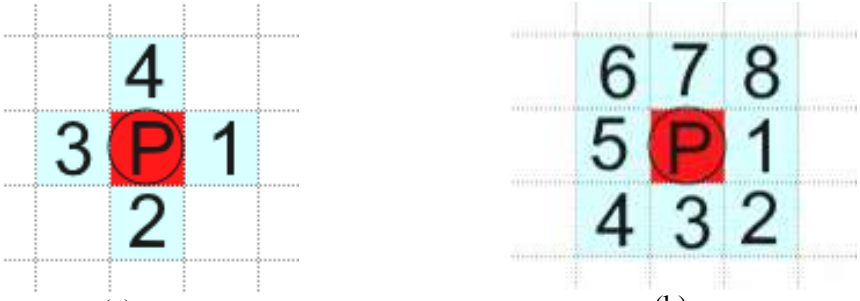
2. Primary graph building stage

Graph is built by tracking spherical wave propagation path over the image.

2.1. Spherical wave propagation

Spherical wave is initiated at any pixel inside the object.

Spherical wave generation is obtained by using 4- and 8-connected propagation alternatively. Connectivity refers to number of pixels adjacent to the current pixel. For 4-connected wave propagation, the pixels above, below, to the left and to the right of the current one are considered its neighbor pixels (see figure 2(a)). For 8-connected propagation, accordingly, all 8 pixels around the current one are considered its neighbor pixels (see figure 2(b)). In figure 2, the current pixel is designated as P, and its neighbors for 4-connected and 8-connected representations are numbered.



(a) (b)

Figure 2. Neighbor pixels and their winding order for 4-connected (a) and 8-connected (b) representations.

When a wave is initiated, the 4-connected propagation is applied first; then 8-connected propagation is applied to each pixel of the previous 4-connected generation, and so on alternately. This way, the wave propagates in octagonal shape. This approach is based on Huygens-Fresnel wave theory principle positing that each element of wave front can be considered as secondary disturbance center producing secondary spherical waves.

There are several special aspects to spherical waves:

- wave propagation becomes steady regardless of its initial point after at most 2N steps, where N is the line width in pixels (see figure 3);
- such wave is “able to” turn and effectively go round various obstacles. Small one- or two-pixel obstacles have little impact on wave propagation. However, it is recommended to eliminate such obstacles at the binary image generation stage for wave stability.

Zero generation consists of one initial pixel that is marked as first in the image. It is located at the center in figure 4.

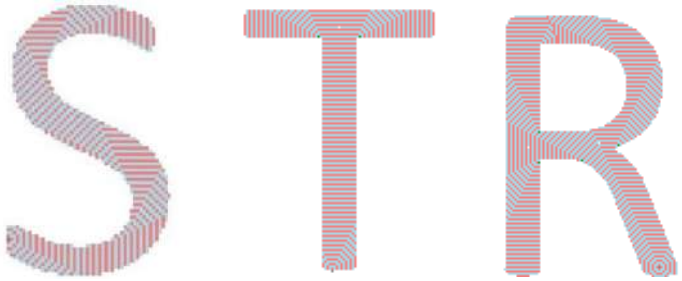


Figure 3. Spherical wave propagation.

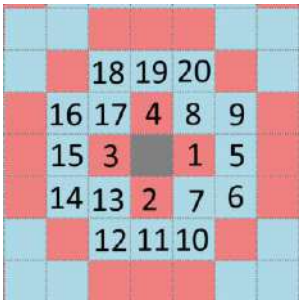


Figure 4. Numbering of first and second generation pixels.

First wave generation consists of unmarked neighbors of the initial pixel at 4-connected propagation. First generation pixels are marked red on the image. In figure 4, they have numbers 1, 2, 3, 4. Secondary generation is unmarked neighbors of the first generation pixels at 8-connected propagation (in figure 3, they have numbers 5-20), etc.

For each point with (i, j) coordinates, checking whether its neighbor pixels have been marked is performed in fixed order - clockwise, always starting from the right middle pixel. This way, the check is done in order shown in figure 2(a) at 4-connected propagation or in order shown in figure 2(b) at 8-connected propagation. You may notice from figure 4 that the winding order for the second generation pixels in the example differs from order of pixels in the image. This is due to fixed order of marking neighbor pixels; however, this simplifies the algorithm.

2.2. Graph's structure

Object's graph is stored in memory as a nested structure [8]. When this method of storing graph structure is used, each of its nodes is determined by a nest consisting of (x, y) coordinates of graph node and a list of addresses for transitions to next (child) nests (graph nodes). Advantages to this method are that it allows to easily modify graph's structure and does not require much RAM as opposed to connectivity or incident matrices.

An example of graph and corresponding nested structure are shown in figure 5. Nests contain graph node coordinates (x_i,y_i) and A_i – a reference to the list of references to neighbor nests (graph nodes).

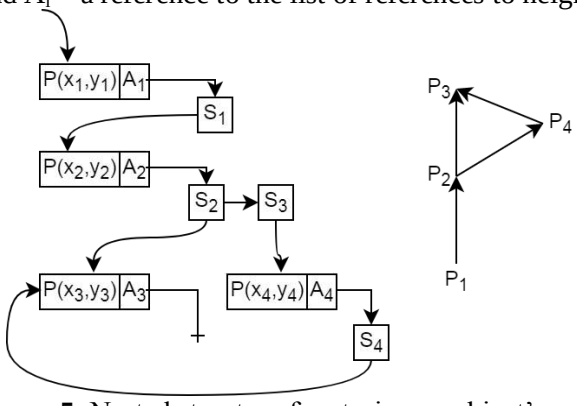


Figure 5. Nested structure for storing an object's graph.

2.3. Key primary graph building algorithms

A diagram of the primary graph building algorithm is shown in figure 6.

Figure's skeleton is built by tracking centers of each odd wave generation. Current wave generation is stored in memory as a list of pixel coordinates of this front. Let us designate the list of points of any odd generation as S_4, because its points are 4-connected neighbors of previous generation's points.

Accordingly, we shall designate lists of even generations as S_8, because they are 8-connected neighbors of previous generation's points.

Each generation of a wave is located between extreme pixels that have coordinates (x_1, y_1) and (x_2, y_2). Extreme pixels of a wave are pixels that have at most one neighbor in their generation. For the

figure skeleton point we shall take the midpoint of segment that connects extreme points of wave generation; thus, the point's coordinates are $((x_1+ x_2)/2; (y_1+ y_2)/2)$.

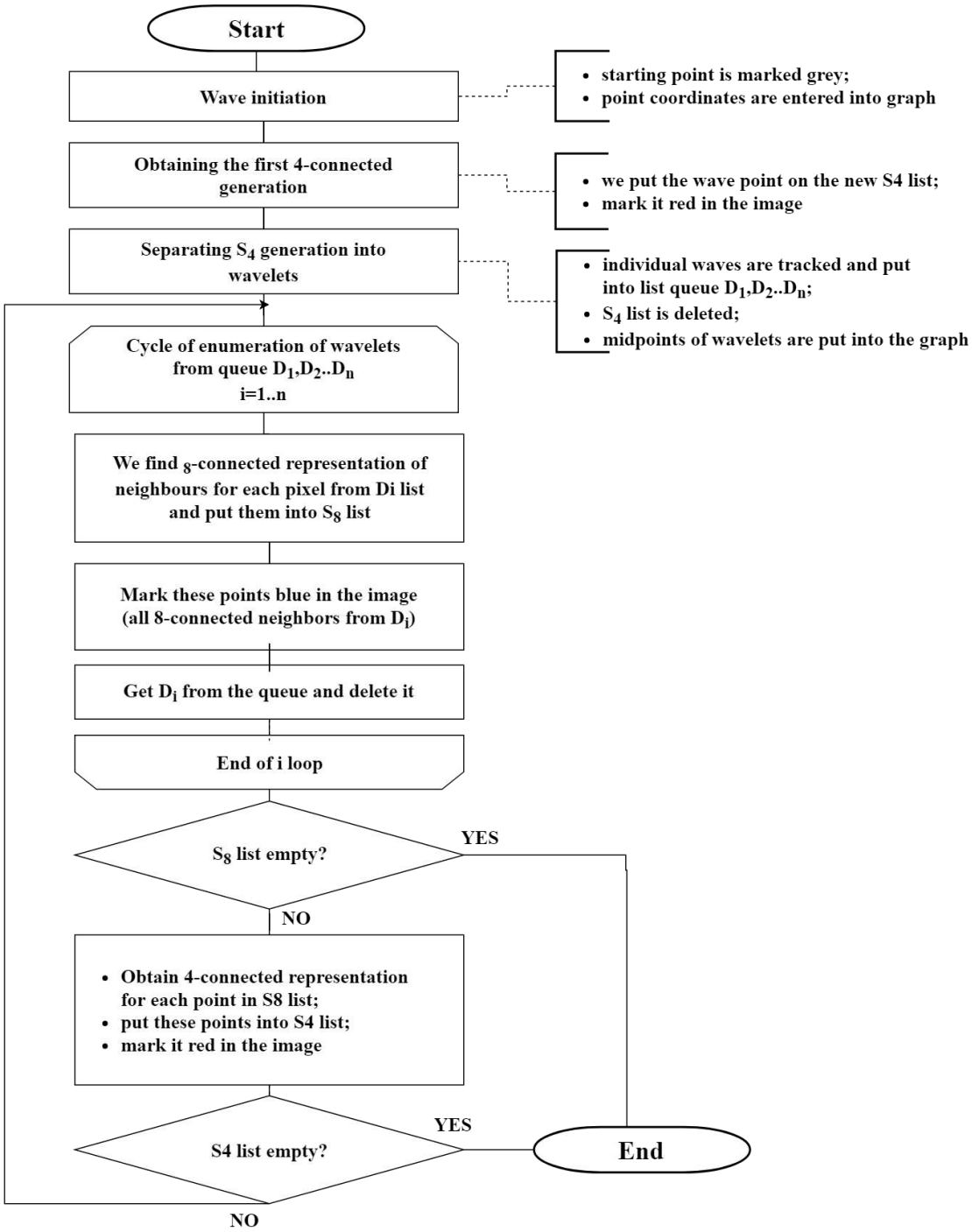


Figure 6. Diagram of the primary graph building algorithm.

Besides, wave generation may include several separate fronts. This happens when wave generation is separated into wavelets.

This situation can be identified using the number of extreme points in S_4 list of generation points. If there are more than two extreme points, separation has taken place and several wavelets are in the list (an example of a wave divided into three fronts is shown in figure 7). In this case, separate lists of wavelets (D_1, D_2,.., D_n) are queued. Separation of S_4 wave generation list into D_1, D_2,.., $D_n$wavelets is shown in figure 8.

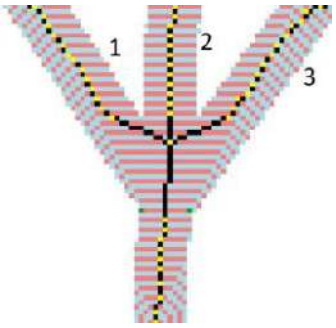


Figure 7. Wave generation separated into wavelets.

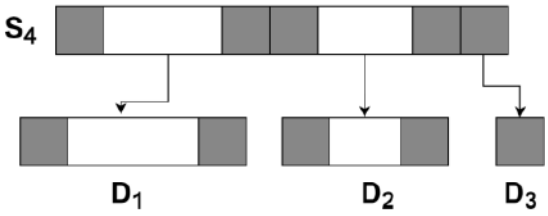


Figure 8. List of wave generation points separated into three fronts.

Algorithm diagram for separating wave generation (S_4) into wavelets (D_1, D_2,.., D_n) is shown in figure 9.

Before separation, wave width steadily grows and the number of wave generation extreme points increases. By tracking this, it is possible to identify junction point of image skeleton segments. Point of junction of separate branches is calculated as barycenter of the polygon formed by extreme points of the parent wave and wavelets [5, 6, 9].

Whenever a wave turns or an image element becomes thicker, the number of extreme points remains constant, thus the wave width increase can be neglected.

A wave attenuates when encountering another wave. In this case, the graph closes.

3. Graph optimization stage

The first method used to optimize the graph (i.e. eliminate redundant information) is successive approximations method. It works as follows: a straight line is drawn through the first two nodes of the graph, and a check of deviation of each successive point from that line is done. If the deviation is within ε limit (which is specified comparable to thickness of lines in the image), the point belongs to this line; otherwise, it forms the next line.

In successive approximations method, equation coefficients of the line $Ax + By + C = 0$ going through points (x_1, y_1) and (x_2, y_2) are calculated using the following formulas:

$$A = y_1 - y_2, B = x_2 - x_1, C = x_1 y_2 - y_1 x_2.$$

Then, for each successive point (x_i, y_i) of graph, its distance from the line $Ax + By + C = 0$ is calculated using the formula $r = |Ax_i + By_i + C| \cdot (A^2 + B^2)^{-1/2}$. If $r <$ ε, the point belongs to the current line.

The obtained set of points belonging to the same line is approximated using orthogonal regression. Orthogonal regression differs from simple linear regression in that it makes allowances for observation errors along both x and y directions (see figure 10).

If points are approximated by the line that is determined by equation $y = \beta_0 + \beta_1 x$, regression coefficients are calculated using the following formulas [10]:

$$\beta_1 = \left(s_{yy} - s_{xx} + ((s_{yy} - s_{xx})^2 + 4{s_{xy}}^2)^{1/2}\right) \cdot \left(2s_{xy}\right)^{-1},$$
$$\beta_0 = \bar{y} - \beta_1 \bar{x},$$

where

$$\bar{x} = \frac{1}{n}\sum x_i,\ \bar{y} = \frac{1}{n}\sum y_i,$$
$$s_{xx} = \frac{1}{n-1}\sum (x_i - \bar{x})^2,$$
$$s_{yy} = \frac{1}{n-1}\sum (y_i - \bar{y})^2,$$
$$s_{xy} = \frac{1}{n-1}\sum (x_i - \bar{x})\,(y_i - \bar{y}).$$

Since the equation y= $\beta_0 + \beta_1 x$ does not determine vertical lines, let us put it to the form $Ax + By + C = 0$. For that purpose, we shall calculate the following coefficients A, B, C: $A = \beta_1, B = -1, C = \beta_0$.

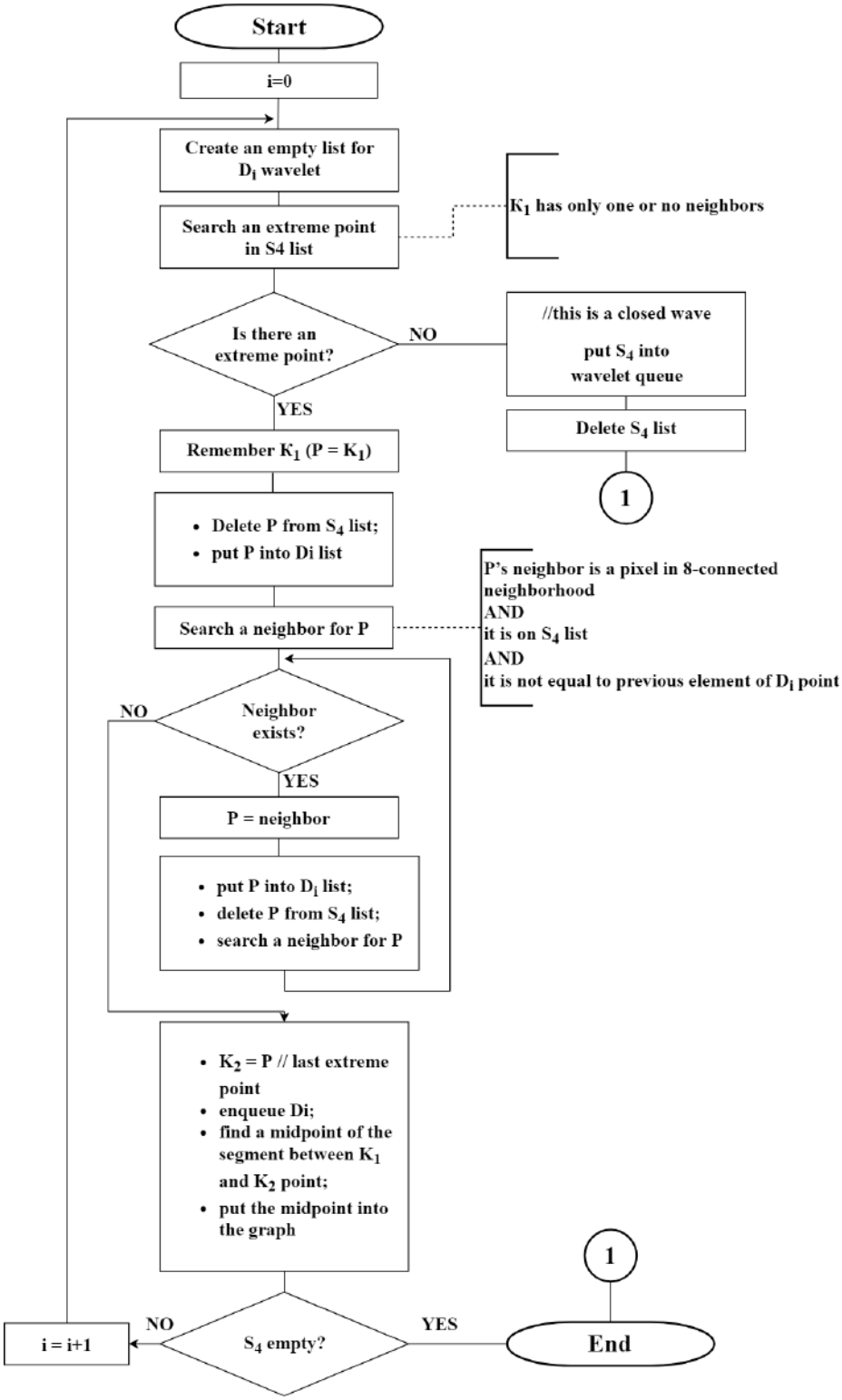


Figure 9. Algorithm diagram for separating S_4 generation into wavelets D_1, D_2,.., D_n.

If $s_{xx} = 0$, the line is parallel to Y axis, and the equation looks like $x = \overline{x}$. In this case, the line has the following coefficients: $A = -1, B = 0, C = \overline{x}$.

If, on the other hand, $s_{yy} = 0$, the line is parallel to X axis, and the equation looks like $y = \overline{y}$. In this case, the line has the following coefficients: $A = 0, B = -1, C = \overline{y}$.

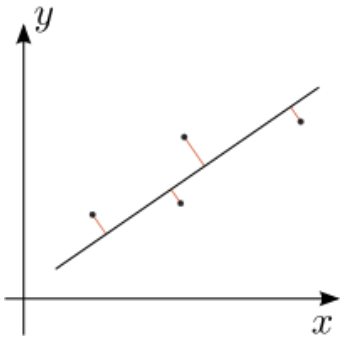


Figure 10. Approximation of graph node set using Deming orthogonal regression.

After the coefficients of approximating line have been found, extreme points are adjusted, i.e. projected onto approximating line, whereas other intermediary points are eliminated from the graph.

Coordinates (x_{proec}, y_{proec}) of projection of extreme point (x_0, y_0) onto line $Ax + By + C = 0$ obtained with orthogonal regression method are calculated using the following formulas:

$$x_{proec} = (B(Bx_0 - Ay_0) - AC)/(A^2 + B^2),$$
$$y_{proec} = (A(Ay_0 - Bx_0) - BC)/(A^2 + B^2).$$

Because, in general, the graph may be closed, all processed nodes are marked so the program does not get caught in an endless loop.

4. Results and conclusions

In this article, diagrams of algorithms for primary graph building and separating wave generation into wavelets are proposed. A suggestion is made to store the structure of the graph representing object's skeleton in nested form instead of connectivity or incident matrices to save memory and make the graph's structure easier to modify. Scientific novelty of this publication is also that Deming orthogonal regression is used in successive approximations method for primary graph optimization.

A wave skeletonization program has been developed using *C#* language. The program proactively converts color raster image into binary image using thresholding where the threshold is an average between maximum and minimum brightness. Then the primary graph is built and optimized.

Wave skeletization algorithm does not require much memory because memory is only allocated to store graph's skeleton and lists of pixels of the current and next wave generations. Spherical wave marking is done in the source image which is easily restorable. The algorithm is easily parallelized when several starting points of wave initiation are used. For wave initiation, it is preferable to choose pixels on lower or upper borders of symbol images; this makes skeletons smoother.

Today, wave skeletization is used successfully to build skeletons of symbols (letters, digits, etc.) for further processing as part of recognition process. However, this algorithm is not as effective for building fingerprint skeleton, because finger papillary picture has more complex structure, thus wave propagation and further analysis have poor accuracy [1].

References

[1] Gudkov V Yu and Klyuev D 2015 Skeletonization of Binary Images and Finding of Singular Points for Fingerprint Recognition *Bulletin of the South Ural State University. Ser. Computer Technologies, Automatic Control, Radio Electronics* **15(3)** 11-17

[2] Khaustov P A 2017 Algorithms for handwritten character recognition based on constructing structural models *Computer Optics* **41(1)** 67-78 DOI:10.18287/2412-6179-2017-41-1-67-78

[3] Glaister P 2001 Least squares revisited *The Mathematical Gazette* **85** 104-107

[4] Zhang T Y and Suen C Y 1984 A fast parallel algorithm for thinning digital patterns *Comm. of ACM* **27(3)** 236-239 DOI:10.1145/357994.358023

[5] Saeed K, Tabędzki M, Rybnik M and Adamski M 2010 K3M: A universal algorithm for image skeletonization and a review of thinning techniques *International Journal of Applied Mathematics and Computer Science* **20(2)** 317-335 DOI:10.2478/v10006-010-0024-4

[6] Klubkov I 2001 Primeneniye volnovogo algoritma dlya nakhozhdeniya skeleta rastrovogo izobrazheniya *Vestnik DGTU* **1(7)** 9
[7] Fisher R, Perkins S, Walker A and Wolfart E 2000 Skeletonization *Medial Axis Transform* (Access mode: http://www.dai.ed.ac.uk/HIPR2/skeleton.htm) (31 May 2018)
[8] Poulovassilis A and Levene M 1994 A Nested-Graph Model for the Representation and Manipulation of Complex Objects *ACM Transactions on Information Systems* **12(1)** 35-68
[9] Mishenev V S and Kudrina M A 2017 Building Raster Image Skeleton Using Wave Algorithm *Trudy Mezhdunarodnogo simpoziuma Nadezhnost' i kachestvo* **1** 378-381 (in Russian)
[10] Haeckel R, Wosniok W and Klauke R 2013 Comparison of ordinary linear regression, orthogonal regression, standardized principal component analysis, Deming and Passing-Bablok approach for method validation in laboratory medicine *Laboratoriumsmedizin* **37(3)** 147-163 DOI:10.1515/labmed-2013-0003

Visible Structures Highlighting Model Analysis Aimed at Object Image Detection Problem

I R Saifudinov[1], V V Mokshin[1], P I Tutubalin[1], L M Sharnin[1] and D G Hohlov[1]

[1]Kazan National Research Technical University named after A. N. Tupolev - KAI, Karl Marx str. 10, Kazan, Russia, 420111

Abstract. The research considers an approach to solving the problem of reducing the data processed in mobile platforms oriented video-analytic systems. Different models of human visual attention has been analyzed and also classified according to the image segmentation. The results obtained are presented in the form of the method used for isolating the borders of the most significant object in the image. They are based on the optimization the length and curvature values of the object borders. This approach allows to filter significant structures in the image and process them with the help of image segmentation techniques. The method was evaluated according to usage the accuracy criterion along with other methods in delineation of boundaries: threshold value, morphological processing and watershed. Software and hardware system for registering dump trucks has been improved. It gives the possibility to automatize the process of registration dump trucks, training them during the construction of roads.

1. Introduction

Automated video surveillance systems (video analytics) used in many spheres of human activity, likewise state institutions and manufacturing enterprises are extremely popular nowadays. The systems usage allows to create options for events automatic alarm, efficiency in employees labour improvement by means of direct control of the performing activity.Herewith, because of the large number of tasks solved in video analytics, as well as their multi-criteria, autonomy tendency, and indistinct nature of tasks, there is a certain necessity to find the effective approach for highlighting image visible structures that allows to work on the basis of mobile platforms. The attempt for analysis of various tasks has been made in the research. [1], [3], [6-8].

The method of highlighting of visible structures in an image based on visibility measuring of length and curvature of the curve that is similar to the concept suggested by Lowe [4] has been taken into consideration. The research deals with the measure of visibility according to the criterion of false positive detection in comparison to the method of delimiting contour of object [2]. In addition to the evaluation of the method, it is compared with the segmentation and accuracy of other methods of optimization of significant information.

2. Visibility network construction

Orientation elements are the basic computational elements of the network [9]. Each element p_i is connected with a processor that normally can perform certain calculations based on the conditions and those with k denotation performed by the neighbour processor. This defines a single network containing kn^2 processing blocks with the local communication. In the current implementation, k is equal to 48, which provides a reasonable angular resolution. Let's appeal to the associated orientation

sequence of the elements $p_i, \dots, p_{i+N}$, , where each element is a linear segment or interval, like length of a curve *N*. (curves can be continuous or with any number of intervals). The optimization task is formulated to maximize the value of Φ_N above overall length of a curve *N*, starting with p_i.

$$\max_{(p_{i+1},\dots,p_{i+N})\in\delta^N(p_i)} \Phi_N(p_i, \dots, p_{i+N})$$

where $\delta^N(p_i)$ is a number of all possible length of a curve *N*, starting with p_i.

For a certain of measures class $\Phi(\cdot)$, the calculation of Φ_N can be obtained by simple local computations made repeatedly. To illustrate, let's take the first curves three elements long only. In this case:

$$\max_{(p_{i+1},p_{i+2})\in\delta^2(p_i)} \Phi_2(p_i, p_{i+1}, p_{i+2})$$

where p_i is defined by p_{i+1} (one of p_i's k of neighbours) and p_{i+2} (p_{i+1} neighbour) for a given element so that the rate $\Phi_2(p_i, p_{i+1}, p_{i+2})$ will be in its maximum. Simple approach (brute-force method) in the analysis of k^2 value in different curves will be required anew. Suppose, however, that Φ_2 corresponds to the condition of:

$$\max_{\delta^2(p_i)} \Phi_2(p_i, p_{i+1}, p_{i+2}) = \max_{p_{i+1}} \Phi_1(p_i, \max_{p_{i+2}} \Phi_1(p_{i+1}, p_{i+2}))$$

In this case, the maximization of the rate Φ_2 can be achieved by application of Φ_1 used repeatedly over shorter curves. The general approach can be formulated the same way, e.i:

$$\max_{\delta^N(p_i)} \Phi_N(p_i, \dots, p_{i+N}) = \max_{p_{i+1}\in\delta(p_i)} \Phi_1(p_i, \max_{\delta^{N-1}(p_{i+1})} \Phi_{N-1}(p_{i+1}, \dots, p_{i+N})) \quad (1)$$

where $\delta(p_i)$ is equal to $\delta^1(p_i)$. Thus, the searching area required for each length of a curve N is being reduced, from p_i to kN, instead of k^N which is essential for a brute-force method approach. The concept (1) is related to the optimality principle, underlining for all multistage decision-making processes. This is a special case in dynamic programming in particular. It refers to the family of functions that follows concept (1) of extensible functions.

There are two factors that are essential for visibility measure. The first one is related to the length of a curve, and the second factor is related to its shape. The length of a curve is determined by the number of its elements that have a factual curve (rather than an interval) passing through these elements. They are called active elements. Whereas elements that are associated with intervals are referred to as virtual elements where local visibility σ_i corresponds to p_i. If p_i is the active element, then σ_i has positive value, which is equal to 1 and 0 for the virtual element σ_i. A measure associated with the length of a curve $p_i, \dots, p_{i+N}$ is determined by the equation:

$$\sum_{j=i}^{i+N} \sigma_j \quad (2)$$

The measure rate above(2) presents the sum of the local values of the visibility of active elements along the curve.

Then, the attenuation function associated with the curve $p_i, \dots, p_j$ is defined as follows:

$$\rho_{i,j} = \prod_{k=i+1}^{j} \rho_k$$

where $\rho_{i,i} = 1$. The measure in (2) is modified by the attenuation coefficients is:

$$\sum_{j=i}^{i+N} \rho_{i,j}\sigma_j \quad (3)$$

The measure rate in equation (3) is weighted contribution of the local visibility values σ_j along a curve, that are in reverse dependency to a number of virtual elements along $p_i, \dots, p_j$. To measure the shape of a curve, measure that in reverse dependency to the total curvature of a curve is used. The total curvature of γ is defined as $\int_\gamma \left(\frac{d\theta}{ds}\right)^2 ds$, where $\theta(s)$ is a slope along a curve and $\frac{d\theta}{ds}$ at the point *P* is known as the local curvature at this point (the reciprocal value of the radius of curvature *R*). It is necessary to use the total curvature to obtain a measure that is limited and in the position of reverse dependency to the total curvature. The next measure is relevant to the following:

$$\exp^{-\int_\gamma \left(\frac{d\theta}{ds}\right)^2 ds} \quad (4)$$

In order to obtain a discrete approximation to the measure in (4), we denote a_k to indicate orientation difference between the k-th element and its successor and Δs as length of the orientation element. A discrete approximation to full curvature of a measure along $p_i, \dots p_j$, will be:

$$C_{i,j} = \prod_{k=i}^{j-1} f_{k,k+1}$$

where

$$f_{k,k+1} = \exp^{-\frac{2a_k \operatorname{tg}\frac{a_k}{2}}{\Delta s}} \quad (5)$$

$C_{i,j}$ is the weight of each value of local visibility σ_j along a curve. A measure that shows a high rate for long curves with low overall curvature is now defined as:

$$\sum_{j=1}^{i+N} C_{i,j}\, \rho_{i,j} \sigma_j \quad (6)$$

The measure in equation (6) is weighted contribution of the local visibility values σ_j along a curve. The curves that will receive a high measure on (6) are long curves, more straight with the least number of intervals.

3. Analysis of the visibility network: discussion and results

To analyse visibility network we compared the percentage of false positive detections. Test samples located around the perimeter of circle at equal intervals consist of short, oriented segments in the field of segments with random position and orientation as it shown in figure 1.

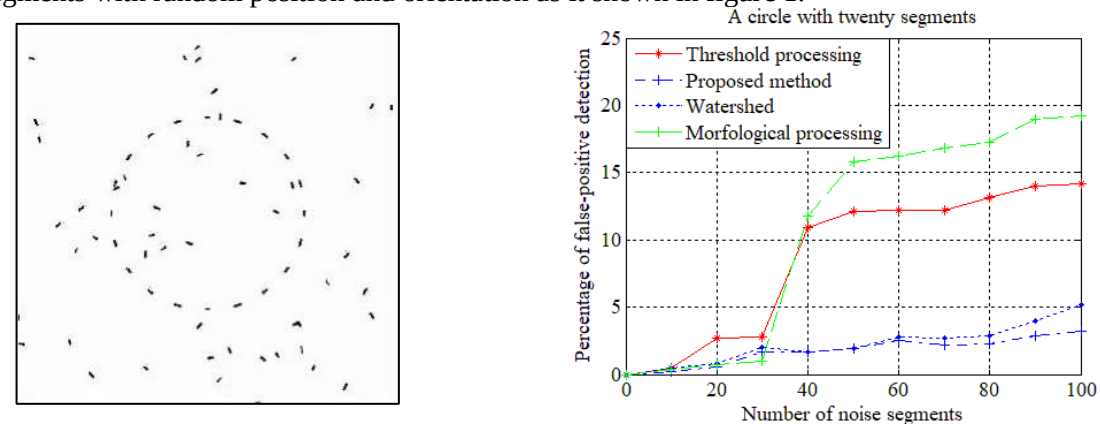


Figure 1. A circle with twenty segments.

The methods that are used for the calculation visibility forms and noise segments are: the threshold value [10], the watershed [11,13], the morphology [12], and the proposed approach. Segments were sorted in ascending order according to their visibility of the most (ϕ_1) and less (ϕ_n) noticeable segments. For given m form segments, false-positive are defined as noise segments, which are assigned a visibility greater than ϕ_{m+1}. A false positive estimate for each method was calculated for samples consisting of different numbers of shape and noise segments rate. A false positive rate for each combination (for example, 20 shape segments and 70 noise segments) was estimated by averaging false positive value by means of more than ten attempts with different noise samples. The picture in the right side of Figure 2 is the graph of false positive rate percentage for a circle with twenty segments.

Each method can be accomplished well enough (less than 10% of false positives) at a low noise level (40 noise segments or less). The results of methods applied begin to diverge at higher noise levels. It should be noted that threshold processing is superior to morphological processing while watershed is relative to the approach suggested, although the latter is more expensive to calculate.

And, finally, at lower signal-to-noise ratios watershed and the proposed approach have significantly lower false-positive estimates. The next comparison was identical to the first one, but shape segments are formed by an unlimited sinusoid (Figure 2).

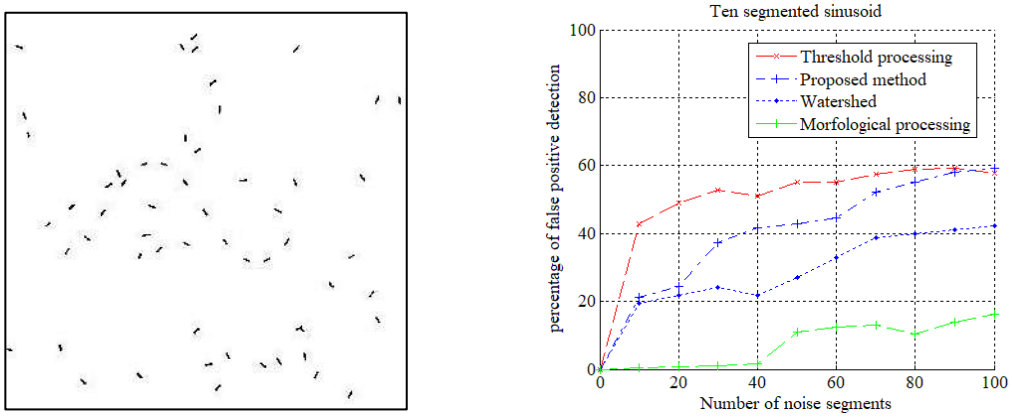


Figure 2. A circle with a ten segment sinusoid.

In the right half of Figure 2 is a graph of the percentage of false positives relative to the number of noise segments for a ten segment sinusoidal curve. Comparatively low indicators of the proposed method compared with other methods can be attributed to its apparent dependence on closure. However, it still outperforms the threshold processing for higher signal-to-noise ratios and has an error rate comparable to threshold processing (i.e., within 5%) at lower signal-to-noise ratios.

In the third comparison a field consisting of correlated noises (i.e., a dipole) was used (Figure 3). The dipole consists of two collinear segments separated by an interval equal to the distance between neighboring segments of the circle. Since the two segments forming the dipole are collinear, the degree of closeness between the segments forming the dipole is greater than between the adjacent segments of the circle. Therefore, it is impossible to distinguish noise segments from the shape segments using only a local measurement. From the graph it can be seen that the methods of threshold processing and morphological processing have almost a 100% false positive level, whereas the watershed and the proposed method are much better able to cope with this task.

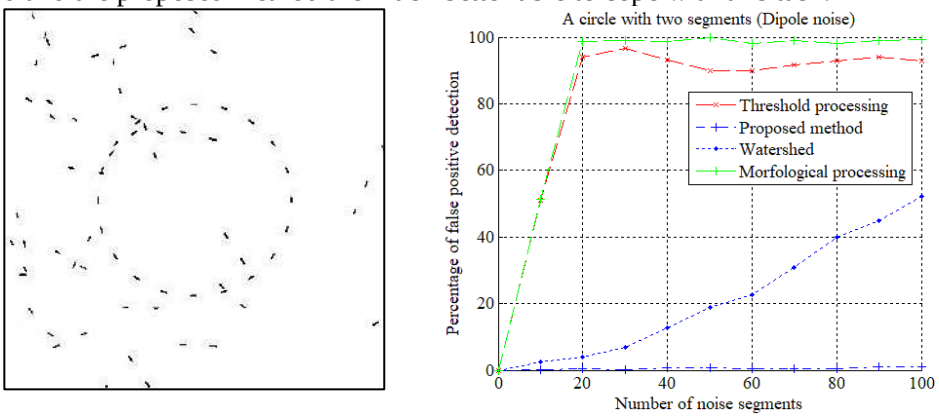


Figure 3. A circle with twenty segments. Dipole noise.

In the fourth comparison (Figure 4), ten segments are used. This is a complex picture, because the sampling frequency is so small, so that there is only one segment at 36 degrees of the circle. Most of methods do not work well even at relatively high signal-to-noise ratios. For noise level 80, threshold processing and the morphological processing method are performed at 90% false positive level. The watershed method is performed a bit better, with a false-positive level of 70%. In contrast, the false-positive level for the proposed method is less than 5%.

Let's consider practical use examples of the specified methods in the decision of the object segmentation problem on the image in the video-analytical system focused on mobile platforms for the

account number of transportations dump trucks. The image segmentation results are shown below. Segmentation for threshold processing is shown in Figure 5.

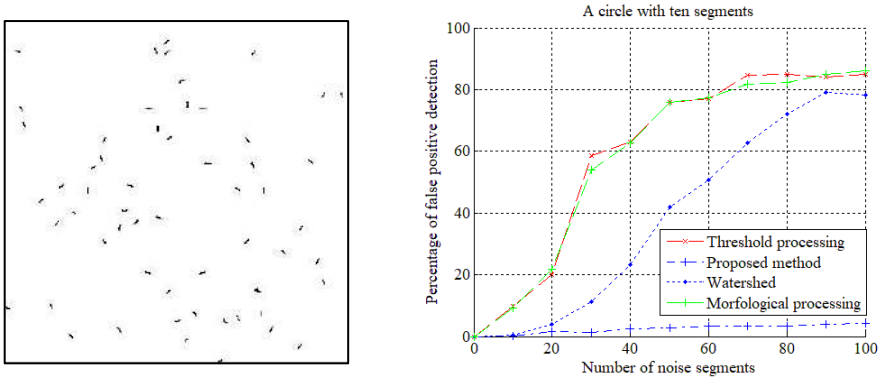


Figure 4. A circle with ten segments.

Figure 5. Segmentation. Threshold value.

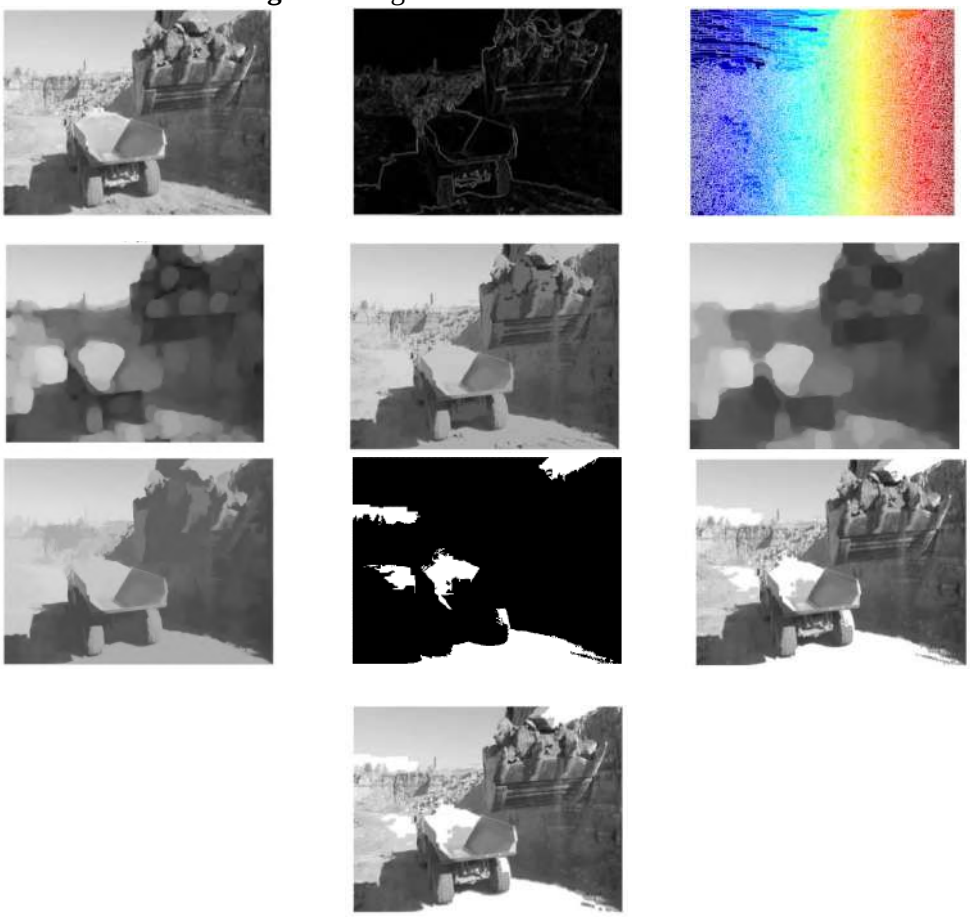


Figure 6. Stages of watershed segmentation. From left to right, from top to bottom.

Segmentation for the watershed is shown in Figure 6. The results of morphological processing segmentation are shown in Figure 7.

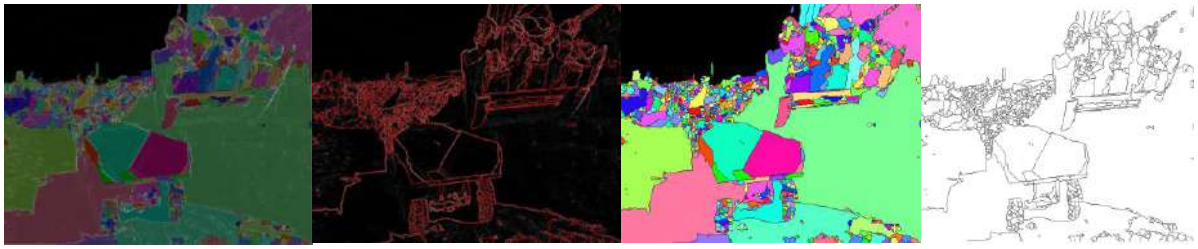


Figure 7. Stages of morphological processing segmentation. From left to right.

The results of the proposed method segmentation for isolating significant structures are shown in Figure 8.

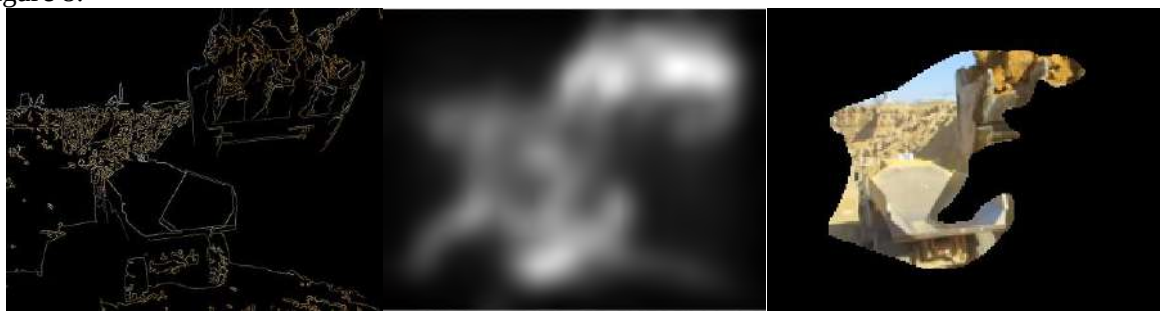
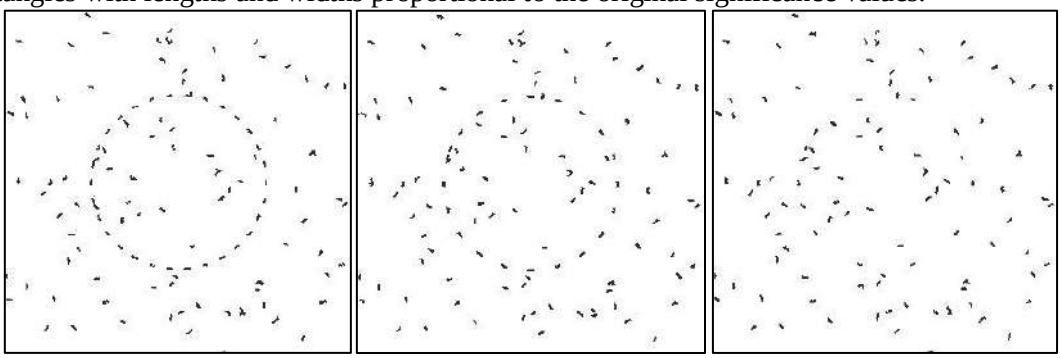


Figure 8. Left. Preprocessing an image using the Sobel operator. The middle. Map of the importance of the image. On right. The selected area of the image.

The proposed method of identifying significant structures was compared with other measures that calculate the optimal structure: contrast discrimination, a combined approach to optimization [18]; quantitative indicators of changes based on the organization of functions: eigenvalues and eigenvectors [19] and stochastic areas of completion: a neural model of illusory shape of the contour and significance [20-23]. Let us consider examples of practical use of these methods in solving the problem of segmentation of important structures in the image.

To get an idea of the strengths and weaknesses of boundary selection measures, it is useful to apply them to a simple test scheme consisting of edges from a circle (thirty, twenty, or ten evenly spaced samples) against a background of one hundred edges of random position and orientation. Three test patterns are shown in figure 9. To visualize the values assigned to each edge, the edges are displayed as rectangles with lengths and widths proportional to the original significance values.

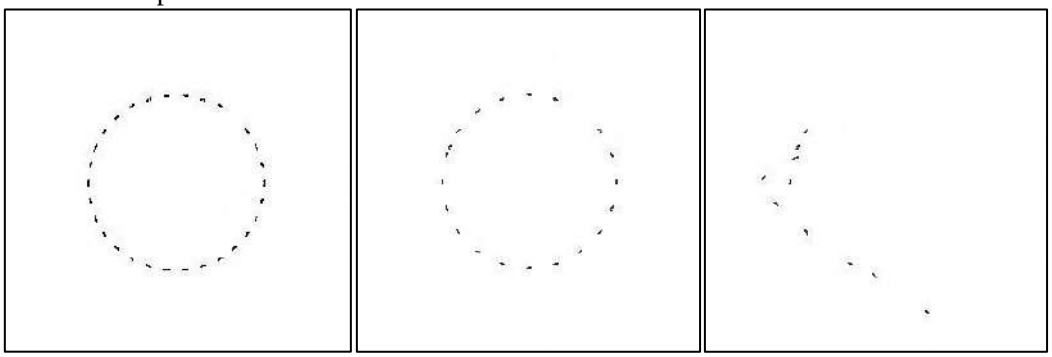


a) b) c)

Figure 9. (a) thirty edge circle against the background of a hundred edges of noise. (b) twenty edge circle. (C) ten edge circle.

First, let us consider the results of measuring the method of contrast discrimination. It is important to note that the previously described modified optimization problem is solved. That is, $y^T Ay$ is optimized for all vectors, $y \in \{0,1\}^n$ and $|y| = m$ where n - is the total number of edges and m is the number of edges. Since the method is given the number of edges, it has a significant advantage over

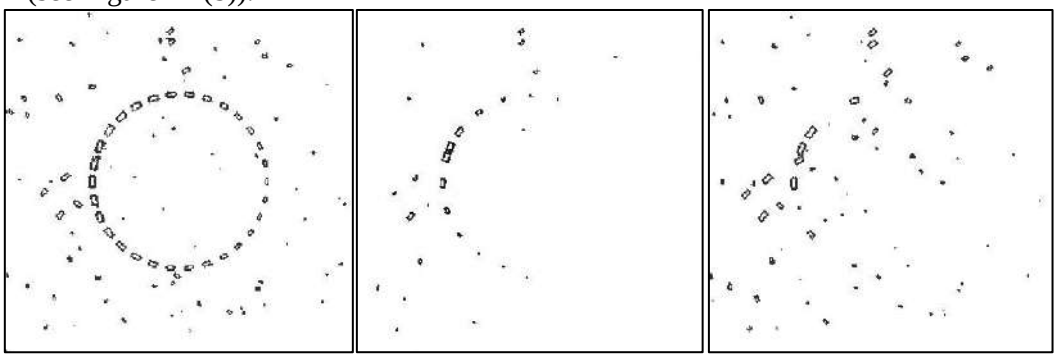
other measures, and these results should be interpreted accordingly. Segmentation for thirty and twenty edge circles is shown in Fig.10 (a) and (b). With the exception of excluding one edge of the circle in the 1 hour orientation, and including a false 10-hour edge, the method calculates perfect segmentation. However, the results are in ten regional districts (see figure 10 (c)) show that the method could not separate the circle from the lantern.



a) b) c)

Figure 10. Edges of edges, (a) computed by contrast discrimination for thirty edge circle, (b) for twenty edge circle, and (C) for ten edge circle.

To better visualize the large range of significance values calculated by the quantitative change measure method, the length and width of the rectangles are drawn proportionally to the $\log(1.0 + 10^6 \cdot x_i)$, where x_i - is the significance of the edge i. In Fig. 11 (a) results for the thirty edge circle calculated using the quantitative change method are shown. In General, the edges of the circle are set to a larger value than the edges of the background. However, it is observed that the significance values in the upper left part of the circle are much larger than the values in the lower right corner. If the eigenvector with the highest positive real eigenvalue is interpreted as a limited distribution of random walks between edges, it is seen that this distribution is dominated by random walks (with reversals in the direction) through the parasitic edge at the 10 o'clock position. Because the measure does not provide tangential continuity or closure, the effect of one inconveniently positioned edge can be profound. For Fig.11 (b) and (c) the asymmetry becomes very pronounced as the circle sample becomes less frequent. The consequence of this is that the measure has failed to isolate the circle from its base, even in the case where a very simple method such as stochastic completion areas has a small problem (see Figure 12 (b)).

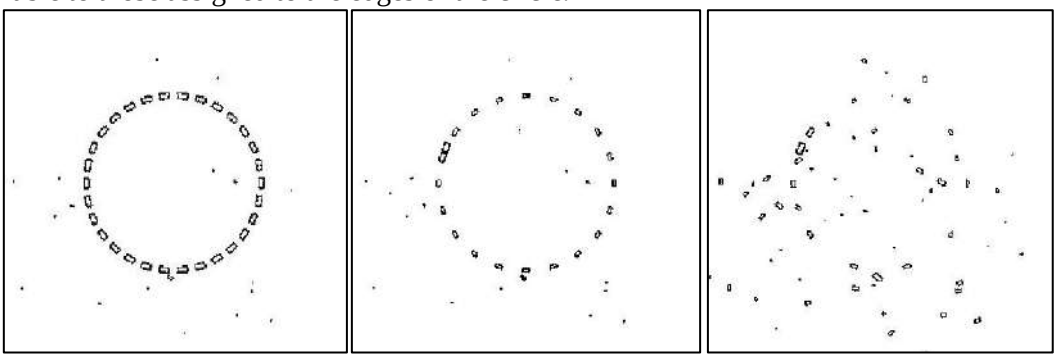


a) b) c)

Figure 11. The values of significance, (a) calculated by the quantitative change in the indicators for the thirty edge circle, (b) for the twenty edge circle, and (C) for the ten edge circle.

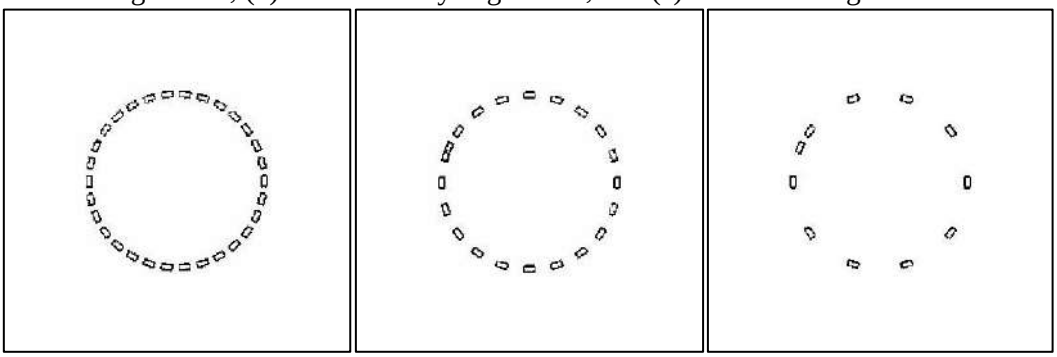
For figure 12 (a) and (b) significance maps for thirty and twenty edge circles calculated by the stochastic completion regions are shown. With the exception of the false 10-hour edge, all edges of the background are assigned gaps with a small value. The contrast between the values calculated by this method and the estimated quantitative indicators of changes is quite impressive. The difference in the power of discrimination is due solely to the use of directivity and the multiplication of vector x by

vector x in this method. This multiplication imposes a restriction that the edge must form a bridge between the other two. The result is intervals with a larger range of values than those with only x or x components . However, this limitation is not enough to distinguish the edges of the ten edge circles from the background borders. A significance map calculated as the stochastic completion regions for the ten edge circle is shown in figure 12 (c). A significant number of background edges have gaps comparable to those assigned to the edges of the circle.



a) b) c)

Figure 12. The significance values calculated by the stochastic completion regions, (a) for the thirty edge circle, (b) for the twenty edge circle, and (c) for the ten edge circle.



a) b) c)

Figure 13. The values of significance calculated by the proposed approach, (a) for the thirty edge circle, (b) for the twenty edge circle, and (c) for the ten edge circle.

Finally, figure 13 (a-c) shows significance maps calculated using the proposed approach. With the exception of the false 10-hour edge, all edges of the background are assigned intervals with a small value. This is true even for ten edge circles, which no other measure could segment from the background.

4. Conclusions

In addition to comparing false positive detection, Table 1 shows the time required to run the algorithms on the Android 5.0 operating system. With a resolution of 640x480 pixels. It can be seen from the table that the proposed method for identifying significant structures shows the fastest and least resource-consuming results. Thus, the proposed model for significant structures identifying is an effective tool for primary segmentation of images in video-analytical systems oriented at mobile platforms. As results have shown, the method of isolating notable structures showed the most accurate results. The method of selecting notable structures is based on the measure of visibility calculation that includes applies the principle of dynamic programming [5, 15]. Also it would be used queueing system to analyse such kind models [16-17]. Two internal properties of the network of significance, such as extensibility and geometric convergence, allow us to optimize the measure of significance and effectively restore optimal curves (in polynomial time).

Table 1. Comparison of the time spent on segmentation and resource intensity.

Method	Calculation time (seconds)	RAM Consumption (MB)
Threshold processing	1.36	70
Watershed	2.03	105
Morphological treatment	1.78	85
Allocation of significant structures	1.09	45
Contrast discrimination	1.66	77
Stochastic completion areas	1.83	98
Quantitative indicators of changes	1.88	95

At the same time, they restrict the range of possible functions that can be used as a measure of significance, so the method has some limitations with scale invariance, fusion of curves, and grouping in the presence of compounds. In addition, overcoming overlapping problems will require asymptotically increasing the complexity of the method, since the discretization is closely related with using dynamic programming to effectively optimize the chosen measure.

5. References

[1] Moksin V V, Saifudinov I R, Kirpichnikov A P and Sharnin L M 2016 Vehicle recognition based on heuristic data and machine learning *Bull. Kazan Techn. Univ.* **19(5)** 130-137 (in Russian)

[2] Saifudinov I R, Mokshin V V and Kirpichnikov A P 2017 Grouping contours of objects of structural images based on the network of element visibility *Bull. Kazan Techn. Univ.* **20(9)** 120-123 (in Russian)

[3] Treisman A 1982 Perceptual Grouping and Attention in Visual Search for Features and for Objects *Journal of Experimental Psychology: Human Perception and Performance* **8(2)** 194-214

[4] Lowe D G 1985 *Perceptual Organization and Visual Recognition* (Kluwer Academic Publishers, Boston)

[5] Kormen T, Leiserson Ch, Rivest R and Shtaine K 2005 Dynamic Programming *Algorithms: construction and analysis (Introduction to Algorithms)* (Moscow: Williams) p 1296 (in Russia)

[6] Gorilik A L and Skripkin V A 1989 *Recognition methods* (Higher School Publishing) p 232 (in Russia)

[7] Vapnik V N and Chervonenkis A I 1974 *Theory of pattern recognition* (Moscow: Science) p 416 (in Russia)

[8] Nesteruk V F and Sokolova V A 1980 Questions of the theory of perception of subject images and a quantitative assessment of their contrast *Optoelectronic Industry* **5** 11-13 (in Russia)

[9] Shashua A and Ullman S 1988 Structural saliency: The detection of globally salient structures using a locally connected network *2nd Intl. Conf. on Computer Vision (ICCV '88), Clearwater, FL*

[10] Laurent H and Radu H 1993 Figure-Ground Discrimination: a Combinatorial Optimization Approach *IEEE Transactions on Pattern Analysis and Machine Intelligence, Institute of Electrical and Electronics Engineers* **15(9)** 899-914

[11] Sarkar S and Boyer K L 1996 Quantitative measures of change based on feature organization: Eigenvalues and eigenvectors *Computer Vision and Pattern Recognition* **40** DOI: 10.1109/CVPR.1996.517115
[12] Williams L R and Jacobs D W 1997 Stochastic Completion Fields: A Neural Model of Illusory Contour Shape and Salience *Neural Computation* **9(4)** 837-858
[13] Shakowat M, Sarker Z, Tan W H and Logeswaran R 2007 Morphological based technique for image segmentation *International Journal of Information Technology* **14(1)** 55-80
[14] Sahoo P K 1988 A survey of thresholding techniques *Computer Vision, Graphics and Image Processing* **41** 233-260
[15] Rutkowski W S 1979 Shape completion *Comput Vis Graph Image Process* **9** 89-101
[16] Yakimov I, Kirpichnikov A, Mokshin V, Yakhina Z and Gainullin R 2017 The comparison of structured modeling and simulation modeling of queueing systems *Communications in Computer and Information Science (CCIS)* **800** DOI: 10.1007/978-3-319-68069-9_21
[17] Tutubalin P I and Mokshin V V 2017 The Evaluation of the cryptographic strength of asymmetric encryption algorithms *IEEE Second Russia and Pacific Conference on Computer Technology and Applications (RPC)* 180-183 DOI: 10.1109/RPC.2017.8168094
[18] Laurent H and Radu H 1993 Figure-Ground Discrimination: a Combinatorial Optimization Approach *IEEE Transactions on Pattern Analysis and Machine Intelligence, Institute of Electrical and Electronics Engineers* **15(9)** 899-914
[19] Sarkar S and Boyer K L 1998 Quantitative measures of change based on feature organization: Eigenvalues and eigenvectors *Computer Vision and Pattern Recognition* **7(1)** 110-136
[20] Williams L R, Jacobs D W 1997 Stochastic Completion Fields: A Neural Model of Illusory Contour Shape and Salience *Neural Computation* **9(4)** 837-858
[21] Borisova I V, legkii V N and Kravec S A 2017 The application of the brightness gradient orientation for automatic object tracking systems *Computer Optics* **41(6)** 931-937 DOI: 10.18287/2412-6179-2017-41-6-931-937
[22] Myasnikov E V 2017 Hyperspectral image segmentation using dimensionality reduction and classical segmentation approaches *Computer Optics* **41(4)** 564-572 DOI: 10.18287/2412-6179-2017-41-4-564-572
[23] Goshin Y V and Kotov A P 2017 Parallel implementation of a multi-view image segmentation algorithm using the hough transform *Computer Optics* **41(4)** 588-591 DOI: 10.18287/2412-6179-2017-41-4-588-591

Algorithm for encoding *n*D spatial objects into GIS

D E Andrianov[1], S V Eremeev[1], Y A Kovalev[1] and K V Kuptsov[1]

[1]State University named after Alexander Grigorievich and Nikolai Grigorievich Stoletov, Gorky str. 87, Vladimir, Russia, 600000

Abstract. In the article it is proposed to apply computer topology methods to create effective data structures that will allow storing and processing spatial scenes in n-dimensional space. The basis for the representation of *n*D-geoobjects is n-dimensional simplexes. For example, the representation of spatial data of high dimension will allow us to describe the topological relationships between 3D objects in time. The mathematical foundations and software for representation and processing of spatial data of high dimension are developed. The algorithm for encoding spatial objects *n*D in GIS will provide an opportunity to solve a wide range of tasks for processing complex graphic information.

1. Introduction

The development of GIS technologies in the modern world is advancing towards the use of 3D maps, multiscale, and also in the analysis of the time component. For high dimensions, those algorithms that work well for 2D and 3D cards are inapplicable. Integration in GIS of additional characteristics, such as the third spatial dimension, time, and scale, has so far been achieved mainly through special adaptations to 2D data structures, and not by creating new ones that extend the capabilities of GIS software.

The urgency of the work is that existing algorithms for processing and analyzing spatial scenes in *n*- dimensional space require significant time-consuming, or expected loss of processing quality. The integration of dimensions raises new requirements for creating new methods and algorithms for processing spatial data. The use of topological data analysis is a promising direction in GIS. The paper proposes to apply the topology of computer techniques to create efficient data structures that will store and process spatial scene in a single *n*-dimensional spatial [1, 2, 3, 4, 5].

For the destruction of objects in the GIS, there are different sets of data structures [6, 7, 8, 9]. In work [10] approaches to storage of spatial relations in GIS, such as: spatial queries on the basis of language SQL and matrix of topological relations are considered. The main attention in the article is given to the representation of natural hierarchical structures of spatial objects. A method for storing information about spatial relationships directly in the object identifier is developed, the structure of such an identifier is given.

The paper [11] discusses the structure of the database table in IBM DB2, which is designed to store and manage spatial data in both a regular table and store spatial data in accordance with the spatial geometry of IBM DB2. DB2 not only saved all spatial information, but also guaranteed spatial data and uniformity of these attributes under any conditions. All data stored in a relational database, so it can implement GIS applications by visiting network technology. Using the method given in this article, you can save a form document or another electronic document card in one kind of relational database. The spatial data table is mainly used to store the space of the geometric data objects of the

*.shp file, each form (such as a point, line and plane) consists of a table form, each spatial object in the table is stored as IBM Geometry type DB2 Spatial Extender, field The spatial type is ST Geometry, because the design of the data table structure saves all the information in the form file and it can be well integrated between the spatial data table and the attribute data table via related properties So, you can use the space data independently of the spatial data file, we just need to read the relevant information from the relational database in order to solve the data consistency problem and the integrity problem. This method of data storage will be useful when working with vector maps mainly for their storage and convenient access to objects, but will not allow access to their topological properties.

In [12] refers to the object-oriented repository (OBS), which was developed in 1998, is he starting point of high-performance data storage. Compared to the traditional file storage of the model, the OBS model uses the object interface and storage of the unloading of the management function from application to storage device.

OBS has the following key features:

1. The object is a logical block of memory that contains object data, object attributes, and methods;
2. the interface of the object is a simple method, such as creating, deleting, opening, closing, reading, writing, etc.;
3. OBS allows intelligent devices, including device and data management, object structure and interpretation relationships, access patterns and security settings.

Advantages of OBS:

1. high performance;
2. security;
3. scalability.

Disadvantages of this method is the inability to work with *n*D objects, since the functional does not provide for storing the properties of these objects.

The article [13] describes the possibility of storing the topology from the data model of the spatial topology of Oracle. The obvious differences from other DBMSs are:

Isolated nodes. The spatial topology data model allows you to integrate isolated nodes into the topology. An important advantage of this is that isolated nodes inside can be identified through links to the database and not resorting to spatial search.

Coordination of storage of the connected nodes. There is a small amount with Oracle's spatial topology data model. Coordinates for connecting nodes are stored twice. This is because the spatial topology data model stores each edge as an Oracle Spatial SDO_Geometry data type. This SDO_Geometry for the edge includes all coordinates, including those that correspond to the start and end nodes.

Thus, for these initial and final nodes, the latitude / longitude coordinate is stored twice; once as part of the SDO_Geometry and once on the connection node record. The API for the Oracle Spatial Topology Data Model manages this data transparently for the user and keeps it consistent.

The authors of Ref. [14] describe a complex approach to constructing a hierarchical structure of a road network for a continuous multiscale representation, especially continuous selective skipping of roads in a network. In this structure, the model of the road network is constructed using a linear and areal hierarchy. A continuous multiscale view of the road network can be achieved by searching in these hierarchies.

But these algorithms are not designed for *n*D objects, including for storing information about the scale of maps and the time of changing objects on the map. Since they allow you to store either only geometric information, or specific topological features.

The basis of representation *n*D geoobjects can be put *n*-dimensional simplexes. For example, the representation of spatial data of high dimension will allow to describe topological relations between 3D objects in time. There is a need to develop mathematical bases and software for the presentation and processing of spatial data of high dimensionality. The algorithm for coding *n*D spatial objects in the GIS will make it possible to solve a wide range of tasks for processing complex graphic information.

2. Encoding algorithm *n*D objects

Vector representations of *n*D objects exclude the following problems: they are usually more compact, they can describe boundaries more accurately than raster ones, and represent object attributes directly. This makes them particularly interesting for more advanced GIS, even if they were explored as data structures. Usually, the objects used in the 2D vector GIS are not considered in other dimensions. However, there are various multidimensional geometric and topological structures that were developed in other areas and can be used for these purposes. Even with the fact that they are difficult to implement and use, and often require additional computation to support some aspects of real data. Therefore, they were almost never used in practice. It was for these objects that a data storage structure was developed based on a compressed simplex tree.

A simplex is a convex hull of n + 1 points of an affine space (dimension n or greater) that are assumed to be affine independent (that is, they do not lie in a subspace of dimension n - 1). These points are called the vertices of a simplex. An example of a simplex is shown in figure 1.

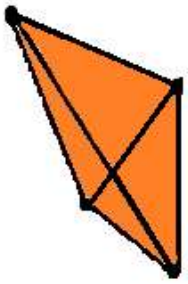


Figure 1. Simplex.

A simplicial complex is a topological space represented as a union of sets homoeomorphic to a simplex and forming a triangulation of this space such that:

1. with any of the simplexes in this set all its faces enter;
2. any two simplexes either do not have a common point at all, or intersect only along a whole face of some dimension, and only one face;
3. for any point x of the complex there is a neighbourhood U such that if it U intersects with the simplex of the complex Δ then $x \in \Delta$.

Let be $K = \{\sigma_0 \ldots \sigma_i\}$ a simplicial complex, which is shown in figure 2.

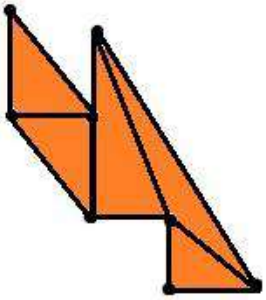


Figure 2. Simplicial complex.

In figure 2 shows a set of *n*-dimensional simplexes σ_i, where $i \in \{1 \ldots n\}$. Formally, let, $\sigma_i \in \{x_i, y_i, z_i, m_j, t_j\}$ where x_i, y_i, z_i the set of coordinates of the object in the set, $i \in \{1 \ldots n\}$, m_j - is the scale on which the object is displayed, and t_j is the time interval in the set into $j \in \{1 \ldots n\}$ which the object on the map was changed.

A simplicial tree of an undirected graph G is the spanning tree of a graph G with a distinguished root with the property that any two adjacent vertices in the graph G are related to each other by an ancestor / child relation. All search trees in depth and all the hamiltonian paths are simplex trees.

In finite graphs, although depth-first searches are inherently sequential, simplex trees can be constructed by a randomized parallel algorithm with a complexity class.

Simplicial trees can be used to determine the depth of the graph tree and as part of the planarity test for checking whether the graph is planar. The description of trees by the one-place logic of second-order graphs makes it possible to recognize the graph-dependent properties of orientation effectively for graphs with bounded tree width using the Course theorem. Not every infinite graph has a tree of simplexes and graphs, which have no such tree, can be described by forbidden minors.

A tree of simplexes exists in any graph with a countable number of vertices, even if the version of infinite depth search cannot successfully verify all vertices of the graph. In an infinite graph a tree must have exactly one infinite path for each ray of the graph and the existence of a simplicial tree characterizes graphs whose topological completions, formed by adding an infinitely distant point for each ray, are metric spaces.

To represent a compressed tree of simplexes K, we store the corresponding numbers in the tree that satisfy the following properties:

1. The nodes of a tree are in a bijection with simplexes of all dimensions of the complex. The first level is associated with an empty face.

2. Each node of the tree, except the first, preserves the label of the vertex. In particular, the node N associated with the simplex σ retains the label of the extreme vertex (σ).

3. Vertexes, whose labels occur along the path from the first level of the tree to the node N, are connected with the simplex σ, are vertices σ. The labels are sorted in ascending order along the given path, and each label appears only once. This data structure is called a simplicial tree K [15].

Consider the simplex tree in figure 3, which contains the following sublevels:

1. the card;
2. scale;
3. The object;
4. coordinates of the object;
5. Time;
6-7. other sublevels

It should be noted that in this figure an area is created, beginning with sublevel 2. These objects are compressed, identical and are on different scales.

The purpose of compression is to identify the common parts of the object and save them only once. More specifically, if the same subtree is located on two different nodes of the simplex tree, then the subtree is stored only once, and the two root nodes now point to a unique copy of the subtree. As a consequence, the nodes are no longer in a bijection with the nodes of the complex, but we still have the property that the paths from the root are in a bijection with simplexes.

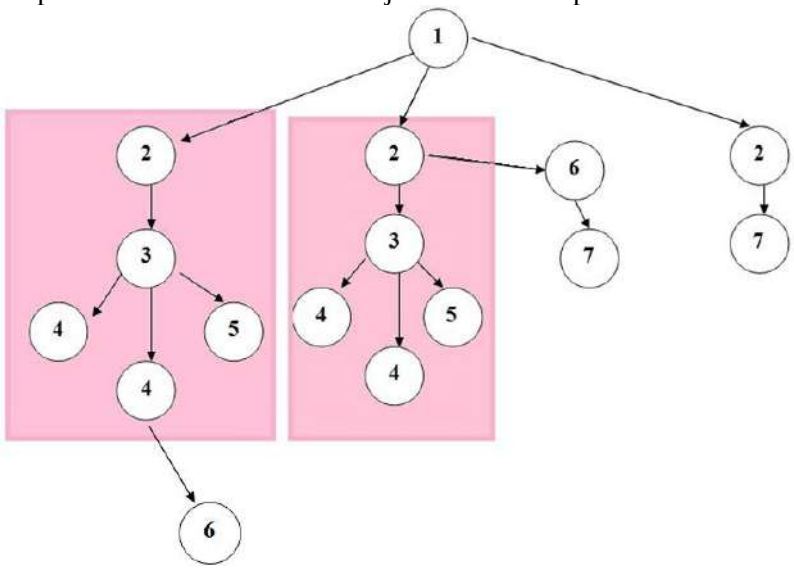


Figure 3. Simplex tree.

Access to all the heirs of a compressed simplex tree can be realized similarly to the usual simplex tree. Allowing an ascending traversal in the tree is also possible (with additional pointers from descendants to ancestors), and this improves the efficiency of some operations, such as surface searches. However, in the compressed simplex tree, the ancestors are not unique. To accommodate this, we mark the ancestors that were available, and use this to return in the upward direction.

3. The results of the algorithm

Figure 4 shows the images of the object on different scales, which changed over time.

The main search takes place on the largest scale, on the remaining scales of the maps, a search is performed and selection of previously found objects by the identifier. Detailed search only occurs in the corresponding buffer zones, and not across the entire map, which allows you to significantly reduce the time to find the right objects.

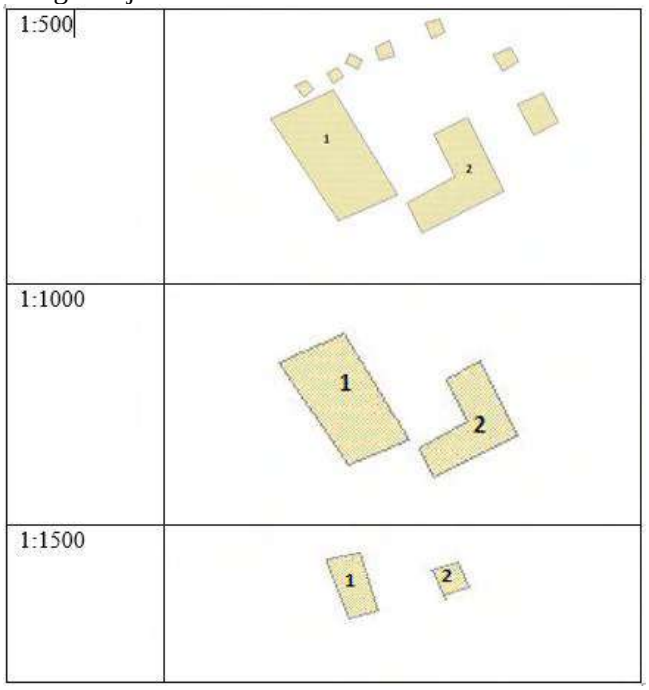


Figure 4. Displaying an object on different scales.

The result of the operation of the *n*D coding algorithm is shown in figure 5 as a graph showing the objects on scales of 1:1000 and 1:1500.

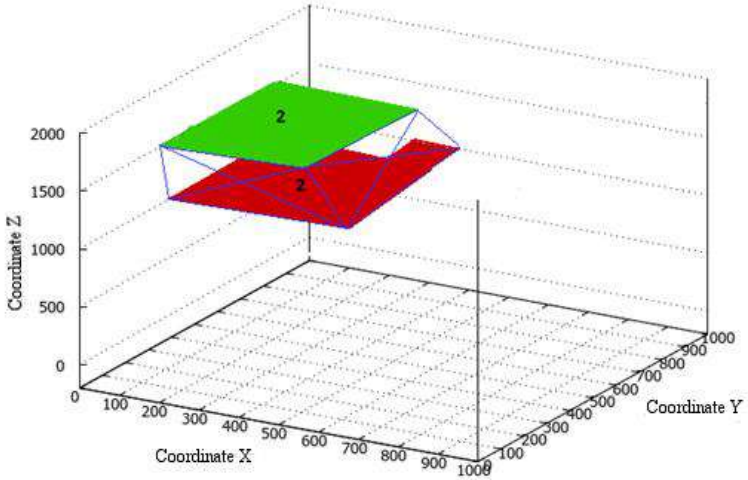


Figure 5. Graph of found objects on different scales. Where X and Y are the coordinates of the object, and the axis Z is the scale of the map (1:500, 1:1000 and 1:1500).

Green on the chart highlighted the object "2" on a scale of 1:1500, red highlighted the object "2" on a scale of 1:1000, the lines that unite them form simplexes.

The work of the algorithm was tested when searching for objects on different scales of the map. To search for objects, an algorithm was used to search for spatial objects based on specified criteria based on buffer zones in multi-scale GIS [16]. The result is displayed in the graph of figure 6.

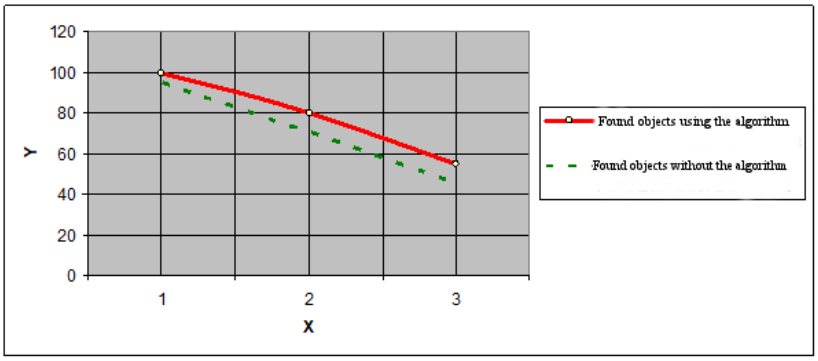


Figure 6. The plot of found objects on different scales.

The graph in Figure 6 shows the search for objects on different scales. Where the axis X is the scale of the map (1:500, 1:1000 and 1:1500), Y is the number of objects on the map. On the basis of the work, a graph was constructed in which an increase in the number of found objects using the developed algorithm by 10% is observed with respect to the search for objects without using it.

Also a graph was constructed based on the search of objects for a certain period of time.

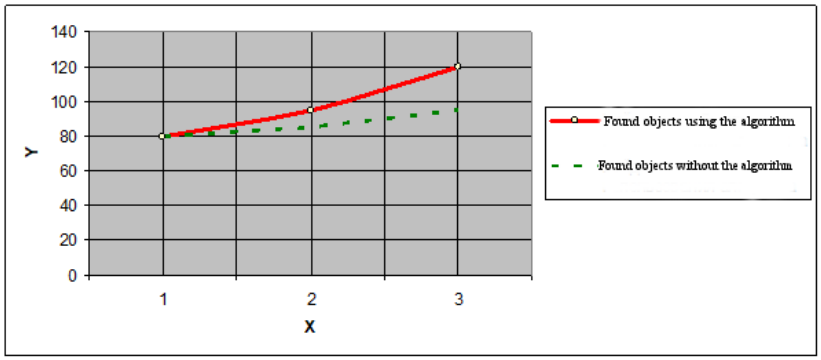


Figure 7. The schedule of the found objects in different time interval.

Figure 7 shows a graph with the search for objects in different periods of time. Where the axis X is the time period in years (2009, 2013 and 2016), Y is the number of objects on the map. The graph shows an increase in the number of found objects using the developed algorithm with an efficiency of 15% relative to the search for objects without using it.

4. Conclusion

The algorithm of encoding of *n*D spatial objects in GIS is developed in the article. A data structure based on a simplex tree was also developed and analogs of data structures for multi-scale GIS were considered.

This algorithm can be used in construction, for example, when testing soil changes in recent years. Also, the algorithm can be used to bypass obstacles, for example a quadcopter, which will analyze the presence of an obstacle in its path on the basis of the map.

5. References

[1] Eremeev S V, Andrianov D E and Komkov V A 2013 Algorithms for the formation of a graph model of urban territory in the GIS *Geoinformatics* **4** 19-24 (in Russian)

[2] Bulaev A V 2008 A formal model for establishing topological relations with objects that contain curvilinear segments *Algorithms, Methods and Systems of Data Processing* **13** 16-24 (in Russian)
[3] Sharapov R V and Varlamov A D 2013 Basic metrics that assess the quality of the work of image search systems *Algorithms, Methods and Systems of Data Processing* **2** 3-11 (in Russian)
[4] Gashnikov M V, Glumov N I, Kuznetsov A V, Mitekin V A, Myasnikov V V and Sergeev V V 2016 Hyperspectral remote sensing data compression and protection *Computer Optics* **40(5)** 689-712 DOI: 10.18287/2412-6179-2016-40-5-689-712
[5] Kopenkov V N and Myasnikov V V 2016 Development of an algorithm for automatic construction of a computational procedure of local image processing, based on the hierarchical regression *Computer Optics* **40(5)** 713-720 DOI: 10.18287/2412-6179-2016-40-5-713-720
[6] Fursov V A, Goshin Ye V and Kotov A P 2016 The hybrid CPU/GPU implementation of the computational procedure for digital terrain models generation from satellite images *Computer Optics* **40(5)** 721-728 DOI: 10.18287/2412-6179-2016-40-5-721-728
[7] Myasnikov E V 2017 Hyperspectral image segmentation using dimensionality reduction and classical segmentation approaches *Computer Optics* **41(4)** 564-572 DOI: 10.18287/2412-6179-2017-41-4-564-572
[8] Sadykov S S, Bulanova Y A, Kanunova E E and Zakharova E A 2014 Development of concepts for constructing an information system for diagnosing neoplasms on mammograms *Information Technologies* **10** 51-56 (in Russian)
[9] Teryokhin A V 2014 The concept of recognition of arbitrarily located three-dimensional objects from two images of projections *Algorithms, Methods and Systems of Data Processing* **2** 29-40 (in Russian)
[10] Eremeev S V and Filimonov M M 2014 Algorithm for coding spatial identifiers in hierarchical topological systems *Algorithms, Methods and Systems of Data Processing* **4(29)** 50-58 (in Russian)
[11] Zhou Yan and Zhu Qing 2006 A scalable distributed spatial data storage mode *International Society for Photogrammetry and Remote Sensing* **XXXVI** 855-859
[12] Zhu Bin and Wang Anbao 2011 The Storage Technology for GIS Data Realization *Journal of Computers* **6(10)** 2229-2236
[13] Zhilin Li and Qi Zhou 2012 Integration of linear and areal hierarchies for continuous multi-scale representation of road networks *Intern. J. of Geographical Information Science* **26** 855-880
[14] Galdi D E 2005 Spatial Data Storage and Topology in the Redesigned MAF *TIGER System* **3** 103-112
[15] Boissonnat J-D, Srikanta K C and Tavenas S 2015 Building Efficient and Compact Data Structures for Simplicial Complexe *An extended abstract appeared in the proceedings of SoCG* **32** 24-31
[16] Kovalev Y A and Eremeev S V 2016 An algorithm for finding spatial objects based on given criteria based on buffer zones in multiscale GIS *Proceedings of the 26th International Scientific Conference "GRAPHICON"* **19** 414-416 (in Russian)

Acknowledgments
The reported study was funded by RFBR and Vladimir region according to the research project № 17-47-330387 and № 326 of 29.09.2017.

High Resolution Imaging for Bistatic P and VHF Band SAR

A V Borisenkov[1] and O V Goriachkin[1]

[1]Povolzhskiy State University of Telecommunications and Informatics, Lev Tolstoy str. 23, Samara, Russia, 443010

Abstract. In the article bistatic SAR imaging algorithm for P and VHF bands is described. The main feature of the system is that it uses two probing signals emitted simultaneously in two different frequency bands. As a consequence the spatial resolution of the proposed SAR can be potentially increased up to two meters. The results of full-scale ground experiment involving the proposed SAR are presented in the article.

1. Introduction

Spaceborne Synthetic Aperture Radars (SAR) for Earth remote sensing using P and VHF bands have been widely discussed over the last several years. Radar images for these bands contain the information about the reflection coefficient inside vegetation and soil [1,2]. Wideband SAR remote sensing using 20-500 MHz band and 100+ MHz bandwidth onboard unmanned aerial vehicles (UAV) and other mobile carriers opens up new application possibilities for Earth remote sensing and defence systems. The most interesting of these possibilities is the detection of small objects concealed by foliage or located below ground.

Airborne ground penetrating radars are known since the late 80's of the 20th century. In the year 1988 JPL started the AIRSAR platform development for VHF band imaging. Around the same time Stanford Research Institute and Swedish Defence Research Agency begun the development of longwave radars. In Soviet Union similar projects were started in the year 1989 by the Institute of Radio-engineering and Electronics in Kharkiv and by the Central Construction Bureau in Samara under the leadership of Professor A.I. Kalmykov.

Currently there exist several airborne radar systems that allow detection through foliage and below ground: Carabas FOPEN Demonstrator developed by Saab Group, Sweden; FOPEN Reconnaissance, Surveillance, Tracking and Engagement Radar; TRACER – a dual band (UHF/VHF) radar; Boeing A160T – a P-band radar operating onboard an unmanned helicopter [2-5].

From the year 2005 to the present time a scientific group of the Earth Radar Remote Sensing Center (ERRSC PSUTI, Samara) have been developing novel remote sensing systems. From global expensive space and aviation systems to local and relatively cheap but sufficient quality technologies for UAVs, ground transports or stationary ground-based infrastructures [4-6].

In contrast to works [1-3] in this paper an imaging algorithm for Bistatic SAR (BiSAR) using P and VHF bands simultaneously is described. The SAR was developed by ERRSC PSUTI during the years 2013-2017. Field test results for the SAR are presented. The test was aimed at obtaining ultra-high resolution images utilizing the combined 80 MHz data composed from 30 MHz P band and 50 MHz VHF band data.

2. Equipment of BiSAR

BiSAR radio equipment consists of two pulse transmitters which can produce stable chirp and/or PSK signals at carrier frequencies of 435 and 145 MHz. Signal bandwidth can be adjusted from 1 to 30 MHz in P band and from 25 to 50 MHz in VHF band. The equipment is mounted on a car and utilizes autonomous primary and secondary power sources.

Both onboard transmitter and ground-based receiver use a Yagi-Uda antenna. Ground-based stationary receiver consists of two dual channel tuned radio frequency receivers with up to 110 dB power gain. Each receiver records in-phase and quadrature components of a signal at 200 MHz frequency using digital representation for subsequent digital signal processing. Receiving antenna height can vary from 8 to 12 meters.

The described configuration of the radar system allows to acquire synchronous radio images for P and VHF bands and also a combined high resolution image using 30 and 50 MHz signals simultaneously. The idea of the method is illustrated by Fig. 1.

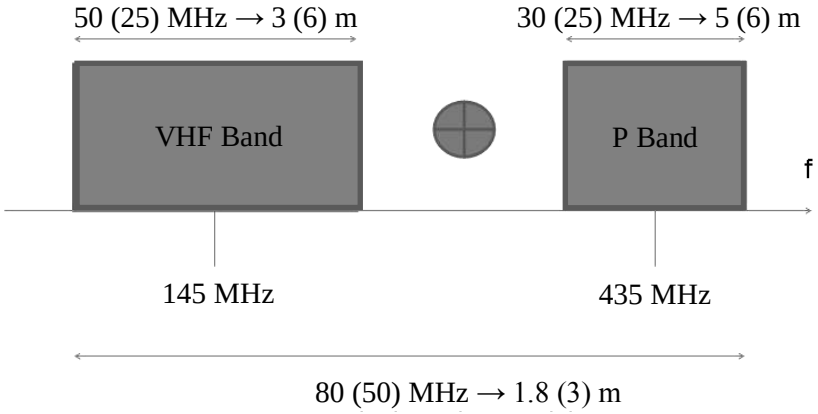


Figure 1. Combination of two frequency bands for SAR range resolution increase.

3. Imaging algorithm

In the paper the specifics of radar imaging algorithm for ground-based experiment are considered. For the BiSAR modification in question Signal-to-Noise Ratio (SNR) for direct channel is high enough, so we can use a simple threshold method to detect probing signals in the presence of noise and reflected signals, which provides the required synchronization of each of the frequency bands independently.

We can describe the steps of radar imaging for each of the frequency bands as follows:

1. Synchronous detection and band-pass filtering of both direct and reflected BiSAR signals recorded by Analog-to-Digital Converter (ADC);
2. Range compression using digital matched filter with chirp or PSK pulses of both direct and reflected BiSAR channels for each of the frequency bands;
3. Synchronization of the direct and reflected channels by producing a file of compressed probing signal sample indices;
4. Producing of two-dimensional (2D) radio hologram files using direct and reflected signals in P and VHF band;
5. Making of a combined 2D radio hologram file for frequency shift of f_S (8-25 MHz during the experiment);
6. Computing of a combined dual frequency radio image using the following formula:

$$\hat{\dot{\xi}}(x,y)=\int_{t_1}^{t_2} G_\Sigma(t,x,y)\,s_2^*\big(t-\tau_2(t,x,y)+\tau_1(t)\big)\,\dot{s}_2(t)\cdot \left(\begin{array}{l}\exp\big(-j2\pi(f_P-f_s)(\tau_2(t,x,y)-\tau_1(t))\big)+\\ +\exp\big(-j2\pi(f_{VHF}+f_s)(\tau_2(t,x,y)-\tau_1(t))\big)\end{array}\right)dt \quad (1)$$

Here $\hat{\dot{\xi}}(x,y)$ is the backscattering coefficient (radar image); $\dot{s}2(t)$ is the combined 2D radio hologram signal which consists of a sequence of reflected and probing pulses after preliminary matched filtering

at zero carrier frequency, frequency shifting and summation of the signals from the two frequency bands; $\tau_2(t, x, y)$ is the time delay of the signal, reflected by a point target at the coordinates (x, y); $\tau_1(t)$ is the time offset caused by transmitter instability and estimated by synchronization system; $G_\Sigma(t, x, y)$ is the weight function, which depends on movement patterns of transmitting and receiving BiSAR antennas.

Fig. 2 shows a block diagram of the described imaging algorithm.

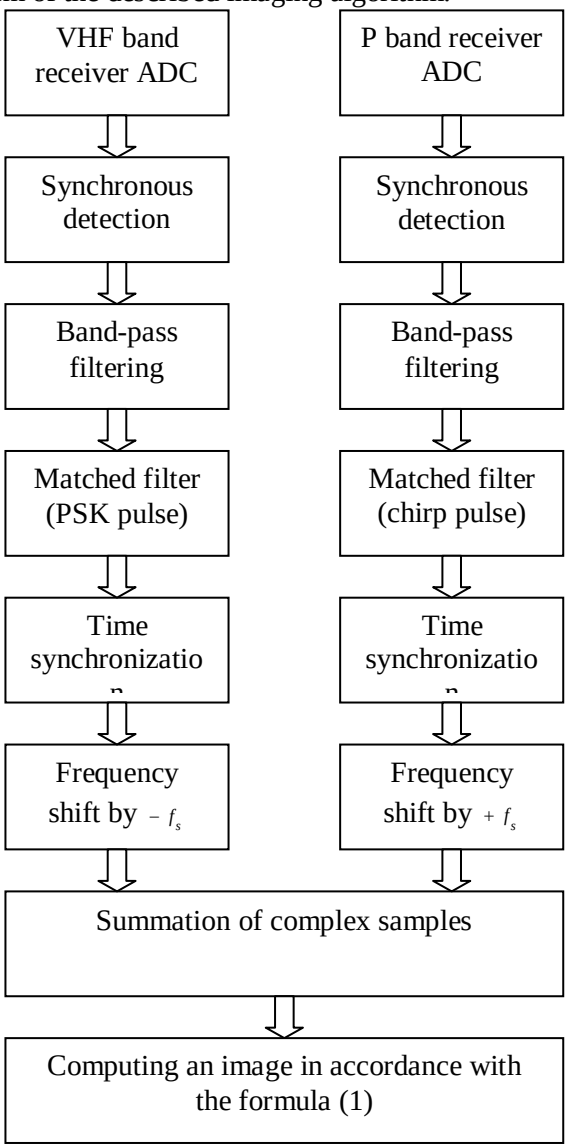


Figure 2. Block diagram of the imaging algorithm.

4. Ground-based experiment results

In the paper the results of the ground-based experiment involving the proposed dual frequency bistatic SAR are presented. For BiSAR system to function properly, the transmitter should be moving while emitting a signal. During the experiment we used a car moving across a bridge. The bridge height was about 20 meters. Fig. 3 shows the geometry of the experiment.

Fig. 6 shows the resulting radar image of the scene of the experiment (4×8 km) combined with optical image of the same scene acquired via Google maps service.

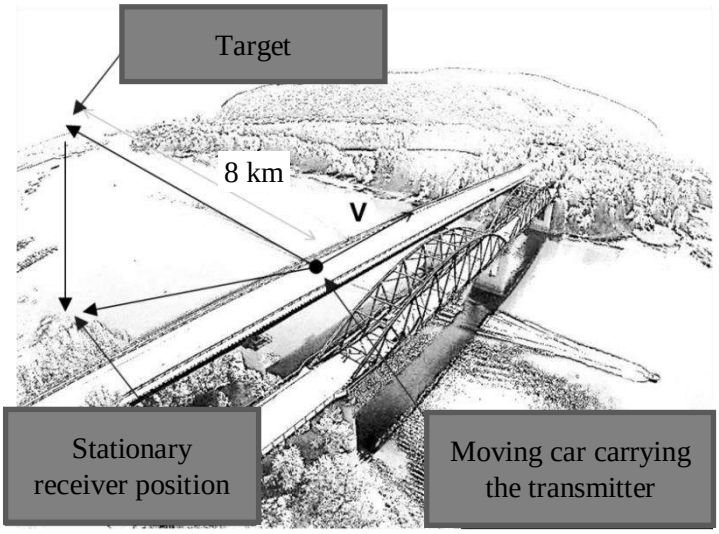


Figure 3. Ground-based experiment geometry.

Figure 4. Stationary receiver position deployed in the field.

Figure 5. Moving car, carrying the transmitter across Sok River (see Fig. 6 for the scene of the experiment).

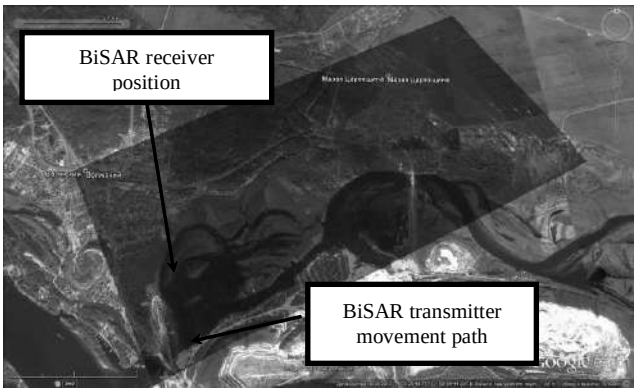


Figure 6. Scene of experiment.

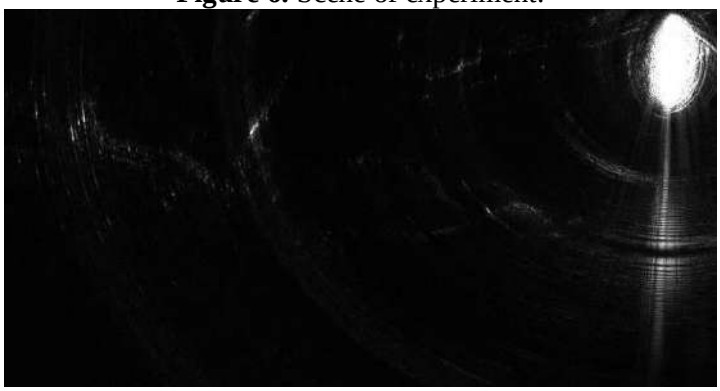


Figure 7. Radar image obtained in P band using chirp signal with 30 MHz bandwidth; pixel resolution is 1.5×1.5 m; synthesis time is 40 sec.

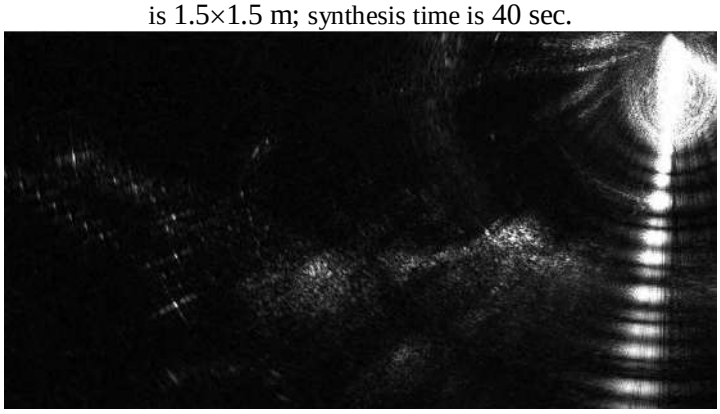


Figure 8. Radar image obtained in VHF band using PSK signal with 25 MHz bandwidth; pixel resolution is 1.5×1.5 m; synthesis time is 40 sec.

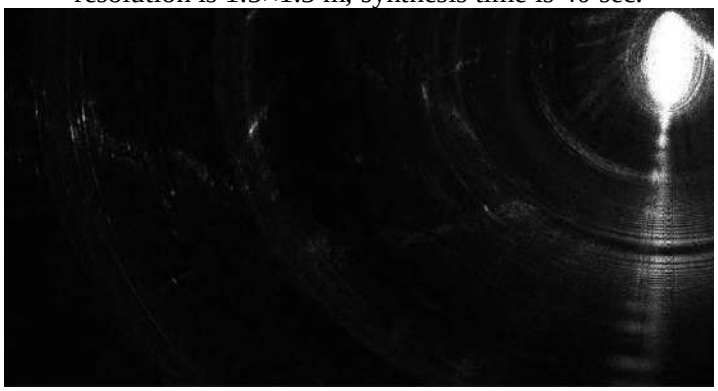


Figure 9. Radar image obtained by combining P and VHF bands; bandwidth is 80 MHz; pixel resolution is 1.5×1.5 m; synthesis time is 40 sec.

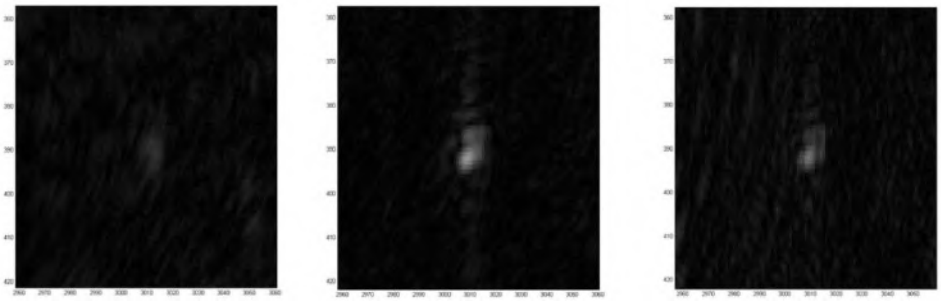


Figure 10. Radar image of the first target, left to right: VHF band, P band, VHF+P bands, pixel resolution is 1.5×1.5 m.

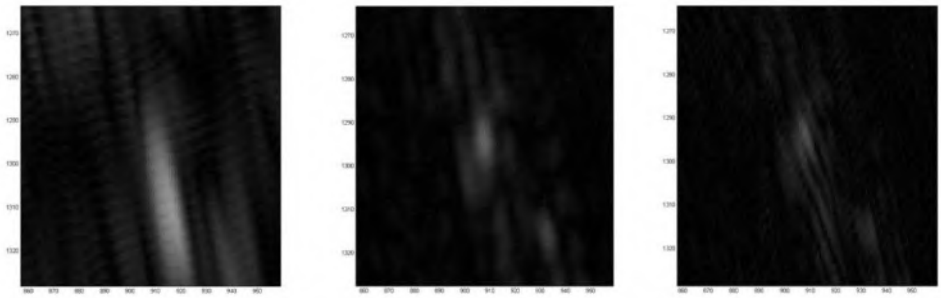


Figure 11. Radar image of the second target, left to right: VHF band, P band, VHF+P bands, pixel resolution is 1.5×1.5 m.

Fig. 7 and 8 show radar images in P and VHF bands respectively obtained during the ground-based experiment; pixel resolution is 1.5×1.5 m; synthesis time is 40 sec. Fig. 7 shows the image with 30 MHz bandwidth that corresponds to potential spatial resolution of 5 meters and Fig. 8 shows the image with 25 MHz bandwidth that corresponds to potential spatial resolution of 6 meters (slant range resolution).

Fig. 9 shows the combined radar image for P and VHF bands with corresponding potential spatial resolution of 3 meters (slant range resolution). Fig. 10 and 11 show the matching fragments of the radar images in VHF, P and VHF+P bands that confirm the resolution increase.

During the experiment the area was observed with a small altitude (5-7 degrees), thus, radar images show only the taller objects, such as masts, power transmission lines, tall buildings, lone tall trees, wooded areas, etc.

The results of the experiment allow us to conclude that there exists a possibility to create a complex of onboard and ground-based equipment together with algorithms and software for the BiSAR system that allows to obtain radar images for VHF and P bands and produce high resolution images for combined VHF+P band with spatial resolution up to 2 meters.

For the year 2018 we are planning to conduct a similar experiment using UAV (octocopter drone) to carry the mobile BiSAR transmitter.

5. Conclusion

The results of the full-scale ground experiment show that using both VHF and P bands with respective resolutions of 6-10 and 6 meters it is possible to obtain a combined radar image with range resolution of 3-6 meters. Thus, the proposed technology demonstrates the possibility of combining VHF and P band probing signals to increase the spatial resolution of bistatic SAR systems.

6. References

[1] Hellsten H and Ulander L M H 2000 Airborne Array Aperture UWB UHF Radar - Motivation and System Considerations *IEEE Aerospace and Electronic Systems Magazine* **15(5)** 35-45

[2] Ulander L, Pierson W, Lundberg M, Follo P, Frolind P-O and Gustavsson A 2004 CARABAS-II SAR Change Detection Performance on Ground Targets concealed by Foliage *Proceedings of 5th European conference on synthetic aperture radar* 297-300

[3] de Macedo K A C, Shiroma G H X, Wimmer C and Moreira J R 2015 Monitoring under foliage with airborne SAR *Proceedings of IEEE Intl. Radar Conference At* 499-505
[4] Goriachkin O V, Zhengurov B G, Bakeev V B, Baraboshin A Y, Nevski A V and Scorobogatov E G 2015 Bistatic synthetic aperture radar P band for small satellite "AIST-2" *Electrosvyaz* **8** 34-39 (in Russia)
[5] Goriachkin O V 2016 Earth Radar Remote Sensing Center in PSUTI *Vestnik svyazi* **2** 24-26 (in Russia)
[6] Zherdev D A, Kazanskiy N L and Fursov V A 2015 Object recognition in radar images using conjugation indices and support subspaces *Computer Optics* **39(2)** 255-264 DOI: 10.18287/0134-2452-2015-39-2-255-264

A comparison of iris image segmentation techniques

M S Semyonov[1] and E V Myasnikov[1]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086

Abstract. The paper compares three different methods of iris image segmentation, namely the method using the Daugman`s integro-differential operator, the method using the Hough transform for the detection of circles, and the method based on the analysis of the distribution of edge points. The accuracy and run time of the implemented methods were estimated in the experimental study conducted using the MMU Iris Image Database. The carried out research has shown that the method using the Daugman`s integro-differential operator has the greatest accuracy, and the method based on the analysis of the distribution of boundary points has the shortest operating time.

1. Introduction

In recent years, identifying a person by his/her biometric parameters becomes a more and more popular field of research. The reason for this is both the increased interest in the biometric identification by public and commercial bodies and the intensive development of technical tools of recording, storage and processing of such information. Today, a broad range of features of human body is used for the identification, including face images [1, 2], fingerprints, iris patterns, voice, manner of walking etc. The identification of a person by an iris pattern is of particular interest in this list, because this method ensures the most accurate identification [3].

In fact, the iris is a circular movable diaphragm with a diameter of about 12 mm, at the core of which there is a round hole, a pupil [4]. An example iris image is presented in figure 1.

The iris color (eye color) depends on the coloration (pigmentation) of iris anterior surface visible through the transparent cornea [5]. However, if the iris color changes during the first 10-12 years of a human life and may change in old age, the iris pattern with clear individual characteristics, which is formed during the intrauterine growth, remains almost the same throughout the human life [6, 7]. These iris characteristics cause significant interest to this identification method.

Generally, the identification systems using an iris image are built according to the classical scheme:

- selection of an informative region (iris) on the image,
- feature description of the selected region,
- comparison of the feature descriptions.

In this case, the quality assessment of both obtained images and effectiveness of individual method stages may be included into the above scheme [8].

It should be noted that the accuracy of the first of the above-mentioned stages (selection of the iris on the image) plays an important role in the achievement of high quality.

The selection of the region corresponding to the iris on the image is the segmentation-related task that can be completed using a variety of methods. Specifically, the approaches based on the edge detection [9], thresholding [10], Daugman's integro-differential operator [11], circular Hough Transform [12] etc. are popular.

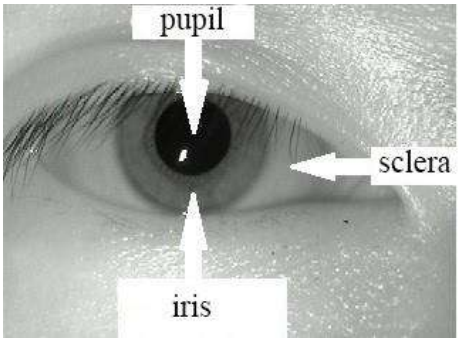


Figure 1. Example image from the data set [15].

When using these approaches, the internal and external iris boundaries (boundaries "iris-sclera" and "pupil-iris") are sought, and the further approximation of detected boundaries with circles is performed. It should be noted that such approximation is only the first stage of a segmentation, because after selecting the ring (which may be non-concentric) formed by these circles, the part of the iris may remain closed and shaded by eyelids and eyelashes.

This paper describes the experimental comparison of three approaches to the selection of the iris on the image.

This paper has the following structure. Sections 2-4 briefly describe the methods considered in this report. Specifically, Section 2 describes the method based on the Daugman's integro-differential operator, Section 3 describes the method based on the circular Hough transform, and Section 4 describes the method based on the edge detection and edge points distribution analysis. The results of experimental studies carried out with MMU Iris Image Database are presented in Section 5. In the end of this paper, there are Section 6 containing the conclusion and Section 7 containing the reference list.

2. The method based on the Daugman's integro-differential operator

According to the method based on the Daugman's integro-differential operator, the possible centers of the iris should be calculated first on an input image I_{in}. To do this, the points on I_{in} with the brightness level below the predefined threshold are selected. Then, the points, which correspond to the local minima of the brightness, are selected from the above-mentioned points.

Local minimum is the minimum of the circular neighborhood with a radius equal to the minimum iris radius. In this method, the minimum and maximum radii of the iris should be set to overlap a little bit the iris radius values from the image database.

Then, the parameters of circle approximating boundary "iris-sclera" should be calculated. This is done by using the Daugman's integro-differential operator [11] defined as the maximum of the smoothed derivative of the average image intensity along the round contour with coordinates (x_0, y_0) and radius r:

$$\max_{(r,x_0,y_0)} \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r,x_0,y_0} \frac{I(x,y)}{2\pi} ds \right|.$$

Here $I(x, y)$ is the image intensity function, $G_\sigma(r)$ is the Gaussian smoothing function, s is the circle contour with the coordinates (x_0, y_0) and radius r. This operator is applied for the points of I_{in} selected as possible iris centres and having r values changing from minimum to maximum iris radius.

Example of how the method described in this section works for figure 1 is presented in figure 2.

3. The method based on the circular Hough transform

According to the method based on the circular Hough Transform, the Canny edge detector [13] is applied first to the input image I_{in} to generate the image of edge points I_{ed1}. The intervals of the pupil and iris radius values used for searching should be established beforehand, depending on an existing image database, so as to overlap a little bit these values.

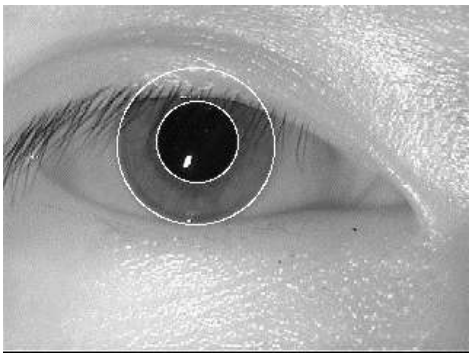


Figure 2. The result of the method based on the Daugman's integro-differential operator for the image depicted in Figure 1.

Then, the parameters of the circle, which approximates the "iris-sclera" boundary, should be determined on the image I_{ed1} of edge points using the circular Hough transform [9]. Maximum and minimum radii used for searching are taken from the pre-determined interval of iris radius values.

Thereafter, image I_{ed2} is made up of the points of the image I_{ed1} that are within the circle determined at the first stage. The parameters of the circle, which approximates the "pupil-iris" boundary, should be determined on the image I_{ed2} using the circular Hough transform [9]. Maximum and minimum radii used for searching are taken from the pre-determined interval of pupil radius values.

An example obtained using this method is shown in figure 3.

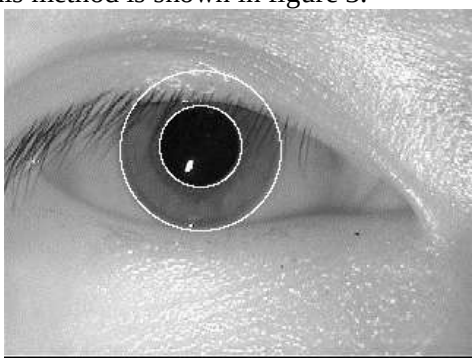


Figure 3. The result obtained using the method based on the circular Hough transform for the image depicted in figure 1.

4. The method based on the analysis of the edge points distribution

As a part of this work, the method based on the circular Hough transform and the analysis of the edge point distribution was proposed. This method includes the following stages. At the first stage, the iris internal boundary (boundary "pupil-iris") should be selected on the input image, and the parameters of the circle approximating the pupil should be determined. At the second stage, the iris external boundary (boundary "iris-sclera") is selected, and the parameters of the circle approximating the iris external boundary are determined using the least squares method. These stages are described below in more detail.

4.1. Pupil boundary determination

To determine the internal boundary of the iris, the Canny edge detector [13] is applied to the input image I_{in} to generate the image I_{ed1} of edge points. Then, using the circular Hough transform [9], the centre and the radius R_{pupil} of a pupil are determined on I_{ed1}. The minimum and maximum circle radii used for searching with Hough transform should be taken so as to overlap a little bit the pupil radius values from the image database.

Figure 4 below provides the example of pupil boundary determination for the image depicted in figure 1.

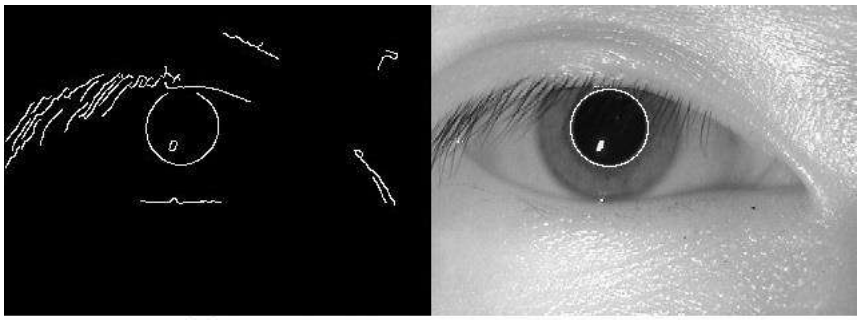


Figure 4. Example of pupil boundary determination for the image depicted in figure 1: (a) - edge points image; (b) - method results.

4.2. Determination of iris external boundary

To determine the iris external boundary, the function of the edge point distribution from the pupil centre, which is described in [14], is analyzed. The Canny edge detector [13] is applied to the input image I_{in} to generate the image I_{ed2} of edge points. The parameters of the Canny edge detector, which used to search for the pupil and iris boundaries, are different and should be selected in an experimental fashion based on the existing image set. It is known [14] that the value of the iris external radius R_{iris} is within the range ($5/4R_{pupil}$, $5R_{pupil}$), so the points, the distance from which to the pupil centre is not within the specified range, shall be removed from I_{ed2}. The function *f(R)* of the edge point distribution on the distance R to the centre of the pupil shall be calculated for the obtained image. It is expected that this function has a local maximum around the target iris radius value, because the pupil and the iris are approximately concentric. However, this maximum may hide among other maxima caused by the noise contamination of the image I_{ed2}.

An example of the graph of the function *f(r)* of the edge point distribution on the distance R to the centre of the pupil is shown in figure 5. The argument r of the function *f(r)* equals to the difference between R and R_{pupil}.

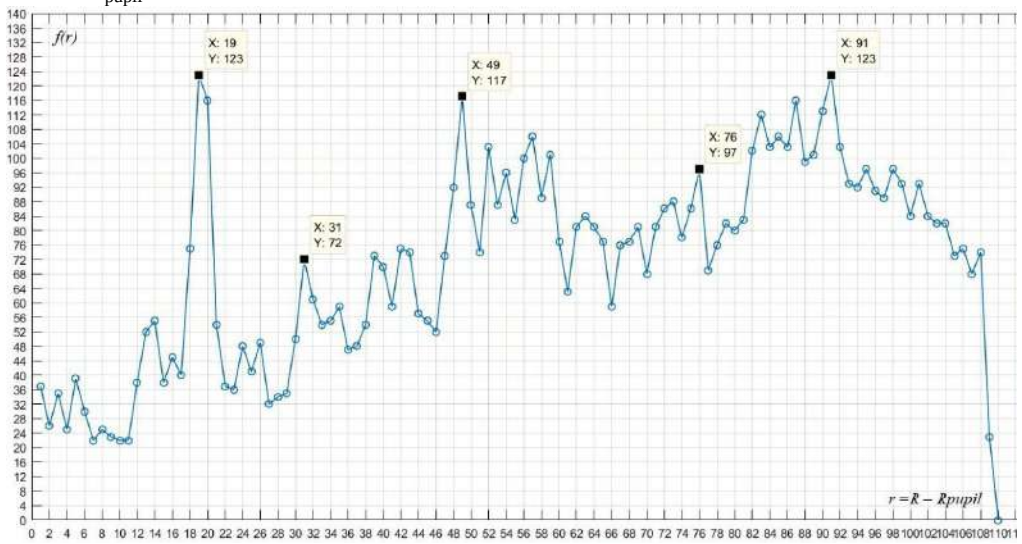


Figure 5. Example of the graph of the function *f(r)* of the edge point distribution on the distance R to the centre of the pupil.

Thereafter, the direct search for the local maxima of the *f(R)* function is performed starting with the highest maximum. For the selected local maximum R_M, the edge points, which fall into the ring defined by the radii $R_M - s$ and $R_M + s$, are approximated by the circle using the least squares method. The circle, for which the approximation error is lower than the predefined value, is selected. It is considered that this circle approximates the iris external boundary.

An example obtained using this method is shown in figure 6.

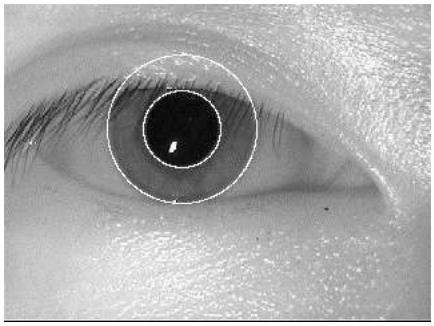


Figure 6. The results of the method based on the analysis of the edge point distribution for the image depicted in figure 1.

5. Experiment

To compare the accuracy of segmentation and operation time of the above methods, the simulation experiment was carried out. All methods were implemented in MATLAB. The experiment is carried out on a personal computer powered by Intel Core i5-4210M with a frequency 2.6 GHz and random-access memory of 6 GB.

When conducting this experiment, 30 iris images taken from MMU Iris Image Database [15] were used. For each image, the expert performed the reference (true) segmentation and defined the true values of pupil and iris boundary parameters (centre coordinates and radii). Thereafter, the similar values of parameters were assessed using each of the methods considered in this paper.

To assess the precision of each method, the Jaccard index [16] averaged over all test images was used. The Jaccard index is defined in this paper as the ratio between the intersection of rings (non necessarily concentric) with reference and evaluated parameters, and the area of interconnection of these rings. Besides, the image-averaged operation time of the assessed methods was evaluated.

The results of the experiments are presented in Table 1 and figures 7, 8. Specifically, “A” section describes the method based on the Daugman's integro-differential operator, “B” section describes the method based on the circular Hough transform, and “C” section describes the method based on the analysis of the edge point distribution.

Table 1. Experimental results.

Method	Precision		Operation time (s)	
	Average value	RMSD	Average value	RMSD
A	0.90	0.03	4.36	1.14
B	0.86	0.03	1.74	0.49
C	0.86	0.05	0.76	0.06

These results demonstrate that the Daugman's integro-differential operator is the most precise, but requires the longest operation time. The method based on the circular Hough transform is a little bit less precise, but has a significantly lower operation time (in 2.5 times).

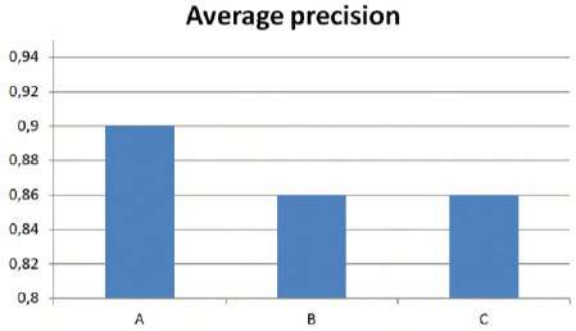


Figure 7. Average precision value for three studied methods: the method based on the Daugman's integro-differential operator (A), the method based on the circular Hough transform (B), the method based on the analysis of the edge point distribution (C).

The method based on the analysis of the edge point distribution was the fastest one (almost 6 times faster than the method of Daugman's integro-differential operator). In this case, it showed the same precision with the used data set as the method based on the Hough transform.

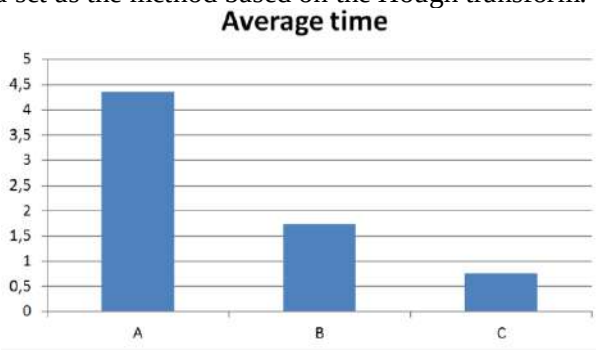


Figure 8. Average operating time (in seconds) for three methods: the method based on the Daugman's integro-differential operator (A), the method based on the circular Hough transform (B), the method based on the analysis of the edge point distribution (C).

6. Conclusion

Three methods of assessment of iris boundary parameters were evaluated in this study, namely: the method based on the Daugman's integro-differential operator, the method based on the circular Hough transform, and the method based on the analysis of the edge point distribution.

For images from MMU Iris Image Database, the reference segmentation of iris images was performed, and numerical experiments to assess the precision and operating time of implemented methods were carried out. The experiments show that the method of the Daugman's integro-differential operator is the most precise with regard to the applied data set, and the method based on the analysis of the edge point distribution proposed in this paper is marked by the lowest operation time and has the same level of precision as the Hough transformation.

In future, it is planned to use study results to develop the method of person identification by the picture of iris.

7. References

[1] Nemirovskiy V B, Stoyanov A K and Goremykina D S 2016 Face recognition based on the proximity measure clustering *Computer Optics* **40(5)** 740-745 DOI: 10.18287/2412-6179-2016-40-5-740-745

[2] Savchenko A V 2017 Maximum-likelihood dissimilarities in image recognition with deep neural networks *Computer Optics* **41(3)** 422-430 DOI: 10.18287/2412-6179-2017-41-3-422-430

[3] *Multiple Biometric Grand Challenge (MBGC - 2007)* (Access mode: http://www.nist.gov /itl/iad/ig/mbgc.cfm)

[4] Petrovsky B V 1988 *Large Medical Encyclopedia* (Moscow: Soviet Encyclopedia)

[5] Pryves M G, Lysenkov N K and Bushkovich V I 1974 *Human Anatomy* (Moscow: Medicine) p 670

[6] Velkhover E S and Ananin V F 1991 *Introduction to iridology. Pupillodiagnosis* (Moscow: UDN) p 212

[7] Konovalov V V and Antonov A A 1990 *Practical Iridology* (Moscow: Okulius) p 88

[8] Matveyev I A 2014 *Methods and Algorithms for Automatic Processing of Images of the Iris* (Moscow)

[9] Gonzalez R and Woods R 2005 *Digital Image Processing* (Moscow: Technosphere)

[10] Ling L L and Brito D F 2010 Fast and Efficient Iris Image Segmentation *J. Medical and Biological Engineering* **30(6)** 381-392

[11] Daugman J G 2004 How iris recognition works *Proc. IEEE Trans. Circ. Syst. Video Technol.* **14(1)** 21-30

[12] Basit A and Javed M 2007 Localization of iris in gray scale image using intensity gradient *J. Optics and Lasers in Engineering* **45** 1107-1114

[13] Canny J F 1986 A computational approach to edge detection *J. IEEE Trans. Pattern Anal. Machine Intelligence* **8(6)** 679-698
[14] Chygrynsky V V, Efimov Y S and Matveyev I A 2016 Fast algorithm for determining pupil and iris boundaries *J. Machine Learning and Data Analysis* **2(2)** 159-172
[15] MMU Iris Image Database (Access mode: http://pesonna.mmu.edu.my/ccteo/)
[16] Jaccard P 1901 Distribution de la flore alpine dans le Bassin des Dranses et dans quelques regions voisines *J. Bull. Soc. Vaudoise sci. Natur.* **37** 241-272

Acknowledgements
The reported study was funded by RFBR according to the research project no. 18-01-00748 a.

Gait analysis for person recognition using principal component analysis and support vector machines

O V Strukova[1], LV Shiripova[1] and E V Myasnikov[1]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086

Abstract. The paper is devoted to the problem of the recognition of a person by gait using a video recorded in the optical range. The method proposedin this paper consists in the detection of a moving person on a video sequence with the subsequent size normalization and dimensionality reduction using the principal component analysis technique. The person classification was carried out using the support vector machine. The experimental studies performed using the CASIA GAIT dataset allowed us to determine the best values of the method parameters. The obtained results showed that with a small number of classes, high classification accuracy canbe achieved.

1. Introduction

The identification of a person by its biometric parameters is popular and widely used all over the world at present. Such specific features as face image, voice timbre, fingerprints, iris pattern and even gait are used for the identification of a person. Although the use of fingerprints or the iris pattern makes it possible to identify a person with little or no error, contactless and remote identification methods are of considerable interest. In this regard, especially important is the problem of recognizing the person using his gait.

Considering the gait as a set of poses and movements, we can distinguish two most common ways of recording (capturing) such information: video [1] (for example, in the optical range) and recording using sensors located on the human body [2]. In addition, there are papers, in which gait analysis is performed based on the readings of the accelerometers built into the smartphone [3].

Considering that the gait allows to identify the person even in cases where it is not possible to produce it in other ways (the object is at a distance, it is impossible to obtain a high-quality image of the face, etc.), the use of a video, for example, from CCTV cameras is of particular interest.

To date, various methods have been used to solve the problem of the person identification on a video by gait.

The approach used in [4] consists in the subsequent segmentation of the background using the background subtraction algorithm based on the mixture of Gaussian distributions (GMM), dimensionality reduction using the principal component analysis technique (PCA), and classification based on the Fisher linear discriminant analysis (FLDA). Another feature of the work is the combination of the signs of movement with the signs of the trace (footprint) of a person.

The first step of the approach proposed in the paper [5] is an improved background subtraction procedure. In this paper, the selected motions are described by the descriptors based on the form statistical analysis (Procrustes analysis) technique. The procedure of the supervised classification is constructed using the appropriate measure (Procrustes distance measure).

In the paper [1], the analysis of the linear (PCA) and non-linear (ISOMAP, LLE) dimensionality reduction techniques, which are used to form features, is performed. A Hidden Markov Model (HMM) is used to classify the generated features.

In the paper [6], the Support Vector Machine (SVM) is used to solve the problem of classification of a person by motion. In particular, the dependence of the classification accuracy on the type of the SVM kernel is studied in the paper.

In general, it can be noted that the problem of recognizing a person by gait attracts the attention of an increasing number of researchers. At the same time, considerable attention is paid to both the methods of the feature description of motion and the choice of effective classification methods.

In this paper, to solve the problem of identification of a person by gait, we follow the general approach used in the above studies [1, 4, 6]. This method is based on the detection and segmentation of a moving person on a video sequence, normalizing the size of frames and reducing the dimensionality of the sequence using the principal component analysis technique. The support vector machine is used as a classifier. Considerable attention is paid to the selection of parameters of a feature description. The study shows the importance of the careful selection of parameters in the solution of the considered problem. This allows to achievehigh quality of the classification with a relatively small number of classes.

The paper has the following structure. Section 2 is devoted to the description of the method used in the paper. Section 3 describes the results of experiments. The paper ends up with the conclusion. The list of used literature is given at the end of the paper.

2. Methods

The method developed in this paper consists of the following steps:

- detection and segmentation of a moving person in the video sequence,
-normalization of the frame size of the selected video sequence fragment,
-dimensionality reduction of the selected video sequence fragment,
-classification of video sequences.

2.1. Detection and segmentation of a moving person on a video sequence

At the first stage of the developed method the moving person is allocated on the video sequence. When the video sequence source is a video surveillance camera, background subtraction methods are used most frequently. The main idea of the methods of this class is to use a certain background model and to decide whether the particular pixel belongs to the background or a moving object. This decision is based on the correspondence of the pixel to the background model. The background model is gradually refined. Although the time-averaged observation image can be used as a background model in simplest applications, better results in this problem are given by more complex models, for example [7-9].

In this paper, we use the background subtraction algorithm based on the mixture of Gaussian distributions (Gaussian mixture model, GMM) [8]. According to this method, each background pixel is modeled by a weighted sum (mixture) of Gaussians. The weights of Gaussians are determined by time periods, during which the corresponding color is present on the video sequence.

As a result of this stage, the set of masks corresponding to individual frames of the video sequence is formed. Each mask reflects the result of the segmentation of a frame into the foreground area corresponding to a moving person and the background. An example obtained using the selected method is shown in figure 1.

2.2. Normalization of the size of detected fragments

At the second stage of the method, obtained masks are processed as follows. First, the center of mass for each foreground region is calculated. Then the linear dimensions (size) of the region are determined, and a framing (truncation of the mask image) is performed. After that, the cropped image is resized to the specified size. The described scheme is shown in the figure 2.Taking into account the time coordinate, the dimensionality of the sequence of masks, which describes the movement of a

person, remains high even after the size normalization. In this regard, the third stage reduces the dimensionality of data describing the movement of a person.

Figure 1. Detection of a moving person on a video sequence: the frame of the original video sequence (on the left), the formed mask (on the right) [10].

Figure 2. Normalization of the size of detected fragments.

2.3. Dimensionality reduction using the principal component analysis technique

To reduce the dimensionality of multidimensional data, both linear and nonlinear methods are used. The most commonly used are linear methods, such as the principal component analysis (PCA) [11] and independent component analysis (ICA). Nonlinear dimensionality reduction methods (for example, nonlinear mapping, ISOMAP, LLE) [12] are used less often due to the high computational complexity of such methods. It should be noted that recent attempts have been made to accelerate such methods [13, 14].

In this paper, we use the principal component analysis technique, as the most often used in such cases (see, for example, [1, 4]). This method searches for a linear projection into the subspace of a smaller dimension that maximizes the variance of data. The PCA method is often considered as a linear dimensionality reduction technique, minimizing the loss of information.

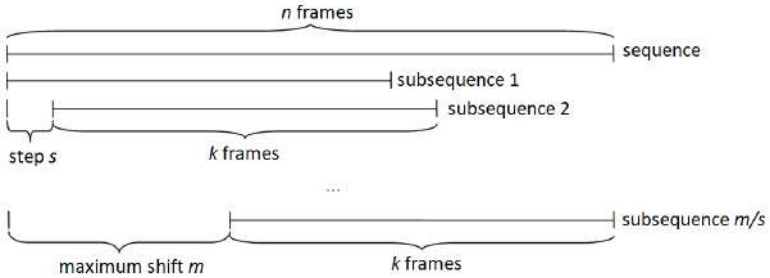


Figure 3. Parameters of the allocated subsequences.

In this paper, before reducing the dimensionality of data, for each sequence of frames we form a set of subsequences of a fixed length. To do this, we successively select subsequences of the predefined length k with the step s starting from the beginning of the whole sequence (see figure 3).

For each selected subsequence, the vector of features is formed as follows: each normalized frame of the subsequence is expanded into a row, and the rows obtained for individual frames are concatenated to each other.

The feature vectors of all sequences for different persons form the input matrix for the principal component analysis technique. When principal components are found, the projection of feature vectors onto the first N principal components is taken as a feature description.

2.4. Classification of video sequences

The features obtained as a result of the principal component analysis are used to train the support vector machine (SVM) [15] classifier. In the considered case, the classes correspond to individual

persons (individuals), and feature vectors obtained for all the subsequences correspond to individual observations (examples).

The description given above is valid for the training mode, in which the parameters of the dimensionality reduction procedure (PCA) and classifier (SVM) are configured. In the testing mode, the data is processed in the same way, except that the parameters of the linear transformation (which is used to reduce the dimensionality) are fixed to the values obtained in the training mode, and the classification is performed by the trained SVM classifier.

3. Experiments

The described above methods were implemented in C ++ using the OpenCV library. A PC based on Intel Core i5-3470 CPU 3.2 GHz was used to perform experimental studies.

For the experimental study, the video sequences from the open CASIA GAIT dataset [16] were used. This dataset contains the sequences of binary images, which contain the silhouettes of moving persons.

In this work, we used sequences of 25 persons, in which the shooting angle is 90 degrees, people are depicted in normal clothes and without bags. There were 6 sequences in each class. The length of each sequence was not less than 60 frames. Classes were divided into training and test samples of 3 sequences each.

To estimate the quality of the considered methods, we used the classification accuracy, defined as the proportion of correctly classified sequences.

In the first experiment, the dependence of the classification accuracy on the maximum shift of the subsequences from the beginning of the sequences was investigated (parameter m in figure 3). In this experiment, to reduce the learning time used in the selection of subsequences, the shift step s was 3 frames. Thus, the maximum step of the subsequences m also changed in step 3, taking values from 0 to 15. The last value was determined from the length k of the generated subsequences (k = 45 frames), and the minimum length of the sequences selected for the experiments (n=60 frames).

The experimental results are shown in figures 4 and 5. It was experimentally determined that the accuracy of the classification as a whole increases with increasing maximum shift. This observation is quite expected, since the greater the maximum shift is used in the formation of characteristics, the more "complete" is the feature description of the video sequence. This growth is also accompanied by an increase in processing time, as the number of processed subsequences increases.

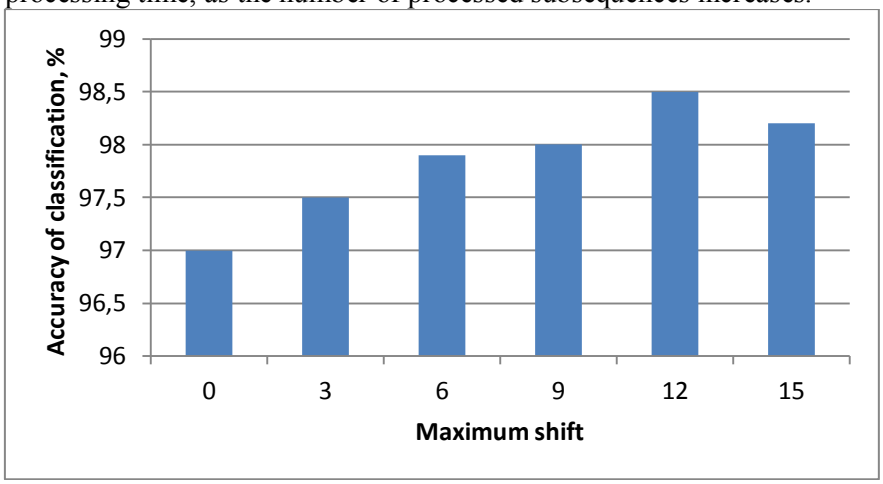


Figure 4. Dependence of the classification accuracy on the maximum shift.

To choose the value of the step s (see figure 3) we considered rather small values from 1 to 4, which allow us to obtain a more “dense” feature description of the video sequence. Our preliminary experiments showed that the best classification accuracy is achieved at small values of the shift s. Taking into account also the computation time, we ended up with the value s=2.For the method proposed in Section 2.3, we studied the dependence of the classification accuracy on the

dimensionality of feature vectors (output dimensionality of the PCA technique). The results of the experiments are shown in figures 6 and 7.

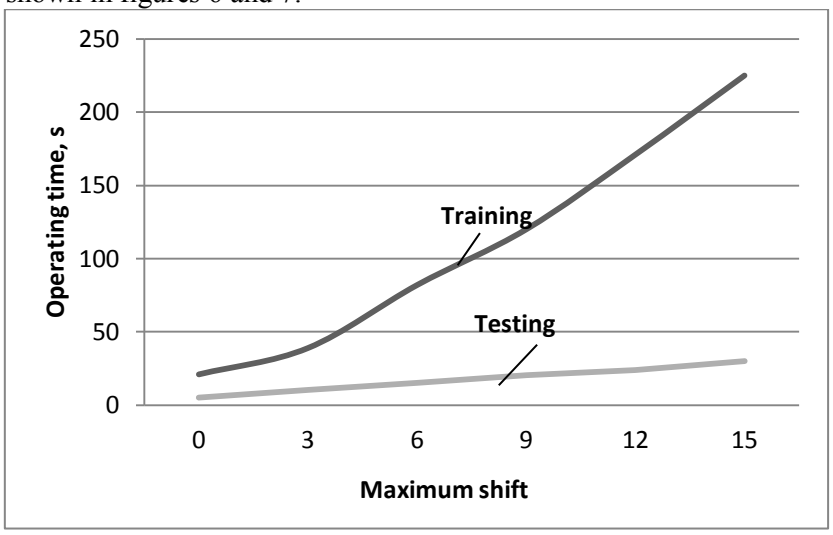


Figure 5. Dependence of training and testing time on the maximum shift.

As it can be seen from the figure, the best values of the classification accuracy are achieved for 64-dimensional feature vectors. The increase in dimensionality is accompanied by the expected increase in processing time, although the changes are not very significant.

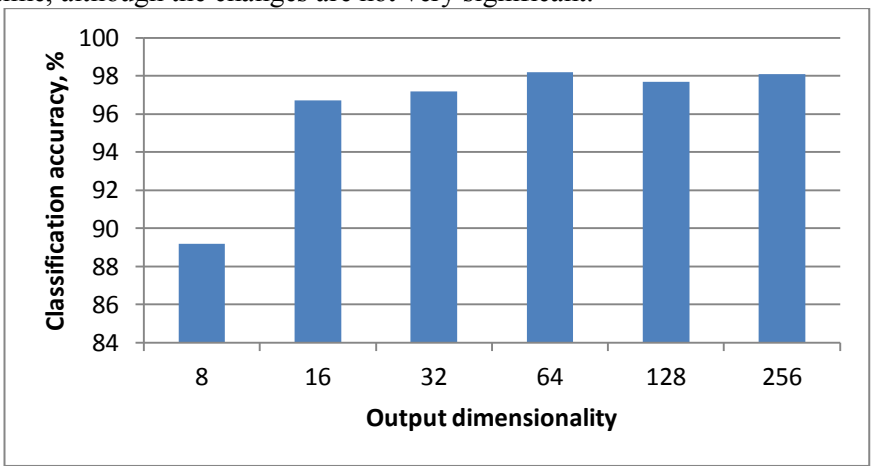


Figure 6. Dependence of the classification accuracy on the dimensionality.

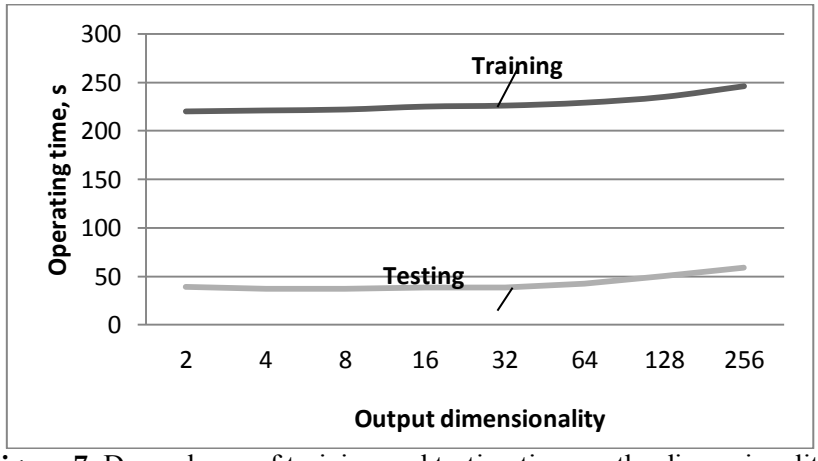


Figure 7. Dependence of training and testing time on the dimensionality.

In the next experiment, we considered the dependence of the classification accuracy on the number of classes (persons). The experiment was carried out for 5, 10, 15, 20 and 25 classes, and other parameters remained fixed. In particular, the step *s* was equal to 2 frames, the maximum shift *m* of the beginning of the extracted subsequences was equal to 15 frames, and the dimensionality of feature vectors was equal to 64. The results of the experiment are shown in the figures 8 and 9.

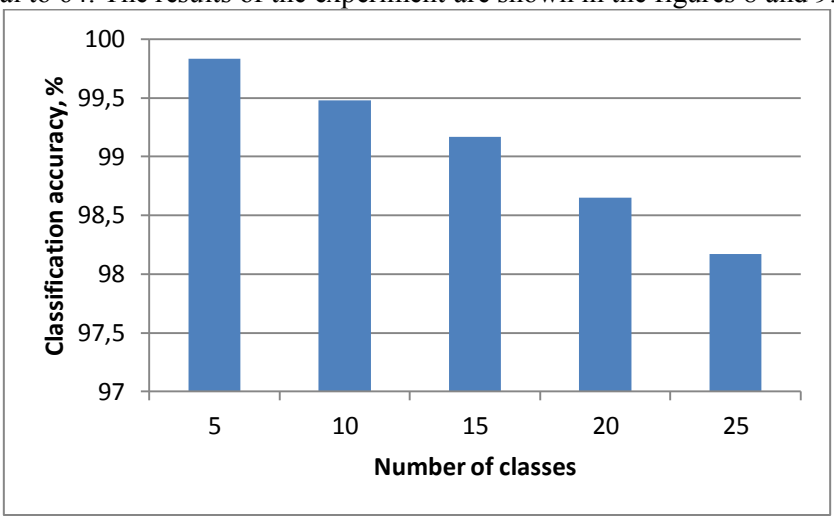


Figure 8. Dependence of the classification accuracy on the number of classes.

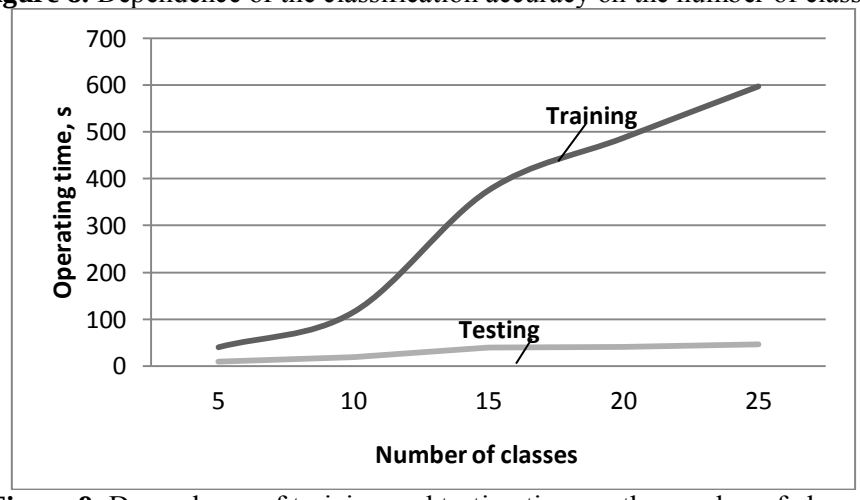


Figure 9. Dependence of training and testing time on the number of classes.

It is worth noting that a direct experimental comparison to other works seems to be quite a challenge in connection with the different data sets used, as well as the potential differences in the experimental conditions. The closest approach to the proposed one is described in the paper [6]. Depending on the classifier configuration, the authors in [6] declared the accuracy from 92.08 to 98.79% for the case with ten objects.

Thus, it can be said that the results obtained in this paper correspond to the current state in the considered field of research.

As it can be seen in the figure 9, the processing takes an increasing amount of time as the number of classes increases. Considerable time is taken in the training mode. This fact becomes especially important in scenarios when the number of classes changes dynamically and it is required to re-train the system regularly.

4. Conclusion

In this paper we proposed the method for human identification by gait. The proposed method consists in the detection of a moving person on a video sequence with the subsequent normalization of size,

generation of subsequences, dimensionality reduction using the principal component analysis technique, and classification using the support vector machine.

The experiments performed on the CASIA GAIT dataset allowed to determine the best values of the parameters of the proposed method.

The drawbacks of the proposed method include its long operating time. In connection with this, a promising line of research is speeding up this method. Another possible direction of further research is the recognition of human actions and the detection of abnormal behavior (see, for example, [17, 18]).

5. References

[1] Josiński H, Świtoński A, Michalczuk A, Kostrzewa D and Wojciechowski K 2013 Feature Extraction and HMM-Based Classification of Gait Video Sequences for the Purpose of Human Identification *Vision Based Systems for UAV Applications. Studies in Computational Intelligence* **481** 233-245

[2] Suutala J, Fujinami K and Röning J 2008 Gaussian Process Person Identifier Based on Simple Floor Sensors *European Conference on Smart Sensing and Context* 55-68

[3] Dingbo D, Guangyu G, Chi H and Jian Ma 2014 Automatic Person Identification in Camera Video by Motion Correlation *Journal of Sensors* **2014** 838751

[4] Murukesh C, Thanushkodi K, Padmanabhan P and Feroze Naina M D 2014 Secured Authentication through Integration of Gait and Footprint for Human Identification *J. of Electrical Engineering and Technology* **9(6)** 2118-2125

[5] Wang L, Tan T, Hu W and Ning H 2003 Automatic Gait Recognition Based on Statistical Shape Analysis *IEEE Transactions on Image Processing* **12(9)** 1120-1131

[6] Shelke P B and Deshmukh P R 2014 Person Identification Using Gait: SVM Classifier Approach *International Journal of Emerging Technologies and Engineering (IJETE)* **1(10)**

[7] Kadew T K P and Bowden R 2002 An improved adaptive background mixture model for real-time tracking with shadow detection *Video-Based Surveillance Systems* 135-144

[8] Zivkovic Z 2004 Improved adaptive Gausian mixture model for background subtraction *Proc.of the 17th Int. Conf. on Pattern Recognition* **2** 28-31

[9] Godbehere A B, Matsukawa A and Goldberg K 2012 Visual Tracking of Human Visitors under Variable-Lighting Conditions for a Responsive Audio Art Installation *American Control Conference*

[10] *Background Subtraction* (Access mode: https://docs.opencv.org/3.3.0/db/d5c/tutorial_py_bg_subtraction.html) (13.11.2017).

[11] Fukunaga K 2003 *Introduction to Statistical Pattern Recognition* (London: Academic Press)

[12] Lee J A and Verleysen M 2007 *Nonlinear Dimensionality Reduction* (New York: Springer-Verlag)

[13] Myasnikov E V 2017 Fast Techniques for Nonlinear Mapping of Hyperspectral Data *Proc. SPIE* **10341** 103411D

[14] Myasnikov E V 2016 The use of Interpolation Methods for Nonlinear Mapping *Lecture Notes in Computer Science* **9972** 649-655

[15] Cortes C and Vapnik V 1995 Support-vector networks *Machine Learning* **20(3)** 273-297

[16] *Database CASIA GAIT* (Access maode: http://www.cbsr.ia.ac.cn/english/Databases.asp) (13.11.2017)

[17] Morozov A A and Sushkova O S 2016 Real-time analysis of video by means of the Actor Prolog language *Computer Optics* **40(6)** 947-957 DOI: 10.18287/2412-6179-2016-40-6-947-957

[18] Shatalin R A, Fidelman V R and Ovchinnikov P E 2017 Abnormal behavior detection method for video surveillance applications *Computer Optics* **41(1)** 37-45 DOI: 10.18287/2412-6179-2017-41-1-37-45

Acknowledgments

The reported study was funded by RFBR according to the research project №17-29-03190.

Detection of malignant breast tumors on the background of fibrocystic breast disease

Yu A Podgornova[1] and S S Sadykov[1]

[1]Murom Institute (branch) Federal State Budgetary Educational Institution of Higher Education Vladimir State University named after Alexander Grigoryevich and Nickolay Grigoryevich Stoletovs, Orlovskaya Street 23, Murom, Russia, 602264

Abstract. The incidence of breast cancer is increasing every year. The mammography screening is the only method of diagnosing the disease. The mammograms analysis is complicated by the existence of various types of mastopathy - the most common breast disease caused by changes in the woman's endocrine profile. More than 40 per cent of the female population is exposed to this disease according to the WHO statistics. An algorithm for the detection of malignant neoplasms was developed by the authors in the context of fibrosis-cystic mastopathy. The experimental research on mammography images from the MIAS database showed that malignant tumors were available in 80 per cent of cases (21 mammograms), in other cases (5 mammograms) the tumor area was vaguely separated from the surrounding tissue.

1. Introduction

In compliance with the statistical data provided by the Russian oncologists, every 5 woman is affected by breast cancer (hereinafter BC) which is, on average, the cause of 16,000 fatal cases per annum [1, 2]. At the early stages breast cancer is virtually impossible to reveal as it is not generally accompanied by any pain. In addition, it is a high risk of missing the diseases signs at mammography screening. One more reason of such missing is the presence of fibrosis-cystic mastopathy (hereinafter referred to as FCM) in a breast. FCM is a group of heterogeneous diseases characterized by a wide range of proliferative and retrograde changes in the breast tissue with a violated ratio of epithelium and connective tissue components.

The radiographic appearance of this mastopathy form is characterized by a mass calcification of a glandular triangle and an abnormal structural pattern. Instead of a usual radiary streaking from the breast base towards a nipple one can observe a chaotic and rough trabecular structure.

Diffuse FCM forms are of interest for research not only from the point of view of a significant incidence of this pathology but mainly because they provide a background for breast cancer development. In the opinion of many authors, malignant changes at non-proliferative FCM forms is 0.9%, at moderately intense proliferation - 2%, at strong proliferation - 31% [3].

Therefore, various mastopathy forms detected in the image complicate the neoplasm diagnostics, especially at the early stages of cancer when the size of a tumor is not more than 1 cm. Applying contemporary information technologies to process mammograms, both digital and digitized, helps to improve the image quality, highlight and magnify the section of interest, automatically highlight the areas of malignant and benign neoplasms. Therefore, as of today there is a problem consisting in urgent necessity to enhance the quality and plausibility of breast cancer detection on the background

of mastopathy at analog mammography images and the absence of contemporary technology to implement these functions in various medical institutions. In view of the foregoing, improving the quality and plausibility of breast cancer detection with the help of contemporary ITs is a relevant applied scientific and technical problem.

The purpose of the paper is to develop an algorithm for breast cancer revealing on the background of FCM allowing for the visualization of malignant tumours to find their exact location and necessary texture characteristics.

2. Algorithm for the detection of cancer tumors on mammography images in the FCM background

Figure 1 shows the proposed algorithm to detect malignant tumors affected by FCM.

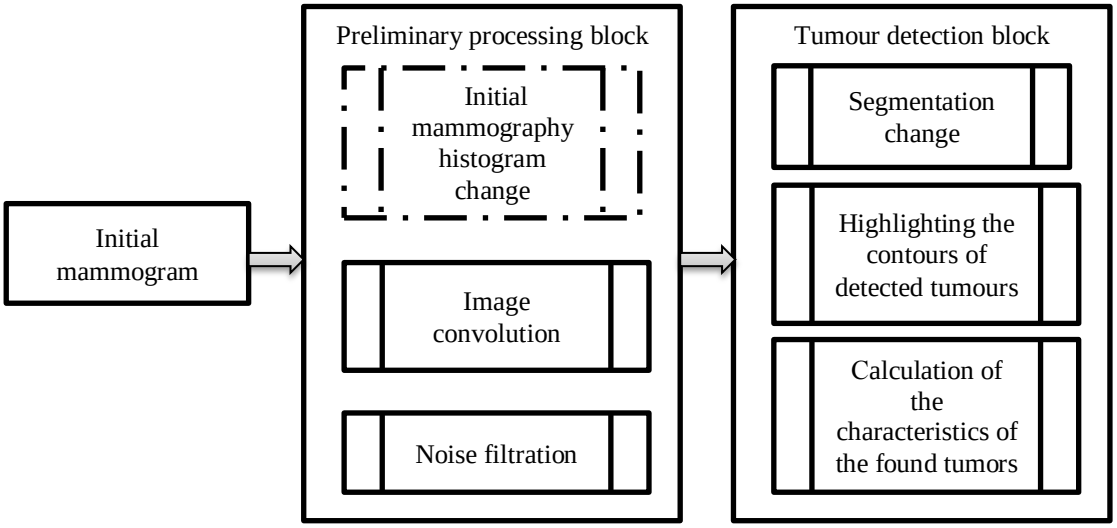


Figure 1. General block diagram for the algorithm of highlighting the breast cancer area on the FCM background.

The initial mammography image is a half-tone image *I(x,y)* digitized by a high-precision scanner and having the dimension of $m \times n$ dots. The pixel brightness gradation is distributed within the range [0. 255].

Fibrous-cystic calcification areas present at mammography images significantly complicate diagnostics as it is impossible to visually detect the neoplasm lesions located under FCM. In this case a radiologist cannot state a plausible diagnosis to a patient. The example of the specified radiogram is a mammography image in Figure 3(б). The authors developed an algorithm for radiographic opacification permitting a specialist to remove a fibrous-cystic calcification layer from the image.

The method is based on historgram transformations [4, 5] of a radiographic image which can be represented as follows:

$$F(x,y) = R \cdot t\left[I(x,y)\right] \cdot (m \cdot n)^{-1} - t[0] + \left(F_{\max} - F_{\min}\right) \cdot \left(I(x,y) - I_{sr}\right) \cdot \left(I_{sr} - I_{\min}\right)^{-1} \quad (1)$$

where *I(x,y)* – current brightness value of the initial image, $I_{\max}$, $I_{\min}$, I_{sr}- maximum, minimum and average values of the initial mammography values, correspondingly, *F(x,y)* – brightness value of the output image dot, $F_{\max}$, $F_{\min}$ - required maximum and minimum brightness values of the output image dots, correspondingly, R = 1.5 – coefficient (1<=*R*=>2).

$$t[i] = \sum_{j=0}^{i} \left(v[j] \cdot (m \cdot n)^{-1}\right), \quad (2)$$

where *i* –brightness gradation value of the output image dot (*i*=0..255), *j* – brightness gradation value of the initial image dot (*j*=0..255, *j*<=*i*,), *v*[*j*] –value of the initial image histogram element at the brightness *j* ($v[j] \in V$), *T* – area of the output image histogram element change *F(x,y)*, *t*[*i*] – value of

the output image histogram element at the brightness i ($t[i] \in T$), $t[0] = t[I(x,y)=0]$ – value of the processed image histogram element at the point brightness equal to , $m \times n$ – image dimension.

If we consider a mammography image showing a breast in the form of a fat involution (Figure 3(a)), the image opacification is not necessary.

The next stage of a preliminary processing is the image mask filtration.

$$d(x,y) = \sum_{k,l} F(x-k, y-l) \cdot g(k,l), \tag{3}$$

where $d(x,y)$ – brightness value of the output image dot, k, l – mask dimensions (k=0..2, l=0..2), $F(x,y)$ – brightness value of the initial image dot, $g(k,l)$ – mask element.

Multiple research have shown the areas of malignant neoplasms can be highlighted with the help of a convolution of a mammography image using the mask given in Figure 2.

2	-1	2
1	-3	1
2	1	2

Figure 2. Mask for mammography convolution.

Such transformation of a mammography image causes some noise, that is why a median filtering is applied with the help of the mask 3×3 [6].

$$r(x,y) = M\left[d(x,y)\right] \tag{4}$$

where $r(x,y)$ – output image, M – median filtering operator.

The contour extraction block in the image locates a necessary area with clear contours.

Before extracting the contours in the image, the image texture map is subject to binarization. The binarization algorithm is presented as follows:

$$b(x,y) = \begin{cases} 1 & \text{если} \quad r(x,y) \geq Q \\ 0 & \end{cases}, \tag{5}$$

where $b(x,y)$ – output image, Q – threshold value equal to the maximum brightness value of the image dots $r(x,y)$,

$$Q = \max\left[r(x,y)\right]. \tag{6}$$

Contour extraction is implemented with the help of the algorithm from [5].

$$k(x,y) = L\left[b(x,y)\right], \tag{7}$$

where $r(x,y)$ – output image, L– contour extraction operator.

The block of characteristics identification [7, 8] calculates the following:

- breast cancer area square $S_к$ (number of dots in the area);
- area perimeter P – approximately equal to the number of the area contour dots;
- form coefficient: $K_f = \dfrac{P^2}{S_k}$

and other geometrical and texture characteristics of malignant neoplasms.

3. Experimental research of the proposed algorithm

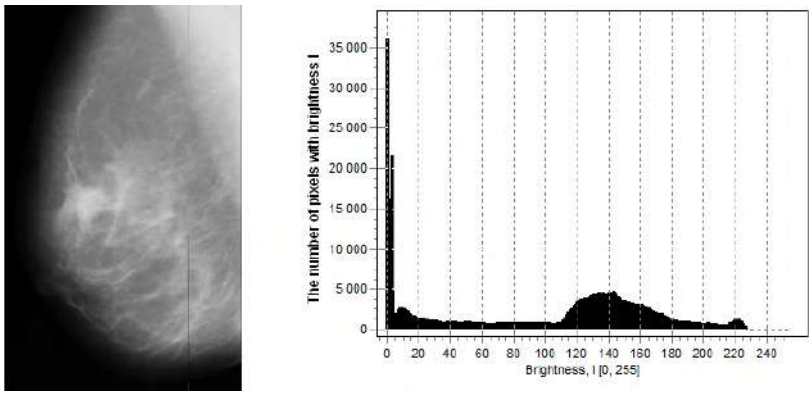
The authors chose the mammography database MIAS for the experiment as all its images have a confirmed diagnosis as well as a clear description of each image with the specification of the tumour location. Figure 3 gives two initial mammography images.

In the radiographic image (Figure 3 (a)) in the breast upper quadrant projection one can identify a lesion with spicular contours, the skin over the them is thickened, the nipple is inverted. Conclusion: breast cancer.

Figure 3(b) shows that, on the background of breast tissue fat involution, one can identify a moderately expressed FCM with predominant fibrosis. In the upper quadrant one can suspect a lesion

presence. At the description of the screening mammography image there is a high possibility to miss the pathologic signs.

a)



b)

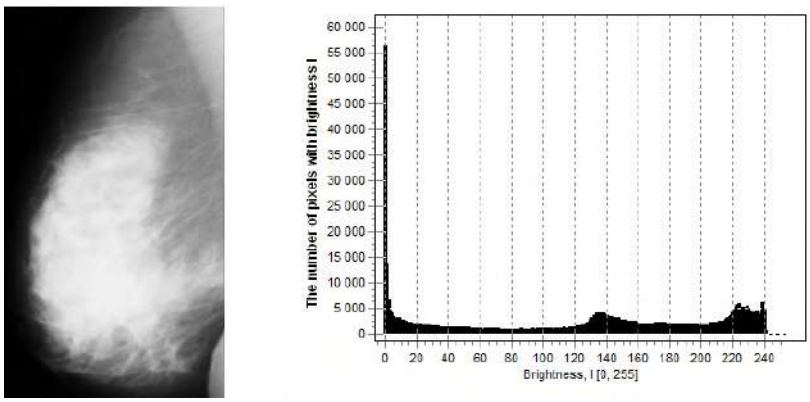
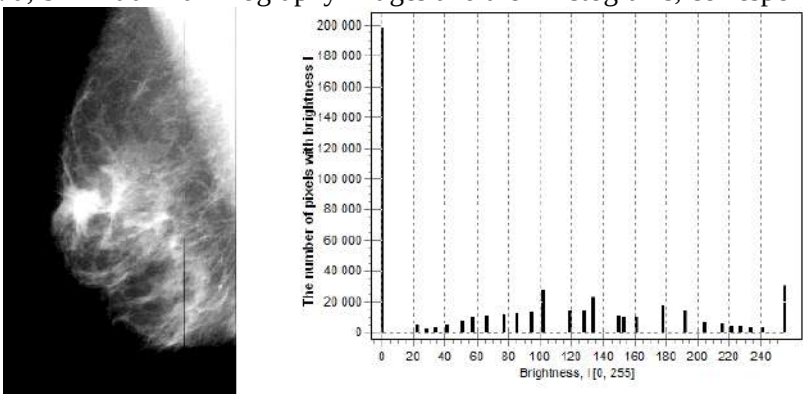


Figure 3. a, b – initial mammography images and their histograms, correspondingly.

a)



b)

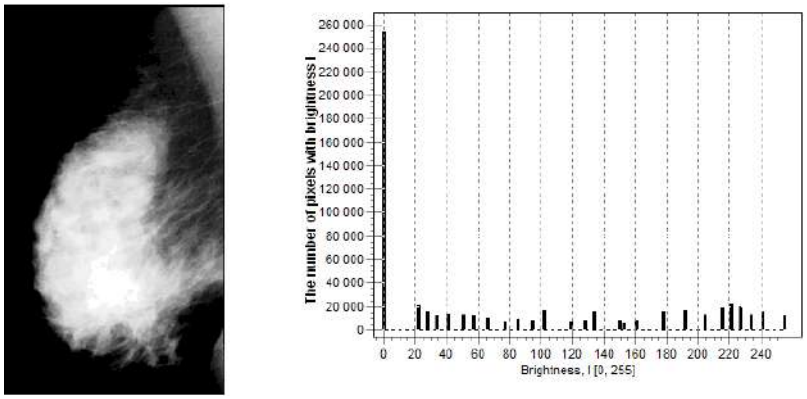


Figure 4. The results of the proposed histogram transformation: a) proposed algorithm of the histogram transformation for the Figure 3(a), b) proposed algorithm of the histogram transformation for the Figure 3(b).

See Figure 3: in the image processing using the proposed opacification method, the neoplasm described above is better visualized due to a larger difference of neoplasm and surrounding tissue densities.

At the opacification of Figure 3(b) one can see a high-density lesion with “blurred” contours, suspected malignant neoplasm. The next processing stage is the transformed image convolution with the help of the mask (Figure 3).

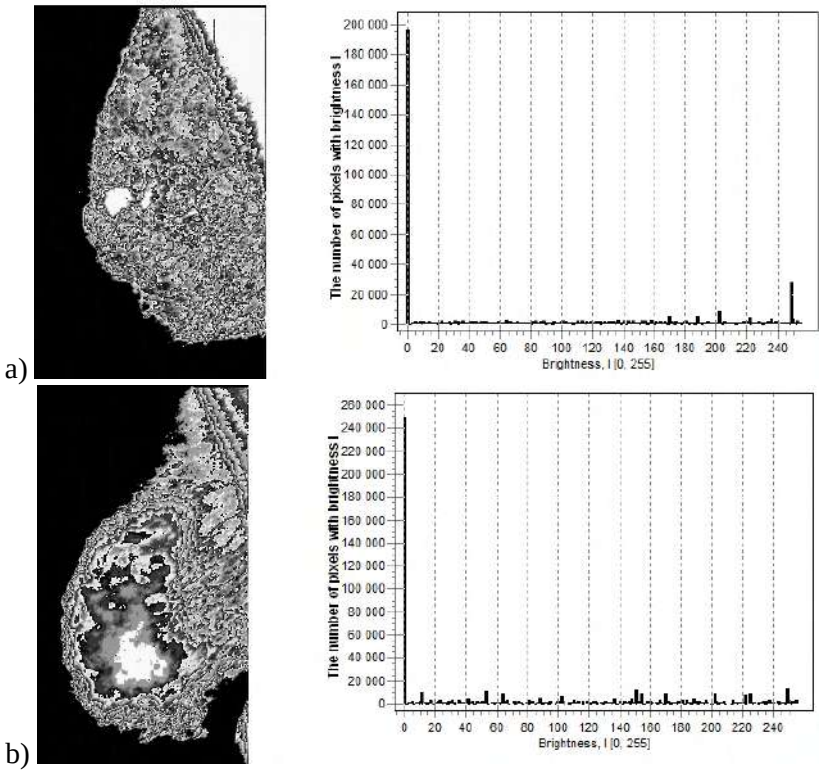


Figure 5. The image convolution result: a) image convolution in Figure 4(a), b) image convolution in Figure 4(b).

Applying the mask convolution method for the image in Figure 4(a), we obtain highlighted lesions on the breast tissues’ background, and small lesions are the tumor “seedings” poorly seen in the analog mammography image.

When processed with the help of the proposed method, this neoplasm (in Figure 5(b)) is highlighted on the background of surrounding tissues which points to a malignant neoplasm with high probability. Such characteristics as maximum and minimum brightness values, mean square deviation and entropy were calculated for all images and united into the Tables 1 and 2 for the purpose of comparison.

The values provided in the tables demonstrate the proposed algorithm of histogram transformations significantly improves the initial images as it increases the MSD value and decreases the entropy value of both initial images.

Table 1. Summary table of image characteristics for the image No. 1.

Image	Minimum brightness	Maximum brightness	Average brightness value	Mean square deviation	Entropy
Initial image	0	240	97	68.60	7.11
Hist. transf.	0	255	80	84.5	3.51
Convolution	0	255	83	91.3	5.21
Filter.	0	249	83	82.7	4.91

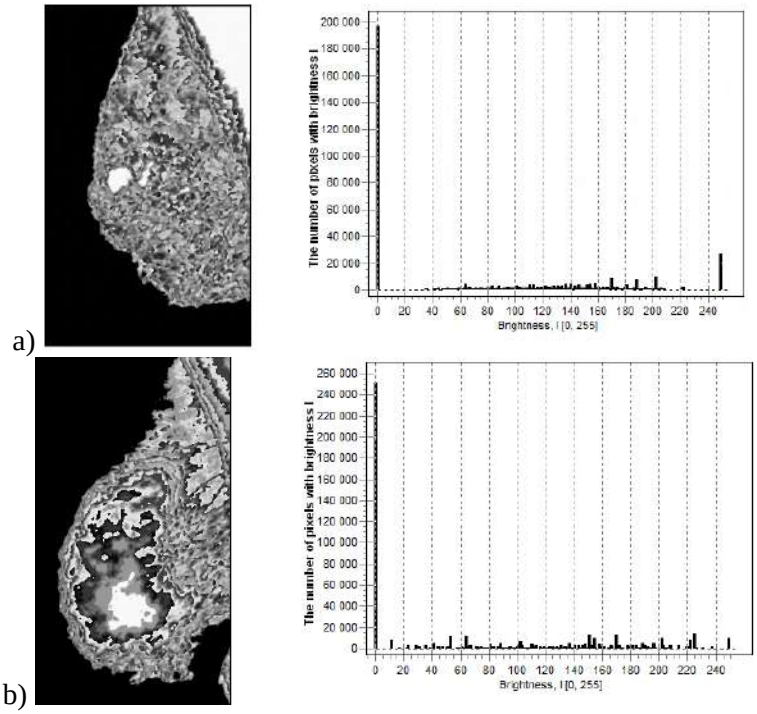


Figure 6. The image median filtration result: a) image median filtration in Figure 5(a), b) image median filtration in Figure 5(b).

Table 2. Summary table of image characteristics for the image No. 2.

Image	Minimum brightness	Maximum brightness	Average brightness value	Mean square deviation	Entropy
Initial image	0	241	122	83.4	7.38
Hist. transf.	0	255	79	90.5	3.6
Convolution	0	255	77	87.3	5.05
Filter	0	249	77	81.4	4.82

Further, we apply binarization to the filtered image in Figure 6(a) and Figure 6(b) and extract the contours of neoplasm highlighted areas.

4. Conclusion

Summing up the aforesaid, one can draw the following conclusions on the research:

1. The proposed algorithm has been used to process 43 mammography images including: 26 images with previously detected and histologically confirmed cancer, 17 mammography images with detected changes assessed as benign neoplasms. At the mammogram processing the authors obtained:

- areas with malignant neoplasms are identified in 80% of cases (21 mammography images), in other cases (5 mammography images) the neoplasm area was vaguely separated from surrounding tissues;
- in 15 cases of 17 ones the lesion is not detected by the proposed algorithm. In 2 mammography images the lesion was highlighted, the interpretation of these results needs additional research conduct.

2. The proposed algorithm for the highlight of a breast cancer area in a mammography image allows detecting the breast cancer area even at the presence of a strong shade from FCM.

3. A set of algorithms for histogram filtering, median filtration, texture segmentation, binarization and contour highlight is presented in the form of the algorithm for highlighting the breast cancer area in mammography images with FCM and corresponds to all requirements of radiologists.

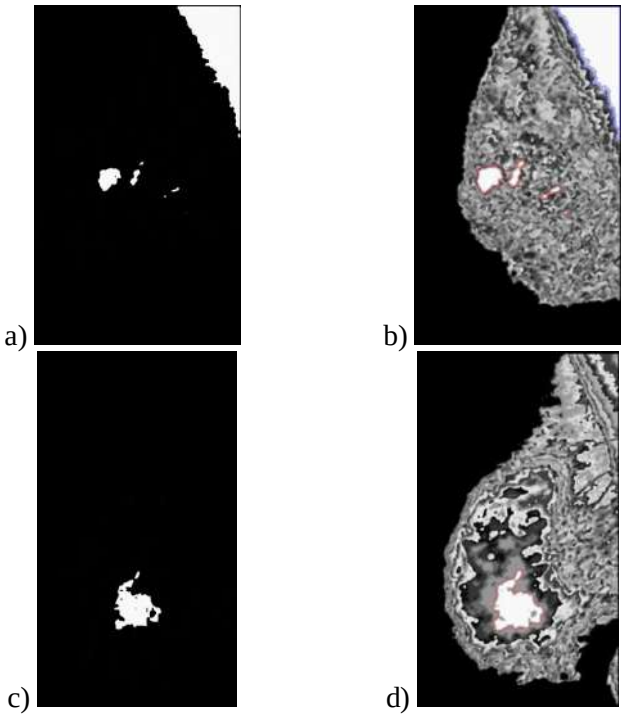


Figure 7. a, c – Image binarization in Figure 6(a) and (b) correspondingly, b, d - Contour application in the image in Figures 6(a), (b), correspondingly.

4. The further application of the proposed algorithm with breast cancer area highlighting on the FCM background provides for the opportunity to automatically detect breast cancer at early stages even at the presence of a severe FCM in a breast.

5. References

[1] Sadykov S S, Bulanova Yu A and Zakharova Ye A 2014 Computer Diagnostics of Neoplasms in Mammography Images *Computer Optics* **38(1)** 131-138
[2] Kaprin A D, Starinskiy V V and Petrovaya G V 2015 *Cancer Care Situation in Russia in 2014* (Moscow: Herzen Moscow Oncology Research Institute) p 236
[3] Korzhenkova G P 2004 *Comprehensive Radiological and Sonographic Diagnostics of Breast Diseases* (Moscow: Firma "STROM", LLC) p 128
[4] Sankar P R and Srinivasa B K N 2007 Parallel Architecture for Implementation of Contrast Limited Adaptive Histogram Equalization *Advanced Engineering Sciences and Technologies* **10(1)** 047-051
[5] Rajesh K R and Puran G 2012 Balvant Singh Underwater Image Segmentation using CLAHE Enhancement and Thresholding *Emerging Technology and Advanced Engineering* **2** 118-123
[6] Pratt U 1982 *Digital Image Processing* (Moscow: Mir)
[7] Gaidel A V and Krasheninnikov V R 2016 Feature selection for diagnosing the osteoporosis by femoral neck x-ray images *Computer Optics* **40(6)** 939-946 DOI: 10.18287/2412-6179-2016-40-6-939-946
[8] Gaidel A V 2016 Matched polynomial features for the analysis of grayscale biomedical images *Computer Optics* **40(2)** 232-239 DOI: 10.18287/2412-6179-2016-40-2-232-239

Acknowledgments

The research was conducted at the financial support of the RFBR in the framework of a scientific project No. 16-37-00227. Podgornova Yu.A. expresses her sincere gratitude to her academic advisor, Dr. of Tech. Sci,. Professor Sadykov S.S. for his help at the research conduct, valuable recommendations in relation to their planning and article preparation as well as for his moral support.

Detecting forgery in image time series based on anomaly detection

N I Evdokimova[1] and V V Myasnikov[1,2]

[1]Samara National Research University, Moskovskoe shosse 34, Samara, Russia, 443086
[2]Image Processing Systems Institute - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001

Abstract. Increasing complexity of image forgery methods is an actual problem nowadays. This problem rises due to the expansion of fields that use digital images in their work. Image time series show the dynamics of the scene and allow it to be compared over time. This paper proposes a new algorithm for detecting forgeries of the single digital image in an image time series described a scene. This algorithm uses analysis of errors set that were computed during reconstruction of the analyzed image using other images of series. The first part of the paper describes the proposed algorithm consisted of three stages. The second part of the paper describes forgery detection using morphological image filtering based on guided contrasting. The third part of the paper contains comparison of considered algorithms and investigation results of intra-image copy-move and inter-image copy-move detection.

1. Introduction

Image time series describe dynamic of an scene. Analysis of an image time series lets modeling an image that can be next in the image time series. Also, it allows deciding authenticity of the image. There are several approaches to image forgery detection. These approaches may use unique artifacts left by the camera, unique artifacts arising after compression and, finally, temporal and spatial correlations [1]. Methods that used temporal and spatial correlation can divide into two categories. Techniques from the first category based on analysis of images pixel data [5], [6], [7] whereas methods from the second category use object level of images [8].

Forgeries may be created to add a new object to the scene or to hide any existing. Image time series forgery detection has its distinctive features compared to image matching. Every image of an image time series is obtained at different moments of time. Two adjacent images of an image time series can be captured under different conditions of illumination, weather or seasonal conditions. In this paper, the algorithm invariant to the conditions for obtaining images of the series is proposed.

The proposed algorithm uses a correlation between corresponding fragments of neighboring images in the series. In this paper, the concept of anomaly applies for image series forgery detection. In the global sense, an anomaly is a fragment of data that does not correspond to the precisely defined concept of normal behavior [2]. In the sense of this paper, all fragments that are marked as anomaly are considered forgeries.

This work consists of three parts. In the first part, an algorithm for image time series forgery detection based on the anomalies detection are presented. An algorithm of forgery detection using morphological image filtering based on guided contrasting [3] is given in the second part. The third part contains the comparison of both algorithms presented in the first part and the second part.

2. Forgery detection based on the anomalies detection

Let there is an image time series $I_t(n_1, n_2)$, t - image number in time series ($t = \overline{0,T}$, $T \geq 1$). Every image has the same size $N_1 \times N_2$ ($n_1 \in [0, N_1]$, $n_2 \in [0, N_2]$) and captures the same scene at different moments of time.

For definiteness, it is assumed that the image $I_0(n_1, n_2)$ is checked for forgeries although it may be located in the image time series anywhere. The fragments $I_t(m_1, m_2)$ of all images are analyzed in the sliding window $D(n_1, n_2) \subseteq \overline{0, N_1 - 1} \times \overline{0, N_2 - 1}$, $(m_1, m_2) \in D(n_1, n_2)$.

2.1. Image fragment description

The fragments $I_t(m_1, m_2)$ are described in k steps, $k = 2^p$, $p > 2$.

On the $k = 1$ step, the fragment $I_0(m_1, m_2)$ is reconstructed by linear combination of corresponding fragments $I_1(m_1, m_2), ...I_T(m_1, m_2)$ for all possible positions of sliding window D:

$$I_0(m_1, m_2) \approx \sum_{t=1}^{T} \alpha_t I_t(m_1, m_2) \tag{1}$$

using mean squared deviation ε_1^2 minimization:

$$\varepsilon_1^2 = \frac{1}{|D|} \sum_{(m_1,m_2)\in D} (I_0(m_1, m_2) - \sum_{t=1}^{T} \alpha_t I_t(m_1, m_2))^2 \to \min_{\alpha_1,...,\alpha_T}. \tag{2}$$

The last action on this step is calculating both types of errors the mean squared deviation and the normalized mean squared deviation that defined by:

$$\tilde{\varepsilon}_1^2 = \frac{\varepsilon_1^2}{\sum_{i=0}^{N_1-1} \sum_{j=0}^{N_2-1} I_0(i, j)^2}. \tag{3}$$

On the $k = 2^p$ step, every fragment $I_t(m_1, m_2), t = \overline{1,T}$ corresponding to window D location is splitted into k fragments using k-means clusterization (by brightness) as shown in the Figure 1. New fragments that was constructed after clusterization can be denoted by $I_t^j(n_1, n_2)$, $j = \overline{0, k-1}$. Then fragment $I_0(m_1, m_2)$ is reconstructed by linear combination of fragments $I_t^j(m_1, m_2)$:

$$I_0 \approx \sum_{t=1}^{T} \sum_{j=0}^{k-1} \alpha_t^j I_t^j \tag{4}$$

using mean squared deviation ε_k^2 minimization:

$$\varepsilon_k^2 \cong \frac{1}{|D|} \sum_{(m_1,m_2)\in D} \left(I_0(m_1, m_2) - \sum_{1\leq t\leq T} \sum_{0\leq j\leq k-1} \alpha_t^j I_t^j(m_1, m_2) \right)^2 \to \min_{\alpha_1^0,...,\alpha_1^{k-1},...,\alpha_T^0,...,\alpha_T^{k-1}}. \tag{5}$$

Then we calculate both types of errors the mean squared deviation and the normalized mean squared deviation similar to (3):

$$\tilde{\varepsilon}_k^2 = \frac{\varepsilon_k^2}{\sum_{i=0}^{N_1-1}\sum_{j=0}^{N_2-1} I_0(i,j)^2}. \tag{6}$$

In the paper, this procedure performs for $k = 4, 8, 16$, so there is a set of three mean squared

$$\begin{bmatrix} 3 & 5 & 12 \\ 1 & 14 & 2 \\ 4 & 9 & 17 \end{bmatrix} = \begin{bmatrix} 3 & 5 & 0 \\ 1 & 0 & 2 \\ 4 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 9 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 12 \\ 0 & 14 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 17 \end{bmatrix}$$

Figure 1. Splitting of the fragment $I_t(m_1, m_2)$ into $k = 2^2$ fragments.

deviation values and three normalized mean squared deviation values for every position of the window D. Calculated values presents as $\bar{x}(n_1, n_2)$ as follows:

$$\bar{x}(n_1, n_2) \equiv \left(\tilde{\varepsilon}_4^2(n_1, n_2), \tilde{\varepsilon}_8^2(n_1, n_2), \tilde{\varepsilon}_{16}^2(n_1, n_2)\right)^T \tag{7}$$

to form the image fragments feature vector. Mean squared deviation values ε_k^2 are not deliberately taken into account because they are directly used in the calculation of $\tilde{\varepsilon}_k^2$.

2.2. Statistic construction method

The obtained vectors $\bar{x}(n_1, n_2)$ set represents in the coordinate system $\tilde{\varepsilon}_4^2\tilde{\varepsilon}_8^2\tilde{\varepsilon}_{16}^2$. This set locates in the three-dimensional cube with sides equal to 1 as shown in Figure 2:

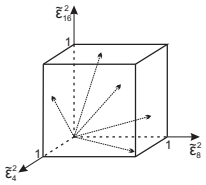


Figure 2. The set of feature vectors in the coordinate system $\tilde{\varepsilon}_4^2\tilde{\varepsilon}_8^2\tilde{\varepsilon}_{16}^2$.

2.3. Anomalies determination

There are no absolute static objects on images obtained in real conditions. This is due both to noises of real cameras and the image compression on the path from the camera to the processing system. It often leads to additional system distortions. Moreover, a scene may contain objects that have specific dynamic characteristics although they are static in the global sense. For example, it may be trees swaying in the wind.

As described above, it can be concluded that it is impossible to obtain a feature vector with coordinates $(0; 0; 0)$ after authentic image fragment representation by described above method. It lets define a rule for assigning fragments corresponding to feature vectors $(0; 0; 0)$ to anomalies. This type of anomalies refers to fragments that were copied from one or several images of the image time series.

On the other hand, the errors $\tilde{\varepsilon}_4^2$, $\tilde{\varepsilon}_8^2$, $\tilde{\varepsilon}_{16}^2$ of an authentic fragment representation must have values that do not exceed a certain threshold. Feature vectors not corresponded to this condition

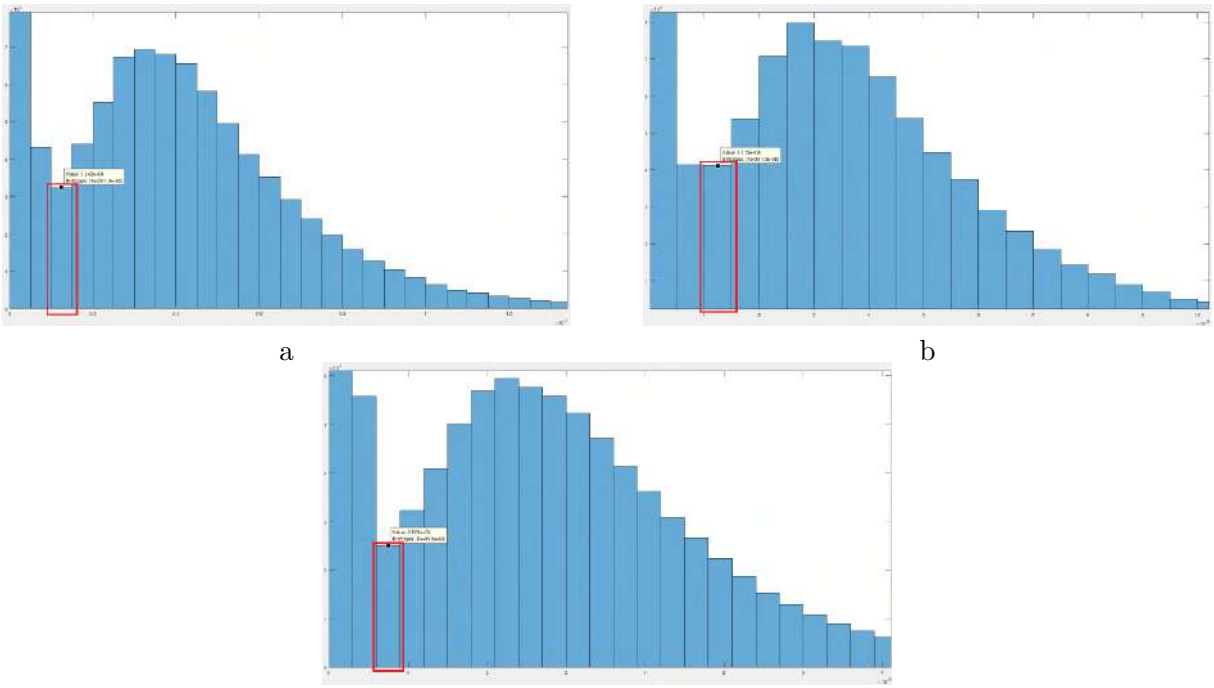
are considered anomalies. This type of anomalies refers to fragments that were copied from the same image or image not included in the image time series.

It is obvious the error value of the same fragment representation decreases with the clusters number increasing. Therefore, it is justified to use different thresholds for $\tilde{\varepsilon}_4^2$, $\tilde{\varepsilon}_8^2$ and $\tilde{\varepsilon}_{16}^2$. The following relation should meet:

$$T_{\tilde{\varepsilon}_4^2} \geq T_{\tilde{\varepsilon}_8^2} \geq T_{\tilde{\varepsilon}_{16}^2} \tag{8}$$

Threshold values (8) are defined by analysis the distribution histograms as shown in Figure (3). The first local histogram minimum is selected and considered a threshold value. So, following threshold values were chosen for histograms shown in the Figure 3:

- $T_{\tilde{\varepsilon}_4^2} = 1.5 \times 10^{-8}$,
- $T_{\tilde{\varepsilon}_8^2} = 1.5 \times 10^{-8}$,
- $T_{\tilde{\varepsilon}_{16}^2} = 0.9 \times 10^{-8}$.



a b c

Figure 3. Selecting the thresholds according to the distribution histogram: a) for $\tilde{\varepsilon}_4^2$, b) for $\tilde{\varepsilon}_8^2$, for $\tilde{\varepsilon}_{16}^2$.

The cube with the feature vectors $\bar{x}(n_1, n_2)$ set (2) is divided into three areas:

1) Origin of the coordinate system;
2) A parallelepiped that is adjacent to the origin;
3) Rest area of the cube.

Per the above, feature vectors from the first area correspond to fragments that were copied from one or several images of the image time series. Feature vectors from the second area refer to authentic image regions. Feature vectors from the third area correspond to fragments that were copied within one image or from an image not included in the image time series. This splitting is shown in Figure 4.

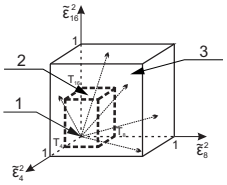


Figure 4. Splitting the cube with set of vectors $\bar{x}(n_1, n_2)$ into three areas.

After extraction of feature vectors from the relevant area and labeling them as suspicious, the corresponded binary mask is created. Then the mask is processed with a noise filter that removes regions with square less than some value. After this, only feature vectors that correspond to forgery regions are kept in the set of suspicious feature vectors.

3. Forgery detection using morphological image filtering based on guided contrasting

Morphological image filtering technique based on guided contrasting was proposed in [3]. This filtering technique makes available detecting changing between two images.

An image forgery detection algorithm based on guided contrasting can be performed in two stages:

1. Background normalization based on guided contrasting;
2. Image forgery detection using normalized background image processing.

3.1. Background normalization based on guided contrasting

Let f - standard image and g - test image. Proposed in [3] background normalization algorithm gives an opportunity to perform background normalization of the image g with considering the shape of the image f. The procedure of background normalization is performed using the window $D(x, y)$. The procedure is applied to the pyramid of images with a constant size of the window $D(x, y)$ to ensure invariance to the window $D(x, y)$ size.

Background normalization based on guided contrasting is as follows:

1. Construction of the pyramid representation $f^t = (f^0, ...f^{t-1})$ and $g^t = (g^0, ...g^{t-1})$ where $f^0 = f, g^0 = g$ and $size(f^i) = \frac{size(f^{i-1})}{2}, size(g^i) = \frac{size(g^{i-1})}{2}, i = \overline{1, t-1}$.
2. Calculation of the filter (9) response $\phi^i(f^i, g^i)(x, y)$ for every pyramid level:

$$\phi(f, g)(x, y) = g_0^{D(x,y)}(x, y) + \mid K(f^{D(x,y)}, g^{D(x,y)}) \mid (g(x, y) - g_0^{D(x,y)}(x, y)), \tag{9}$$

where $g^{D(x,y)}(u, v) = g(x, y), if(u, v) \in D(x, y); 0, otherwise$, $g_0^{D(x,y)}(x, y) = mean(g^{D(x,y)}(x, y))$ and $K(f^{D(x,y)}, g^{D(x,y)})$ is the local normalized correlation coefficient defined by (10).

$$K(f, g) = \frac{(f^{D(x,y)} - f_0^{D(x,y)}, g^{D(x,y)} - g_0^{D(x,y)})}{\parallel f^{D(x,y)} - f_0^{D(x,y)} \parallel \parallel g^{D(x,y)} - g_0^{D(x,y)} \parallel} \tag{10}$$

3. Calculation of absolute difference Δg_f^i between g^i and corresponding filter response $\phi^i(f^i, g^i)(x, y)$ for $i = \overline{1, t-1}$.
4. Reconstruction of a difference image from the pyramid. It is performed from level $t - 1$ with averaging on every level by following:

$$\Delta m_f^i(x, y) = \{g_f^i(x, y), if\ i = t - 1;\ max(g_f^i(x, y), h_f^{i+1}(x, y)),\ otherwise\} \tag{11}$$

where h_f^i - twofold spatial increased image g_f^i.

Image m_f^0 is the normalized background of the image g with considering the shape of the image f.

3.2. Image forgery detection using normalized background image processing

Image g forgery detection performs by normalized background image m_f^0 processing. This processing carries out as following:

1. Image m_f^0 binarization. For this procedure, thresholding an image with an opportunity to set a threshold by the user was chosen.
2. Morphological filtering of the binary image from step 1 such as erosion and dilation in the square window 3×3.
3. Segmentation and enumeration of all non-zero fragments.
4. Computation of an minimal convex hull of every suspicious segment and fill it.
5. Calculation of the local morphological correlation coefficient (MCC) K_M^n [4] for every suspicious segment by formula (12) and comparison it with a threshold:

$$K_M^n = \frac{\parallel g_F^n \chi^n(x,y) \parallel}{\parallel g \parallel} \tag{12}$$

where n is the number of an analyzed suspicious fragment, χ^n is the indicator function that has value "1" for pixels from analyzed suspicious fragment and "0" otherwise and g_F^n defined by:

$$g_F^n = \frac{(g, \chi^n)}{\parallel \chi^n \parallel^2} \tag{13}$$

6. Creating global mask with all forgery fragments which correspond to segments with MCC greater than the threshold.

4. Experiments

The experiments were carried out on a desktop PC with Intel Core i5-4460 processor and 16 GB RAM.

Five image time series were obtained using the same camera. The camera was still all the time. It has captured the scene and token image every 10 sec. As result of this procedure, there are five image time series with six images in every series. Obtained images have 920×1380 dimension. Next, every image has been transformed to gray-scale. These time series were chosen as the objects of experiments.

Series consisted of six images were used for image forgery detection through the algorithm based on anomaly detection. The algorithm based on guided contrasting used series from two images that are first and last.

Copy-move various type embedding procedure was developed for forgeries generation.

Experiments of two type forgeries detection were carried out:

1. Copy-move within one image - intra-image copy-move.
2. Copy-move from another image of the image time series - inter-image copy-move.

4.1. Intra-image copy-move detection

The experiment results with duplicate taken from the same image are shown in the tables 1 and 2. Example of detection using these algorithms is shown in the Figure 5.

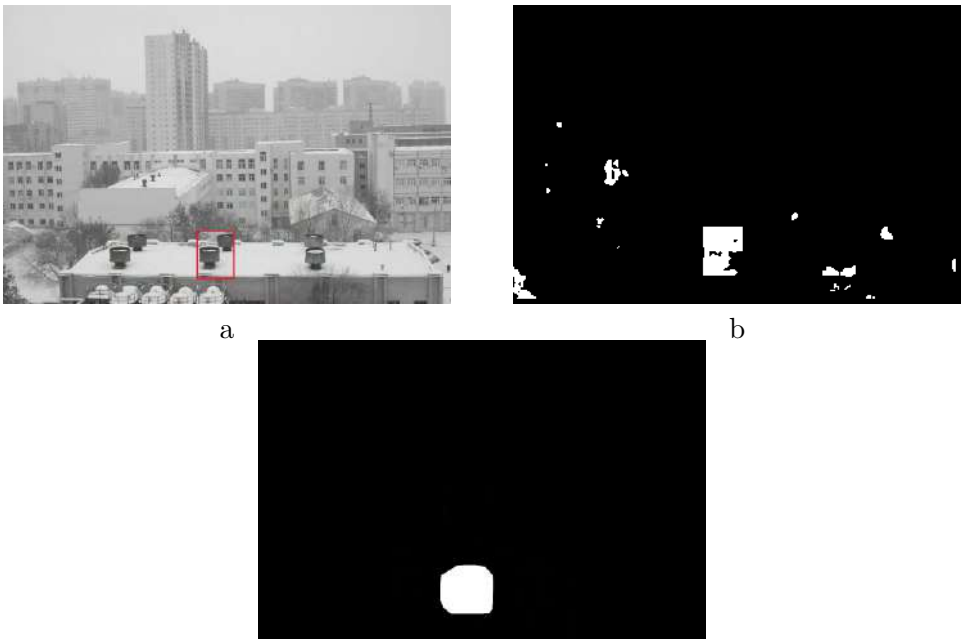
As shown in the tables 1 and 2, both algorithms give about the same results of F1 (0.66 is mean F1 for the algorithm based on anomalies detection, and 0.59 is mean F1 for the algorithm based on guided contracting). However, these algorithms reach these values on account of different components as so as Precision and Recall. So, the algorithm based on anomalies detection has high values of Recall. It means a greater portion of forgery pixels are detected using the algorithm based on anomaly detection than using the algorithm based on guided

Table 1. Intra-image copy-move detection using the algorithm based on anomalies detection.

Series	Precision	Recall	F1
1	0.55	0.84	0.66
2	0.53	0.74	0.62
3	0.62	0.85	0.72
4	0.55	0.68	0.61
5	0.63	0.78	0.70

Table 2. Intra-image copy-move detection using the algorithm based on guided contrasting.

Series	Precision	Recall	F1
1	0.93	0.40	0.56
2	0.94	0.67	0.78
3	0.38	0.03	0.06
4	0.50	1	0.68
5	0.86	0.93	0.89



a b c

Figure 5. Example of intra-image forgery detection: a - forgery image; b - result of detection using the algorithm based on anomalies detection; c - result of detection using the algorithm based on guided contrasting.

contrasting. On another hand, the algorithm based on guided contrasting has high values of Precision. It means this algorithm has less false detection than the algorithm based on anomalies detection.

4.2. Inter-image copy-move detection

The experiment results that were carried out on all image series with duplicate taken from another image of the image series are shown in the tables 3 and 4. Example of detection using these algorithms is shown in the Figure 6.

Table 3. Inter-image copy-move detection using algorithm based on anomalies detection.

Series	Precision	Recall	F1
1	1	0.84	0.91
2	1	0.95	0.97
3	0.58	0.74	0.65
4	0.8	0.82	0.8
5	0.87	0.84	0.85

Table 4. Inter-image copy-move detection using algorithm based on guided contrasting.

Series	Precision	Recall	F1
1	0.	0.	0.
2	0.75	0.004	0.009
3	0.	0.	0.
4	0.16	0.06	0.09
5	0	0	-

As shown in the tables 3 and 4, the algorithm based on guided contrasting doesn't give opportunity detecting inter-image copy-move forgeries while the algorithm based on anomalies detection does it and give high values of F1.

5. Conclusion

The algorithm for image time series forgery detection based on anomaly detection was proposed in this paper. Also, comparison of the proposed algorithm and the algorithm for image forgery detection based on guided contrasting carried out. Experiments showed that both algorithms have about the same quality of detection intra-image copy-move in the sense of metric F1 (0.66 and 0.59 respectively). On another hand, experiments let to conclude that the algorithm of image forgery detection based on guided contrasting doesn't give opportunity detecting inter-image copy-move, unlike proposed algorithm.

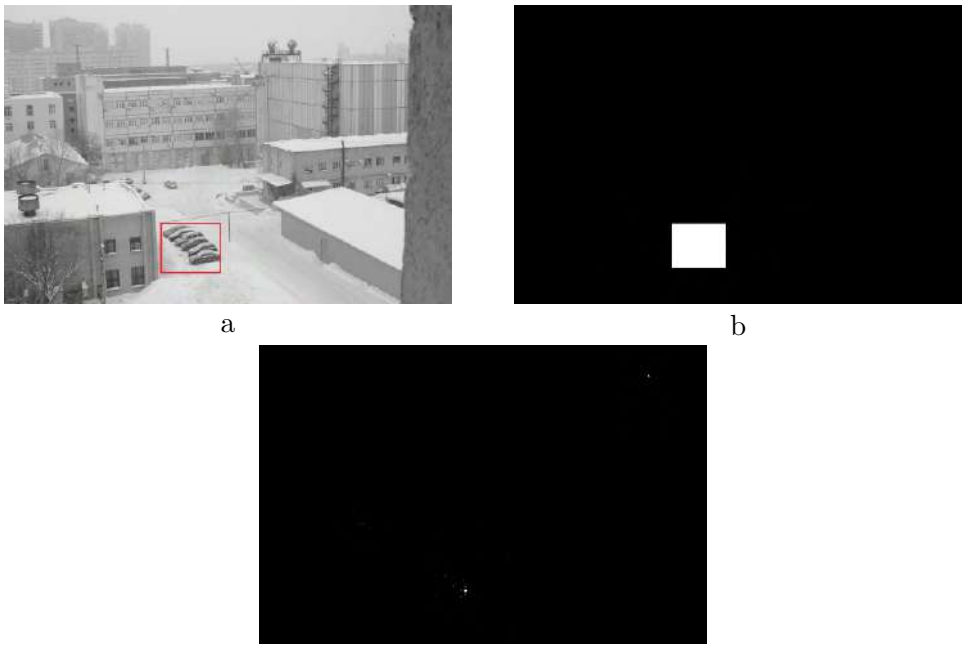


Figure 6. Example of inter-image forgery detection: a - forgery image; b - result of detection using the algorithm based on anomalies detection; c - result of detection using the algorithm based on guided contrasting.

6. References

[1] Christian A and Sheth R 2016 Digital video forgery detection and authentication technique - a review *International Journal of Scientific Research in Science and Technology* **2(6)** 138-143
[2] Chandola V, Banerjee A and Kumar V 2009 Anomaly detection: a survey *ACM Computing Surveys* **41(3)** 51-58
[3] Rubis A Yu, Lebedev M A, Vizilter Yu V and Vygolov O V 2016 Morphological image filtering based on guided contrasting *Computer Optics* **40(1)** 73-79 DOI: 10.18287/2412-6179-2016-40-1-73-79
[4] Pyt'ev Yu P and Chulichkov A I 2010 *Methods of Morphological Image Analysis* (Moscow: Fizmatlit)
[5] Evdokimova N I and Kuznetsov A V 2017 Local patterns in the copy-move detection problem solution *Computer Optics* **41(1)** 79-87 DOI: 10.18287/2412-6179-2017-41-1-79-87
[6] Kuznetsov A V and Myasnikov V V 2016 A copy-move detection algorithm based on binary gradient contours *Computer Optics* **40(2)** 284-293 DOI: 10.18287/2412-6179-2016-40-2-284-293
[7] Kuznetsov A V and Myasnikov V V 2014 A fast plain copy-move detection algorithm based on structural pattern and 2D rabin-karp rolling hash *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **8814** 461-468
[8] Hussain M, Chen D, Cheng A, Wei H, Stanley D 2013 Change detection from re-motely sensed images: from pixel-based to object-based approaches *ISPRS Journal of Photogrammetry and Remote Sensing* **80** 91-106

Acknowledgments

The reported study was funded by RFBR according to the research project 17-29-03190, research project 18-01-00748 and by the Federal Agency of scientific organization (Agreement 007-Г3/43363/26).

Algorithm for detecting block-like cracks in facies of human biological fluids

V R Krasheninnikov[1], L I Trubnikova[2], O E Malenova[1], A S Yashina[3], M L Albutova[2] and O A Marinova[2]

[1]Ulyanovsk State Technical University, 32, Severny Venetz Street, Ulyanovsk, Russia, 432027
[2]Ulyanovsk State University, Lev Tolstoy Street42, Ulyanovsk, Russia, 432017
[3]Research-and-Production Association Mars, Solnechnaya Street 20, Ulyanovsk, Russia, 432022

Abstract. One of the most effective methods of early medical diagnosis is based on the analysis of facies (thin films) images of dried human biological fluids. The presence of special structures (markers) in facies indicates various pathologies of the organism at their earliest stages. To accelerate and reduce the cost of mass preventive medical examination of the population, it is required to create algorithms that allow markers detection on images of facies. In this paper we present an algorithm for detecting the "block-like crack" marker on images of facies of the follicular fluid obtained by puncture in vitro fertilization process. This marker indicates hypoxic and ischemic brain lesions. When developing this algorithm, first, through a visual analysis of the marker, its characteristic features were revealed. Further, methods of algorithmic detection of these features were developed. The decision about the presence of the marker was taken if there was a combination of its features. When testing this algorithm, 86% of images containing these markers were found, with 11% of false alarms.

1. Introduction

The development of algorithms for the automated analysis of images of the facies of biological fluids (BF) is of interest for a number of reasons. On the one hand, the processes that occur during the drying of drops of various BF (saliva, blood serum, cerebrospinal fluid, urine, etc.) are intensively studied, which can provide additional information in the diagnosis of diseases and make early diagnosis in the absence of visible symptoms, allowing treatment to begin on the early stages. On the other hand, this method belongs to the group of non-invasive methods of investigation, which is especially important in the diagnosis of diseases in newborns and premature infants. In addition, the acceleration achieved and the reduction in the cost of image analysis makes it possible to conduct mass preventive surveys of the population, helping to improve the quality of health care.

The method of studying biological fluids (BF) by means of their dehydration and analysis of the crystallization of contained substances has a long history. The theoretical description of the process of evaporation of a drop was made by D.K. Maxwell [1]. Bohlen investigated the facies of capillary blood and noticed a connection between markers and gastrointestinal tumors [2]. In the series of works by V.N. Shabalin and S.N. Shatokhina (for example, [3, 4]) the features of the crystalline structures of BF and their connections with pathologies were analyzed. The comprehensive review and bibliography on the study of BFare given in the book by SA. Kraevoy and N.S. Coltovoy "Diagnosis by drop of blood, crystallization of biofluids" [5]. Currently, the wedge-shaped dehydration method is used to diagnose diseases in oncology, gerontology, pediatrics, obstetrics, gynecology and other fields

of medicine. However, most of these works are aimed at identifying markers of pathologies and improving visual perception: modernization of microscopes, addition of chemical reagents in BF, etc.There are considerably fewer articles on computer processing of facies images [6-12]. And even among these works, methods of finding image zones suspicious of the existence of markers are usually developed, for example, in [10] wavelet analysis is used for this purpose. Algorithms presented in [13-21] are elaborated to detect several peculiar markers with high probability while the probability of false alarms is low.

In this paper, the algorithm for detecting a marker is a block-like crack. This marker indicates the presence of structural changes in tissues, hypoxic and ischemic brain lesions.The markers on facies of the BF are very diverse in shape, size, orientation, etc., which is the main difficulty in their algorithmic recognition. In the present work, the following approach for the development of recognition algorithms is used. First, a visual analysis of the markers is carried out to reveal their characteristic features. Then the methods of algorithmic detection of these features are developed. The decision about the presence of the marker is made if a combination of its necessary characteristic was found in the image section.Note that only a medical worker can make the final diagnosis to the patient. Computer analysis of the facies is only auxiliary. Its destination is to identify images on which specific set of markers is present. Therefore, it is not necessary to find all markers of this type on the image. It is enough to find at least one of them and inform about it. Then the operator will perform a more thorough analysis of this image. In fact, it is required to select images which contain at least one of considered markers. This is an indicator of the effectiveness of algorithms for detecting and recognizing markers.

2. Algorithm for block structure detection

To develop the algorithm for detection of the marker, its main features were first revealed. They distinguish the marker from the basic structure of the facies and other markers.

2.1. Marker description

Most often the term "blocky" is used to denote the type of joint. The articular joint (artic trochlearis), or ginglium, is formed by a part of the block, cylinder or truncated cone on one bone and corresponding to the shape of a depression on the other [22]. It is schematically depicted in figure 1. On the left is a diagram of the block-shaped joint. On the right is its simplest schema to geometric primitives. This is more like the desired marker shape on the facies images.

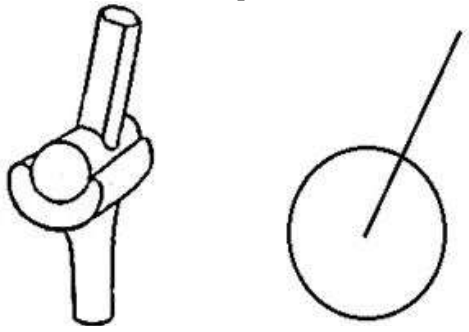


Figure 1. Scheme of the block joint.

Figures 2-6 show images of the facies of a healthy person and facies with block-like structures. Marker is recognized by a long crack and one end is enclosed in an oval. Fractions of the facies look like dark lines on the image. As can be seen from the figures, the marker has a lot of variations of form. At the end of a long crack, there may be a closed circular or oval crack or an open arcuate crack. These markers appear at the ends of radial cracks (long faults directed to the center of the facies). An exception may be some pathological conditions. With them, the line of intersection of radial cracks shifts from the center of the facies. Radial cracks are clearly visible in figure 3.

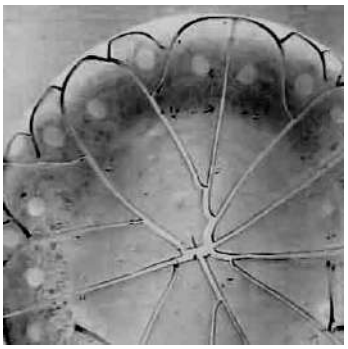


Figure 2. Bloodfacia of a healthy person.

Figure 3. Block-like cracks.

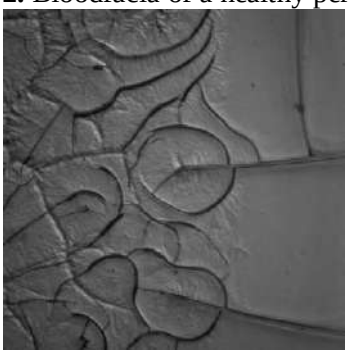


Figure 4. Block-like cracks.

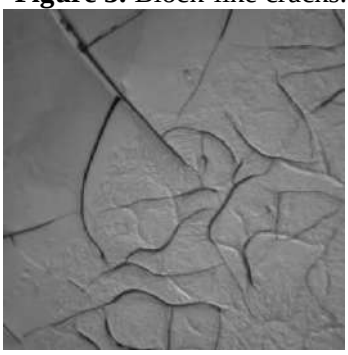


Figure 5. Block-like cracks.

2.2. Markersearching algorithm

Based on the characteristics described above, an algorithm for searching for a marker was compiled. It includes the following steps:

- Preparing the image for processing.
- Segment the image in brightness for highlighting cracks.
- Selection of contours.
- Search for radial cracks.
- Search for rounded cracks that intersect radial cracks.

Decision making about the presence of a marker. If the image shows the intersection of a radial crack with an arc or an ellipse, then a decision is made about the presence of a marker on the image.

3. Description of the algorithm steps

Let's consider the steps of the proposed algorithm operation on the example of image processing of the facies. The image shown in figure 6 is taken as example.

3.1. Preparing the image for processing

When you run the program, a color image is loaded into the memory. In our case, the colors of the marker do not carry any important information. Therefore, the whole image is converted to grayscale (figure 6). To reduce random noise and the number of false contours, we use a median filter.

3.2. Brightnessimage segmentation

Cracks are darker than the main part of the facies. For their selection, the average brightness of the image is calculated. All areas with brightness above the average are removed from further consideration (figure 7).

3.3. Extraction of contours

To isolate the contours, the Canny algorithm was used [23]. The result of applying it to figure 7 is shown in figure 8. To isolate the contours boundaries by this method, the values for the two thresholds were experimentally selected.

Figure 6. Uploaded image in grayscale.

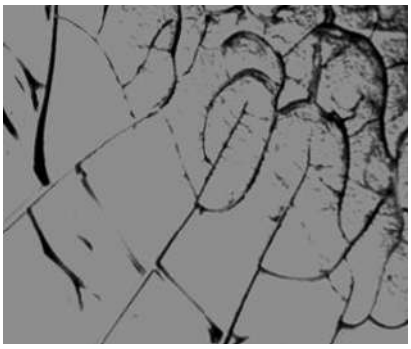


Figure 7. Threshold segmentation result.

The minimum threshold is 100 and the maximum threshold is 180. They determine the significance of the boundary points. If the gradient value on this fragment is below the minimum threshold, then the boundary is considered to be insignificant. If there is no value on the border fragment above the maximum threshold, it is deleted. It is noticeable that there are many extraneous, useless contours (short ones). Therefore, by using per-pixel bypass of the contour, we leave only the boundaries of sufficient length. The empirically calculated threshold is 50 pixels. The bypass is implemented using the circuit bypass algorithm called "Beetle" [24]. The result is shown in figure 9.

Figure 8. The result of the Canny algorithm.

Figure 9. Insignificant contours removing.

3.4. Search for radial cracks

After isolating the contours, it is easy to find long approximately rectilinear cracks. For this, Huff's transformations were used [25]. With its help, long enough straight lines have been found, which can be attributed to radial cracks. It is empirically established that as x400 magnification, the length of the radial crack exceeds 150 pixels. The found radial cracks are shown in figure 10.

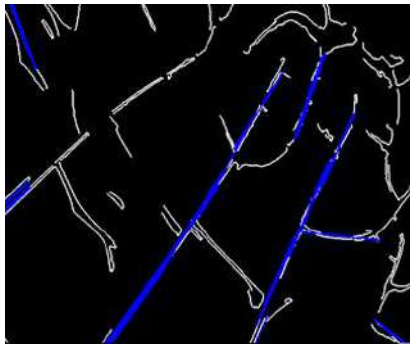


Figure 10. The result of searching for radial cracks.

3.5. Search for rounded cracks

The shapes of our marker vary greatly and not everyone can be an ellipse. Therefore, we use an algorithm based on the search for arcs in an image. It is described in [26] and is based on the following idea. A line similar to an arc must have the same direction of rotation (left or right) along its course.

And the angles between consecutive segments should be in a certain confidence interval (figure 11). If the angle is too small, then the segments are considered to be collinear. They cannot be part of the arc. If the angle is too large, then the lines are part of the contour of another type (square, rectangle, etc.). In our case, the confidence interval of the angle of rotation was chosen experimentally from 6 to 60 degrees. Since the bypass goes in a specific direction, parts of the arc will be considered as vectors (figure 12). The angle between them is determined by the relation:

$$\theta_i = \arccos\left(\frac{\overline{v_{i+1}}\,\overline{v_i}}{|\overline{v_{i+1}}||\overline{v_i}|}\right). \quad (1)$$

The detected arcs are shown in figure 13.

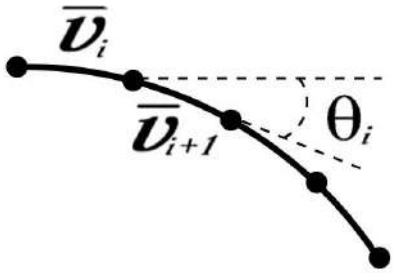


Figure 11. Arrangement of arc segments.

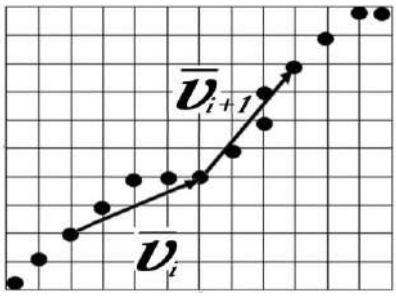


Figure 12. Vectors on the line.

Figure 13. The found arcs.

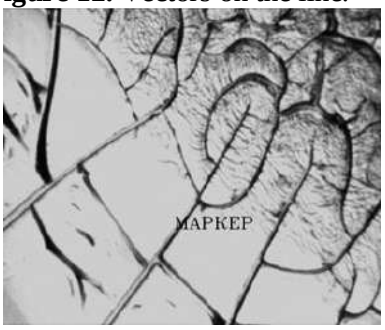


Figure 14. Discovered marker.

3.6. Decision-making

We check the angles of intersection of the found arcs and radial cracks. If the angle found is close to a direct angle, then a decision is made about the presence of a block-like crack in the facies image. In figure 14, the inscription indicates the found marker.

4. Experimental results

The purpose of image processing facies is to identify images on which there are desired markers (in our case - block-like cracks). In this case, it is not necessary to find all markers of this kind, it suffices to indicate that such markers (at least one) are available. Therefore, when testing the described algorithm, it was considered that the detection occurred if at least one block-like crack was found on the image on which they were available.

The algorithm was tested on 300 images of blood serum facies of patients examined at the medical faculty of Ulyanovsk State University. As a result of the tests, 86% of images with the desired markers were correctly identified. To find the probability of false alarms, this algorithm processed image facies, on which this marker is absent. It turned out that on 11% of such images false detections of structures were made, somewhat similar to the desired marker. This suggests that in the future it is advisable to extend the set of features of this marker for more accurate identification.

5. Conclusion

As a result of this study, an algorithm and software were developed for the automated detection of facies images with suspicion of the presence of a "block-like crack" marker and the screening of

images without this marker. Tests showed that 86% of the images on which markers were present were properly identified. False detections of structures were made on 11% of images without this marker. But it found somewhat similar to the desired marker. This indicates that in the future it is expedient to extend the set of characteristics of this marker for more accurate identification. It should be noted that among the images from the list of false alarms there were other markers, which are some kind of "suspicious areas". Their search is also included in the tasks of detection algorithms for markers.

6. References

[1] Maxwell J C 1890 *The Scientific Papers of James Clerk* (Niven University Press Cambridge) p 211
[2] Bolen H L 1942 The Blood Pattern as a Clue to the Diagnosis of Malignant Disease *J Lab Clin Med* **27** 1522-1536
[3] Shatokhina S N and Shabalin V N 2013 *Atlas of structures of human noncellular tissues in norm and pathology: Morphological structures of blood serum* (Moscow-Tver: Triad) p 240
[4] Shatokhina S N and Shabalin V N 2001 *Morphology of human biological fluids* (Moscow: Chrysostom) p 304
[5] Kraevoy S A and Koltovoy N A 2016 *Diagnosis using a single drop of blood. Biofluid crystallization* (Moscow-Smolensk)
[6] Abdumanonov A A, Botirov M T and Karabaev M K 2014 Development of computer programs and algorithms for automatic morphometry bioliquid facies *New look. International Scientific Gazette* **6**
[7] Buzoverya M E, Shishpor I V, Sevastyanov V I and Perov N V 2011 *The hardware and software complex "Morfo"* (Sarov: RFJC-VNIIEF) 254-255
[8] Ilyasova N Yu, Kupriyanov A V and Khramov A G 2012 *Information technologies of image analysis in problems of medical diagnostics* (Moskow : Radio and communication) p 424
[9] Lukashevich M M 2013 Analysis of images of biological fluids *Bulletin of the Fund for Fundamental Research* **3** 94-105
[10] Makarevich V E, Kirilenko E A, Petrashenko V A, Zablotskaya T U and Bilokon T A 2014 Methods of wedge dehydration of biological fluids *Morphology* **8(1)** 113-117
[11] Petrov V O 2009 *Automation of the analysis of raster images of a solid phase of a biological fluid of biomedical preparations* (Volgograd: VSTU) 150 p
[12] Pogodin S V 2010 Automatic diagnosis on the analysis of medical images *Collection of scientific works of the 1st International teleconference "Problems and perspectives of modern medicinebiology and ecology* **1** 71
[13] Krasheninnikov V R and Kopylova A S 2011 Identification of Pectinate Structures in Images of Blood Serum Facia *Pattern Recognition and Image Analysis* **21(3)** 508-510
[14] Krasheninnikov V R and Kopylova A S 2012 Algorithms for automated processing images blood serum facies *Pattern Recognition and Image Analysis* **22(4)** 583-592
[15] Krasheninnikov V R, Trubnikova L I, Albutova M L, Kopylova A S and Tarasova A V 2013 Algorithm for detecting language structures on images of blood serum *Ulyanovsk medico-biological journal* **4** 70-73
[16] Krasheninnikov V R, Trubnikova L I, Albutova M L, Kopylova A S and Tarasova A V 2013 Detection of sickle-shaped cracks in images of facies of cervical mucus of a woman *Ulyanovsk medico-biological journal* **4** 81-85
[17] KrasheninnikovV R,Trubnikova L I, Albutova M L and Yashina A S 2015 Algorithm for detecting a marker of gallbladder disease on images of blood serum *Ulyanovsk Medical and Biological Journal* **4** 104-110
[18] Krasheninnikov V R, Yashina A S and Malenova O E 2017 Markers detection on facies of human biological fluids *Procedia Engineering* **201** 312-321
[19] Krasheninnikov V R,Yashina A S and Malenova O E 2017 Algorithms for detection of markers on the facies of human biological fluids *Proceedings of the III International Conference and Youth School "Information Technologies and Nanotechnologies" (ITNT-2017)* 655-662

[20] Krasheninnikov V R, Yashina A S and Malenova O E 2017 Algorithms of crescent structure detection in human biological fluid facies *The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences* **4** 169-172
[21] Zakharova G P and Shabalin V V 2014 Computer identification of distinctive features and anomalies in samples of multicomponent solutions of biological fluids *Russian Otorhinolaryngology* **6** 37-42
[22] Vorobiev V P 1940 *Atlas of human anatomy* (Leningrad: Medgiz) p 382
[23] Canny J A 1986 Computational Approach to Edge Detection *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6** 679-698
[24] Potapov N N and Tupikov P A 2007 Methods for selecting contours in images *8th international conference Image recognition and image analysis: new information technologies* 105-108
[25] Ballard D H 1981 Generalizing the Hough Transform to direct arbitrary shapes *Pattern Recognition* **13** 111-122
[26] Akinlar C and Topal C 2013 ED Circles: A real-time circle detector with a false detection control *Pattern Recognition* **46** 725-740

Methods of detection and ecological state evaluation of protective forest belts using complex ground survey and remote sensing data processing

L M Kavelenova[1], E S Korchikov[1], N V Prokhorova[1], D A Terentyeva[1] and V A Fedoseev[1,2]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086
[2]Image Processing Systems Institute - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001

Abstract. The protective forest belt system of Samara region consists of several components created from the late 19th to the early 20th century, in the middle of the 20th century, in the 60-70s of the 20th century, and more recent plantings. In the forest-steppe and steppe regions with a high level of agricultural transformations, the natural biological equilibrium of ecosystems is significantly disturbed. The protective forest belts play an important role in natural and anthropogenic landscapes conservation in Samara region. Because of numerous dispersed forest belt distribution throughout the Samara region territory, their different ownership and the lack of sufficient funds for their monitoring, it is difficult to organize belt state monitoring on the ground. The use of space imagery, which is processed during verification using data obtained from reference polygons, can help to overcome this situation. The peculiarity of forest belts as an extended object of relatively small width actualizes the task of developing the methods for their condition assessing. In this paper, some results of this work are analyzed for reference areas of field shelterbelts and roadside forest belts of the Samara region.

1. Introduction

In the forest-steppe and steppe regions with a high degree of land plowing, the natural biological equilibrium of ecosystems is significantly disturbed. In these regions, the protective man-created forest plantations seem to be here a sort of important connections between natural and anthropogenic landscapes. The forest belts reduce wind and water erosion by their participation in the redistribution of moisture and lowering of wind currents speed. When absorbing aerial pollution, they carry out its redistribution from air to phytomass and then to the soil environment, forming zones with increased accumulation of technogenic pollutants. The ecosystems developing on the basis of forest belts are also characterized by ambiguity. On the one hand, in conditions of the prevalence of agrocenoses, they become shelters of species of natural flora. On the other hand, introduced plant species used in the forest belts can spread in the environment beyond plantations, becoming agents of biological invasions. Nevertheless, the positive role of forest belts seems to be predominant, what makes necessary the realization of measures aimed at their preservation.

The system of protective forest belts of Samara region includes the components related to different periods of creation: from the late 19th to the early 20th century; in the middle of the 20th century; in

the 60-70s of the 20th century; in the last decades of the 20th century and first years of the 21st. The first attempts to create forest belts in the Samara region were undertaken in 1872 but were not almost successful [1]. That time ash, yellow acacia, and oak trees were planted mainly on solonetsous soils that are of little use for afforestation, so most of the plantations were almost destroyed, remaining only 16%. Later, in 1894-1897, about 60 hectares of forest belts, called "Timashevskie," were laid. Based on the results of these works, Nestor Genko, an enthusiast of national forestry, developed the basic provisions of agrotechnics for the regional plantations cultivation [2]. The watershed forest belts created under his leadership represent broad stripes of artificial forest with a width of more than 600 m, located in the form of broken lines with the length from 2.5-3.5 to 25 km along the watersheds of rivers Samara and Chapaevka (Dubovskie and Teplovskie), Chapaevka and Chagry (Kamyshinskie , Bezenchukskie and Vladimirskie), Sok and Padovka (Shilanskie) [1, 2]. These forest belts stretch mainly from the southwest to the northeast, perpendicular to the direction of the dry winds and are divided by longitudinal, and every 80 to 200 m and by transverse clearings used as forest roads [1].

The surviving fragments of Genko forest belts, which are currently assigned the status of natural reserves, have been significantly transformed in comparison with the original plantations. The oldest specimens of trees were lost, and their place was taken by the individuals formed due to self-seeding, which led to a distortion of the rock composition of the strips. Ordinary placement of woody plants is also not completely traceable. The period of 1892-1906 accounts for the greatest amount of work on the creation of forest belts, in 1906 the planting was stopped, and then resumed only in 1949. It was precisely in 1949-1962 that the second large-scale period of the creation of forest shelterbelts called the "Stalin's plan for the transformation of nature" designed for 15 years. During this time, about 2 million hectares of forest shelterbelts were planted, the total length of which exceeded 5000 km, and the direction was designed to block the way to the dry southeast winds. By 1952 the system was basically created and started working. The system of State forest belts along the reservoirs of the rivers the Northern Donets, the Volga and the Ural was divided into eight zones. Each of them consisted of 1 to 6 parallel narrow stripes of trees with 30 to 60 meters width, located at a distance of 300 meters from each other. On the territory of Samara region, the fourth zone begins, with a length of 580 km, from Chapayevsk (not far from Samara), which stretches to Vladimirovka-on-Volga. According to the design, the strip consists of 4 parallel forest belts 60 meters wide and 300 meters between the bands [1, 3]. From 1966 to 1982, more than 3 thousand hectares of protective forest plantations per year were created annually. The last years of XX in Samara region were lost for new forest belts creation, but later such practice was revived (figure 1).

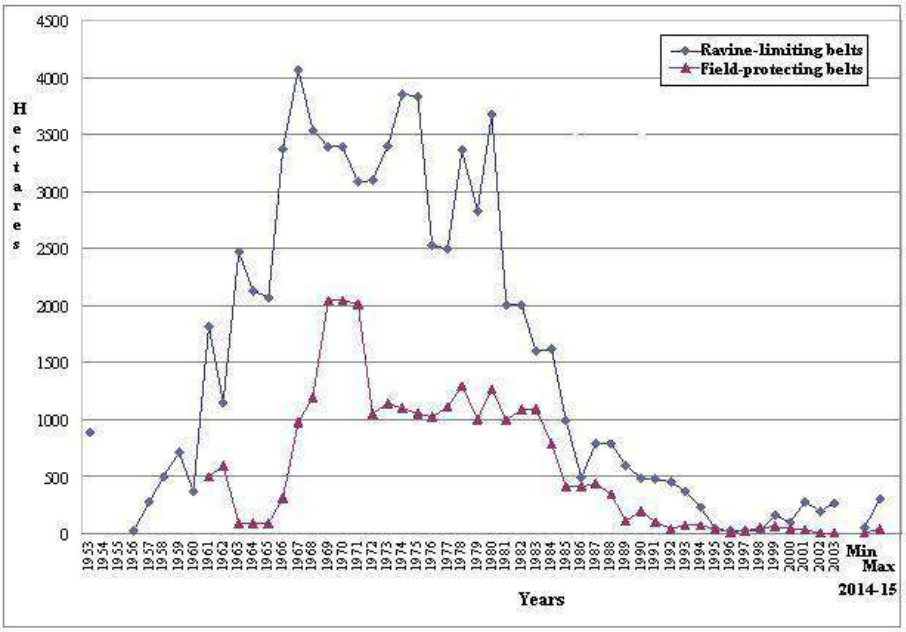


Figure 1. The dynamics of forest belts creation in Samara region (based on [1, 3]).

According to the forest fund inventory, the total area of forest belts is more than 8.5 thousand hectares. Currently, forest belts are represented mainly by second-generation coppice stands, which resulted from clear cuttings. The average age is 55 years. Plantations of seed origin in the age of 90 - 100 years occupy a smaller area, because in wartime were the only source of wood. Massively represented forest belts of "middle" age, both field-protective and roadside, are characterized by different life conditions of the stand. In recent years, there has been a clear trend towards the loss of birches in many plantations, for which the consistent impact of stressful conditions in the winter of 2009/2010. and subsequent drought in 2010 caused severe damage to the root systems. The dying of trees, which lasted for several years, resulted in the actual loss of a number of roadside belts with birch as the main tree species.

2. Problem statement

Because of the forest belts multiplicity and dispersed distribution over the entire territory of Samara region, also as their location on the different subordination lands, as well as the lack of sufficient funds for their monitoring, the organization of a ground survey of the forest belts current state is extremely unlikely. That is why it is expedient to use for this task remote sensing images which are very helpful in many other problems on vegetation analysis [4-6]. The peculiarity of forest belts as a linear extended object of relatively small width actualizes the task of developing methods for their detection and assessment of the state, including using the spectral characteristics of these plantations. In this paper, some results of this work are analyzed for reference areas of field shelterbelts and roadside forest belts of the Samara region.

3. Brief review of field study results

Three polygons were chosen as models taking into account their target membership (field, roadside) and species composition. Further, during the ground survey, the coordinates of the start and end points of the sites and their complex survey were coordinated, including the determination of the state of the plantations (figure 2):

1. “The environs of Bobrovka village" (located in the Kinelsky district, 3 km to the east of the village. Bobrovka, two reference plots of forest field shelterbelts reference plots 1, 2). The forest belts are four-row plantations (2 - *Fraxinus lanceolata*, 2 – *Ulmus glabra*). The plot 1 is practically impassable, represents a continuous array of closed crowns, with separate downward-facing shrunken branches, withered trunks, light impregnations of trunks without bark. Planting is quite uniform, with a well-developed wood self-sowing from the north-west side, in the vegetative state it carries out active light absorption (with maxima in the red and blue regions of the spectrum). The plot 2 is rarefied, with "windows" and glades, which could have resulted from the death of a part of trees and due to unauthorized cuttings. Here also are the trunks of dead trees, facing the shrunken branches, The planting has inhomogeneous character, especially noticeable in the tier of crowns of the main rocks, will manifest itself in the features of the light absorption pattern.

2. “The environs of Formalny village" (located in the Kinelsky district in 2 km to the north-east of the Formalny village, 3 reference plots, one of which belonged to a group of field shelterbelts, and two to roadside strips (reference plots 1 - 3). The plot 1 (forest field shelterbelt) also is a four-row plantation with a relatively sparse arrangement of trees. On both sides of the strip, self-seeding of tree species is expressed, its strips do not reach 4 m in height. Planting openwork in the area of the crowns, in the lower part of the partially openwork, should give a relatively homogeneous pattern of light absorption. Plots 2 and 3 (fragments of the roadside strips) are also a four-row plantations (*Betula pendula, Ulmus glabra, Betula pendula, Ulmus glabra*). Plot 2 is characterized by irregular tree placement as a result of the death of individual specimens. The plantation is open-spaced in the crowns area, in the lower part is partially opened. The light absorption within the forest belt should give a non-uniform picture due to the loss of the woody places in places. Plot 3 in the result of a part of birches missing has glades in the tree layer. Partial filling of vacant places in the plantations occurs due to self-sowing and birch shoots originated from the trunks of weakened trees and individual stumps. For these roadside forest belts, we assumed future difficulties in distinguishing them from the images, which is due to their relatively small width (< 20 m).

3. “The environs of Malaya Malyshevka village” (located in the Kinel region, 6 km to the west of the Malaya Malyshevka village, four reference sites were laid, among which three belonged to the group of field shelterbelts, and one to roadside) (reference plots 1-4). Plot 1 is the forest belt is a plantation with a relatively uniform arrangement of trees, blown along the entire height. For plots 2 and three the high life state of birch trees as the main species were distinguished. Plot 4 as a fragment of a roadside forest belt, initially formed as a four-row plantation, is practically impassable now due to the windfall and the inclined crowns closing.

Figure 2. Scheme of the location of the forest belts reference areas: Reference area 1 - environs of Bobrovka village, Reference area 2 - the environs of Formalny village, Reference area 3 – environs of Malaya Malyshevka village.

4. Detection of forest belts in space images

Further, for this territory, the search for actual space images of spatial resolution as much as possible was carried out. As a result, fragments of images from the Resurs-P spacecraft obtained by the Geoton instrument with a resolution of 2.4 meters in four spectral channels (near infrared, red, green, and blue) were selected. The image size was 3158 × 3159 pixels and 10 bits per pixel. Due to the narrowness of the forest belts (about 10-15 m, as the survey showed), and also the predominantly mixed rock composition, the resolution of the Resource-P data (both spatial and spectral) was insufficient to classify the state of the belts (vitality of plantations). Therefore, the only problem of automatic detection of bands in the selected terrain from the available space images was considered.

For the detection of forest belts, we chose a technology consisting of two stages. At the first stage, a thematic classification of the territory into different types of the underlying surfaces was carried out to identify all areas of the forest. The second stage was devoted to the allocation of narrow forest belts among the found areas of the forest.

For the first stage of thematic classification, we used the previously developed technology [7], which includes the following main steps:

1. Pixel-wise preprocessing of input data using the combination of NAPC and PCA transformations.
2. Spatial preprocessing of input data through median filtering with a small window.
3. Classification procedure which takes into account both the cluster structure of data and spatial information. This procedure is based on k-means++ clustering followed by the connected components filling. The resulting regions were classified by majority voting.
4. Spatial post-processing of classification results using nonlinear filters.

The results of our previous studies [7] showed that the performance of this technology is the possibility of achieving high classification accuracy even in the case of a small training set selected from the spatially localized fragments of the analyzed image. It is convenient to use parts of the selected control and measuring polygons as such fragments, which should be supplemented with examples of other samples of the underlying surface.

In our investigation, this technology was compared with a classification algorithm based on hierarchical histograms [8]. This algorithm analyzes a feature vector structure and performs the construction of a hierarchical histogram of the training sample at the learning stage. The histogram is stored in the form of an unbalanced fixed-depth tree, each node of which contains information about the frequency of sample objects that hit the interval of the histogram given by this node. As a result of the truncation of the tree by the principle of node preservation, the frequency of the hit of the sampling elements is greater than a given threshold, the generalizing ability of the classifier is ensured. At the classification stage, the cell of the hierarchical histogram to which the feature vector belongs is determined, and the pixel is included in the class with the maximum frequency of the training sample elements entering this cell. Pixels for which the histogram cell was not found are included in the additional class "not classified".

Figure 3a shows a fragment of an image from the Resurs-P satellite (the Geoton sensor) that contains a part of the analyzed territory including the selected measuring polygons.

To evaluate the classification quality, we prepared manually a ground-truth markup of the territory which includes six classes of the underlying surface: asphalt, forest and various types of crops (see figure 3b). The amount of the marked data is about a half of all pixels at the image. As can be seen from figure 3a, the forest class is represented by forest belts bordering fields, small forest areas at the upper part and at the lower part, on the banks of small reservoirs.

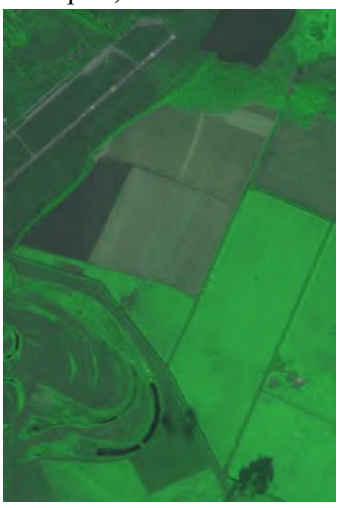
a)  b) 

Figure 3. The analyzed satellite image of the satellite with Resurs-P (a) and the marking of the main classes of the underlying surface (b).

Figure 4a shows the results of the automatic classification of this image using the selected technology. The training sample is formed from the pixels of localized areas, and its size is 1% of the all marked pixels in figure 3b. Figure 4b gives the results of the pixel-wise classification for comparison. The chosen technology provides a higher quality solution to the problem. The numerical estimation of the classification accuracy calculated for all the marked data was 95% for the figure 4a.

5. Detection of forest among the found forest areas

In figure 5a, areas classified as forests are marked black. Further, we need to distinguish forest belts among these areas. Forest belts are characterized by narrowness and consist of long straight sections. To find them, we first applied the following mathematical morphology operations [9]:

1. Closing with a small window to merge separated fragments of forest belts.
2. Opening with a window exceeding the maximum forest belt width, to remove all the belts.
3. Subtraction of the step 2 result from the step 1 result.

As a result of these operations, fragments of forest belts and groves, comparable to them in width, should remain on the image, as shown in figure 5b.

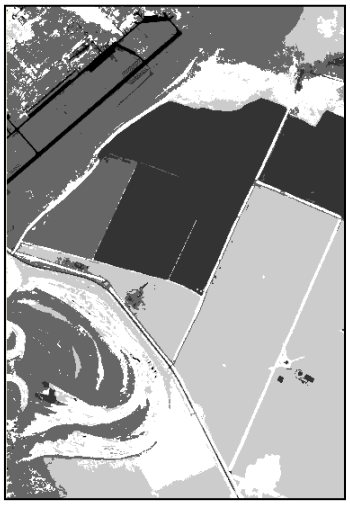
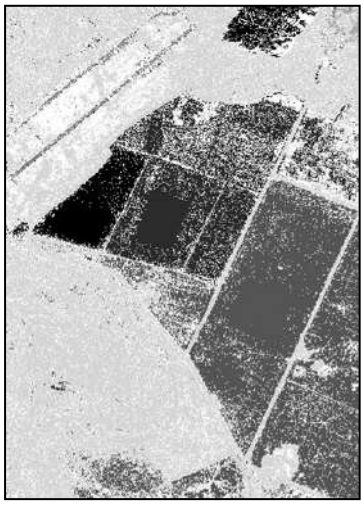
a)  b) 

Figure 4. Classification results of the image in Figure 3a: (a) using the selected technology, (b) using pixel-wise classification.

Next, to reveal linear fragments from the remaining objects, we used Hough transform [9]. On the Hough spectrum, we found local maxima corresponding to the most vivid linear components. The coordinates of the obtained maxima uniquely determine the straight lines. Their intersection with the regions selected in figure 5b allowed to determine the ends of the desired segments, which ideally should coincide with the boundaries of the forest belts. The final results are depicted in figure 6 (the segments are marked in red). One can see that despite some false segments all the required forest belts of sufficient width (more than 8 pixels in the given Geoton image that corresponds to 20 meters on the ground) were successfully found.

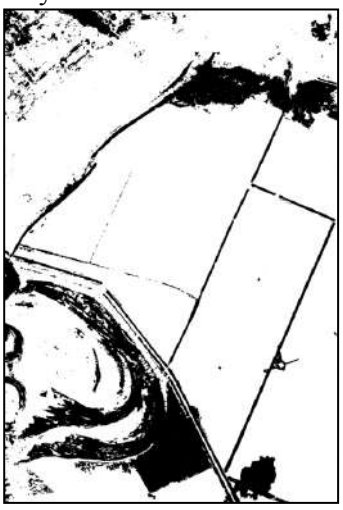
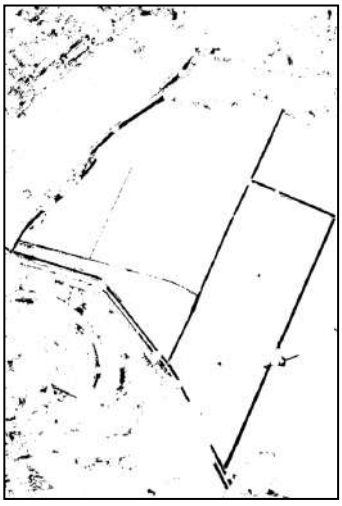
a)  b) 

Figure 5. Forest areas (a) and the result of their morphological processing (b).

6. Conclusion

In this paper, we presented the results of forest belt detection on satellite imagery. This problem is part of a practically meaningful study aimed at assessing the ecological state of forest belts in the Samara region. To solve this problem, a two-stage method was used, consisting of thematic classification of remote sensing data to identify forest areas and subsequent detection of forest belts from the classification results. The belt detection stage was performed using a technique based on mathematical morphology and Hough transform. To get the data used in this study, a field survey was conducted, as a result of which control and measurement polygons were identified, the coordinates of which were subsequently used for numerical evaluation of the quality of the solution of the classification problem. The obtained results after analysis of four-channel images with a spatial resolution of 2.4 m allow us

to conclude that the selected technique can be used to solve the problem of finding forest belts. For the subsequent analysis of the ecological state of plantations, data having a greater spatial and spectral resolution are required.

Figure 6. Results of forest belt detection.

7. References

[1] Shabalin I M 2005 *The Samara forest* (Samara)
[2] Pajusova T V 2010 "Green Bastions" by Nestor Genko *Samara Luke: Problems of Regional and Global Ecology* **4** 177-182
[3] Kazantsev I V and Matveeva T B 2016 Ecological state of forest belts in the Samara region *New science - from idea to result: International scientific periodical on the results of the International Scientific and Practical Conference, Surgut* 5-8
[4] Afanasyev A A and Zamyatin A V 2017 Hybrid methods for automatic landscape change detection in noisy data environment *Computer Optics* **41(3)** 431-440 DOI: 10.18287/2412-6179-2017-41-3-431-440
[5] Terekhin E A 2017 Recognition of abandoned agricultural lands using seasonal NDVI values *Computer Optics* **41(5)** 719-725 DOI: 10.18287/2412-6179-2017-41-5-719-725
[6] Vorobiova N S, Sergeyev V V and Chernov A V 2016 Information technology of early crop identification by using satellite images *Computer Optics* **40(6)** 929-938 DOI: 10.18287/2412-6179-2016-40-6-929-938
[7] Fedoseev V A 2018 Hyperspectral satellite image classification using small training samples from the image itself *Journal of Physics: Conference Series* (in print)
[8] Denisova A Y and Sergeyev V V 2017 Supervised multichannel image classification algorithm using hierarchical histogram representation *Procedia Engineering* **201** 213-222
[9] Gonzalez R C and Woods R E 2009 *Digital Image Processing* (Pearson Education)

Acknowledgments

This work was supported by the Russian Foundation for Basic Research (grants 16-29-09494 ofi m, 18-07-00748 a), by Project 5-100, and the Federal Agency for scientific organizations (Agreement 007-GZ/43363/26). Satellite images of the apparatus "Resurs-P" are provided by JSC RCC "Progress".

CFA Artifacts Analysis for Image Splicing Detection

A A Varlamova[1] and A V Kuznetsov[1, 2]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086
[2]Image Processing Systems Institute - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001

Abstract. Image splicing is one of the widespread image forgery techniques. It represents pasting in image parts of other images. The paper is devoted to one of the methods of image splicing localization based on analysis of CFA artifacts that appear in an image during the capturing process. The proposed method is based on measuring a feature that characterizes the presence/absence of CFA artifacts for each image block. Obtained values of the feature define the probability of each block to be pasted. In the experimental part of the paper, we analyse the accuracy of the splicing detection method and its robustness against different types of distortions such as additive Gaussian noise, JPEG compression, and linear enhancement. The results showed that the suggested method reveals pasted regions of different shape, size, and nature in images. The method possesses stability against additive Gaussian noise and linear enhancement, but it is not steady against JPEG compression. The advantage of the method is the ability to localize splicing regions even at the smallest 2×2 block level.

1. Introduction

Today the appearance of a large number of digital devices, which allow capturing images has led to a decrease in their cost and, as a consequence, to their wide availability for each person. At the same time, a large number of software tools for image editing has also increased significantly. These tendencies have resulted in widespread image forgeries.

Nowadays, any user can make changes to the image which can be visually imperceptible. Besides, if it comes to professional forgery, the majority of existing services for verifying the authenticity of images cannot reveal it too.

There are many examples from the military and political spheres, the media, litigation, the activities of insurance companies and many other areas when forged images were used for the purpose of committing crimes or concealing any facts, or for copyright infringement, or to cause public outcry [1]. In this regard, image protection and verification of image authenticity have become the urgent problems.

Depending on the purpose of forgery, images can be subjected to such modifications as retouching and embedding of duplicates (copying areas of the image and pasting them into other areas of the same image). Methods for retouching and duplicates detection in images are considered in [2, 3] and [4, 5], respectively.

This work is devoted to detection of another commonly applied type of forgeries – photomontage. Photomontage represents pasting areas into an image that were taken from another image [6].

In some cases to protect images from this type of forgery embedding a digital watermark into an image can be done. [7]. However, this approach has several substantial disadvantages. Its application is limited: authentication can only be done by an owner of the data.

There are other solutions which do not require embedding additional information into an image. They include the methods that use the characteristics of device's sensor by which the image was obtained in order to detect the forgeries. One of such characteristics is color filter array (*CFA*) artifacts. These are local artifacts in an image caused by the presence of a CFA filter in the camera that captured the image. [8]. CFA is a mosaic of tiny color filters placed over the pixel sensors of the image sensor to capture color information. It is presented in most modern cameras. Note that CFA artifacts are unique for each camera model.

In [9] the authors describe the method of CFA artifacts detection. Under paradigm, they calculate the probability map of presence/absence CFA artifacts in the image and calculate its Fourier transform (FT). Spikes in the Fourier domain are the evidences of the map periodicity which means that the image contains CFA artifacts. With a small modification, the method can be used for detection of 256×256 forged image blocks.

Similarly, based on the fact that CFA artifacts have a periodic structure, the authors in [10] proposed an algorithm for determining the nature of images (whether they were obtained with a digital device or artificially generated). This method is also based on the analysis of the FT. The absence of CFA artifacts in an image area indicates that area has been altered or artificially generated, or a CFA filter was not used in the registering device. Due to the use of the FT, this method is applicable for detection of tampering regions on images of size 64 × 64 or greater.

Methods of this group are not robust against the forgeries when the image was reinterpolated after splicing was made. This problem will be solved in the course of further research. In this paper, this case is not considered, since it is another kind of forgeries and goes beyond the scope of the task. Also, we do not consider the case when the original image and the pasted region were obtained using the same recording device – in this case, CFA artifacts are the same.

This paper is devoted to the investigation of one of the methods for detecting pasted regions in images based on the analysis of CFA artifacts [8]. It allows detecting forgeries on areas with a minimum size of 2 × 2. The result of applying the method is a tampering probability map – a two-dimensional array, each element of which contains a probability of tampering of a corresponding local area in an image.

2. Image splicing detection based on CFA artifacts analysis

Despite the fact that image forgeries can be visually imperceptible they alter its statistical characteristics. In particular, they destroy the inter-pixel connections that arise during the process of obtaining an RGB image [11].

In most modern cameras CFA is used to produce an RGB image. There are a lot of types of CFA filters, but the most commonly used is the Bayer filter, which is shown in Figure 1.

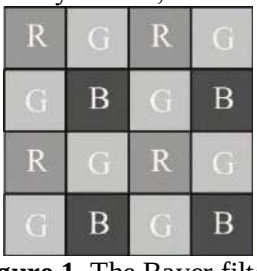


Figure 1. The Bayer filter.

Once the light passes through a CFA and a camera's sensor a RAW image is generated. A value of each pixel of a RAW image is defined for only one channel, whereas the other two values are not known. Hence only the third part of the color information is presented in a RAW image. A RAW file also contains the EXIF data – information on the date and time of photo capturing, a model of the capture device, and other parameters of recording the photo, etc.

Since for each sample of the RAW image the value of only one channel of three is determined, a demosaicing algorithm (an interpolation) is used to obtain a three-channel image. This leads to a correlation between the samples within each channel, and, as a consequence, CFA artifacts appear in the image due to the characteristics of a used camera [12].

The task of the demosaicing algorithm (demosaicing is interpolation of Bayer's templates) consists in obtaining an RGB image from the Bayer pattern. In other words, demosaicing algorithm is an interpolation of each of the three color channels in those samples where the value of the corresponding color component is unknown.

When obtaining a three-channel image by interpolation, in each channel the missing pixels values are calculated from the values of known, neighboring pixels. This process can be interpreted as a filtering process, in which the interpolation kernel (mask) is periodically applied to the original RAW image to obtain a resulting three-channel image.

There are many interpolation algorithms. The most detailed review of them is given in [9]. When calculating the missing values of pixels, the values from all channels can also be used for calculations, which lead to the appearance of interchannel connections. Further, for simplicity, we will consider the algorithm without interchannel connections. It means that missing channel values will be calculated based on known samples from the same channel only. All the calculations given in the paper are performed for the green channel, for the other two they can be produced in a similar way.

The simplest interpolation algorithm is the bilinear interpolation algorithm. Propose the bilinear interpolation algorithm was applied by the camera during the capturing process. Figure 2 shows a schematic view of the green image channel after interpolation. The pixels with known values (acquired with camera) are located in the positions A, whereas the pixels with interpolated values are located in the positions I.

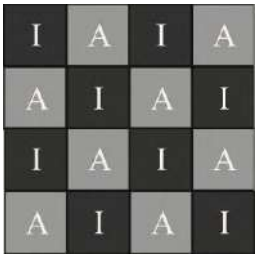


Figure 2. The green channel of the image (*A*– acquired green samples, *I* – interpolated green samples).

When using bilinear interpolation, the values of the green channel $s(x,y)$ are determined by (1):

$$s(x,y)=\begin{cases} G_A(x,y), & (x,y)\in A \\ G_I(x,y)=\sum_u\sum_v h(u,v)\times G_A(x-u)(y-v), & (x,y)\in I \end{cases}. \quad (1)$$

where $G_A(x,y)$ – values of the acquired samples in the green channel; $G_I(x,y)$ – values of the interpolated samples in the green channel; $h(u,v)$ – the interpolation kernel.

In this paper, the interpolation kernel is defined as:

$$h(u,v)=\frac{1}{4}\times\begin{bmatrix}0 & 1 & 0\\ 1 & 4 & 1\\ 0 & 1 & 0\end{bmatrix}. \quad (2)$$

Usually, color filter arrays (including the Bayer filter) have a periodical structure. Hence the correlation between samples has a periodical structure too. Forgeries destroy or alter the correlation between image samples. Thus, by analyzing the correlation of the pixels in local areas, whether the image was tampered or not can be determined.

2.1. CFA Modeling

For simplicity, the one-dimensional case will be considered, since the conclusions drawn from the calculations are also valid for the two-dimensional case.

Let $s(x)$ be a one-dimensional green channel of image that was obtained by interpolation using the Bayer filter. Then its values are determined by the equation (3):

$$s(x)=\begin{cases} G_A(x), & x(\mathrm{mod}2)=0 \\ G_I(x)=\sum_u h(u)G_A(x+u), & x(\mathrm{mod}2)\neq 0 \end{cases}, \quad (3)$$

where $G_A(x)$ – values of the acquired samples in the green channel; $G_I(x)$ – values of the interpolated samples in the green channel; $h(u)$ – the interpolation kernel.

In practice, only odd values of u (values of the acquired samples) contribute to the above summation, hence, $h(u)=0$ for odd values of u. Otherwise, $G_A(x+u)=0$ and the prediction error for the green channel can be determined by the equation (4):

$$e(x)=s(x)-\sum_u k(u)s(x+u), \quad (4)$$

where $k(u)$ – the prediction kernel.

Note that, in case the interpolation kernel $h(u)$ used by the camera is known, the prediction kernel coincides with the interpolation kernel, i.e. $k(u)=h(u)$ and there is no prediction error. In case the type of the used filter is not known, the prediction error occurs.

After the substitution of (3) in (4), the prediction error can be written as:

$$e(x)=\begin{cases} G_A(x)-\sum_u k(u)s(x+u), & x(\mathrm{mod}2)=0 \\ \sum_u h(u)G_A(x+u)-\sum_u k(u)s(x+u), & x(\mathrm{mod}2)\neq 0 \end{cases}.$$

Assume $k(u)=h(u)$, then the prediction error is equal to zero in the odd positions of x (interpolated) and differs from zero in the even positions of x (acquired with the camera). Hence, the variance of the prediction error is identically zero in the interpolated samples, whereas it differs from zero in the acquired samples.

In general, the exact interpolation coefficients may not be known, however, we can assume that $k(u)=0$ for odd u. Moreover, the equality $\sum_u k(u)=\sum_u h(u)=1$ usually holds for any used interpolation kernels.

Since only the values corresponding to odd values are meaningful, we consider only them to estimate the prediction error. Therefore, the prediction error can be expressed as follows:

$$e(x)=\begin{cases} G_A(x)-\sum_u k(u)\sum_v h(v)G_A(x+u+v), & x(\mathrm{mod}2)=0 \\ \sum_u (h(u)-k(u))G_A(x+u), & x(\mathrm{mod}2)\neq 0 \end{cases}. \quad (5)$$

By assuming that the values of the acquired samples are independent and identically distributed (i.i.d.) with the mean μ_G and the variance σ_G^2, the prediction error of the mean can be evaluated as (6):

$$E[e(x)]=\begin{cases} \mu_G-\mu_G\sum_u k(u)\sum_v h(v), & x(\mathrm{mod}2)=0 \\ \mu_G\left(\sum_u h(u)-\sum_u k(u)\right)=0, & x(\mathrm{mod}2)\neq 0 \end{cases}. \quad (6)$$

The variance of the prediction error in even samples x is calculated as (7):

$$Var[e(x)]=\sigma_G^2\left[\left(1-\sum_u k(u)h(-u)\right)^2+\sum_{t\neq 0}\left(\sum_u k(u)h(t-u)\right)^2\right]. \quad (7)$$

The variance of the prediction error in odd samples x is calculated as (8):

$$Var[e(x)]=\sigma_G^2\sum_u (h(u)-k(u))^2. \quad (8)$$

According to the above calculations, the variance of the prediction error is proportional to the variance of the acquired signal. If the prediction kernel is close to the interpolation kernel, the variance of the prediction error will be much higher at the positions of the acquired pixels than at the positions of the interpolated pixels.

2.2. Modeling of the method for image splicing detection

Thus the variance of the prediction error is higher in the acquired samples (samples A) than in the interpolated samples (samples I). This statement is also true for two-dimensional case. If the image was not obtained by applying the demosaicing algorithm or was forged, the variance of the prediction error for both types of samples will have close values within a range ε. Therefore, in order to identify the presence/absence of artifacts that arise after the application of the interpolation, it is necessary to calculate the variance of the prediction error for the samples A and I.

Let *s(x, y)* – the green channel of image, then the prediction error can be calculated by (9):

$$e(x,y) = s(x,y) - \sum_{u\neq 0}\sum_{v\neq 0} k(u,v)s(x+u, y+v). \quad (9)$$

where $k(u,v)$ – the two-dimensional prediction kernel.

Assume the used demosaicing algorithm is unknown, thus $k(u,v) \neq h(u,v)$, where $h(u,v)$ – the two-dimensional prediction kernel that was used by the camera to obtain the image.

Note that the values of the acquired samples usually are independent and identically distributed only locally, so the estimation of the local variance of the prediction error for both I and A samples is locally calculated.

Let the prediction error be local stationary within a range $(2K+1)\times(2K+1)$, $c = 1 - \sum_{i=-K}^{K}\sum_{j=-K}^{K}\alpha^2(i,j)$ – a scale factor that makes the estimator unbiased, $\mu_e = \sum_{i=-K}^{K}\sum_{j=-K}^{K}\alpha(i,j)e(x+i, y+j)$ – a local weighted mean of the prediction error, $\alpha'(i,j) = W(i,j)$ if $e(x+i, y+j)$ and $e(x,y)$ belong to the same class of samples, else $\alpha'(i,j) = 0$, W – a $(2K+1)\times(2K+1)$ Gaussian window with standard deviation $\sigma_W^2 = \frac{K}{2}$. A Gaussian window is a two-dimensional smoothing filter whose elements are distributed in accordance with the normal law of distribution. The value of the variance of W was chosen experimentally by comparison with other values from the following set: $\sigma_W^2 = \left\{K, \frac{K}{2}, \frac{K}{4}, \frac{K}{8}\right\}$.

Hence, the local weighted variance $\sigma_e^2(x,y)$ is defined by the equation (10):

$$\sigma_e^2(x,y) = \frac{1}{c}\left(\sum_{i=-K}^{K}\sum_{j=-K}^{K}\alpha(i,j)e^2(x+i, y+j) - \mu_e^2\right), \quad (10)$$

where $\alpha(i,j) = \frac{\alpha'(i,j)}{\sum_{i=-K}^{K}\sum_{j=-K}^{K}\alpha'(i,j)}$ – weights.

2.3. Feature modelling

After finding the locally-weighted variance of the prediction error, a feature characterizing the ratio between of variances of the prediction error in the acquired and interpolated samples is calculated. From the obtained values of the measure, it is possible to determine the presence/absence of CFA artifacts in the image.

Let the size of the analyzed image be $N\times N$, then we can calculate the feature for each of the disjoint image blocks of the $B\times B$ size. The size of block value must be related to the period of the Bayer filter, the smallest period and block size is 2×2. The matrix of the obtained values of the

variance of the predictor error is divided into blocks of size $B \times B$. Each block $B_{k,l}$ contains the values of variance of the acquired and interpolated samples, which we denote as $B_{Ak,l}$ and $B_{Ik,l}$, respectively, where $k,l = \overline{0,\left(\frac{N}{B}\right)-1}$.

To calculate the feature for each image block, we use the geometric mean of the locally weighted variances of the prediction errors within the selected image fragment. It is worth noting that any other averaging measure can be used to get some characterization of the "tampering" of the fragment, for example, the arithmetic mean.

Let $GM_A(k,l)$ be the geometric mean of the prediction error for A samples within the block $B_{k,l}$ and defines by (11):

$$GM_A(k,l) = \left[\prod_{i,j \in B_A(k,l)} \sigma_e^2(i,j)\right]^{\frac{1}{|B_{Ak,l}|}}, \tag{11}$$

$GM_I(k,l)$ – the geometric mean of the prediction error for I samples within the block $B_{k,l}$ and defines by (12):

$$GM_I(k,l) = \left[\prod_{i,j \in B_I(k,l)} \sigma_e^2(i,j)\right]^{\frac{1}{|B_{Ik,l}|}}, \tag{12}$$

then the measure characterizing the ratio between prediction errors in the acquired and interpolated samples can be calculated by the (13):

$$L(k,l) = \ln\left[\frac{GM_A(k,l)}{GM_I(k,l)}\right]. \tag{13}$$

If CFA artifacts are present in the image block $B_{k,l}$, which means that this block was obtained using the demosaicing algorithm, the variance will be higher in A samples. Thus, the value of measure $L(k,l)$ will be positive. However, if the image was obtained in a different way, the prediction errors of variances for the two types of samples will have close values within a range ε, since the sample values will be equally distributed and will have the same statistical characteristics. Hence, the value of $L(k,l)$ will be close to zero within the range ε.

2.4. The tampering probability map estimation. Expectation-maximization algorithm (EM algorithm)

If in the image were pasted frames from other images, in order to make the insertion more realistic, it is usually accompanied by other processes: smoothing, compression, etc. These processes destroy the traces caused by the interpolation process, that is, leads to the destruction of CFA artifacts. Therefore, the values of the feature L in the image are be non-uniform: in some areas its values are much higher than zero, which is a consequence of the presence of CFA artifacts, and in other areas where CFA artifacts are absent, the feature values are close to zero within the range ε. This fact can be used to detect forgeries in images by using values of L to find the probability of the presence of CFA artifacts in each image block $B_{k,l}$. Thus, using the obtained measure values, it is possible to determine the probability map of the presence of CFA artifacts. For this aim, the EM algorithm is used [13].

Let there are two hypotheses: M_1 – CFA artifacts are present in the image; M_2 – CFA artifacts are absent in the image. Assume the $L(k,l)$ values is Gaussian distributed under both hypotheses and for any possible size of the blocks $B_{k,l}$. For a fixed $B \times B$, we can characterize the feature using the following conditional probability density functions:

$$P\{L(k,l)|M_1\} \sim N(\mu_1, \sigma_1^2),$$

where μ_1 – mean under the truth of the hypothesis M_1, $\mu_1 > 0$;

σ_1^2 – the variance under the truth of the hypothesis M_1;

$$P\{L(k,l)|M_2\} \sim N(\mu_2, \sigma_2^2),$$

where $\mu_2 = 0$ – mean under the truth of the hypothesis M_2; σ_2^2 – the variance under the truth of the hypothesis M_2.

Assume the distribution parameters in both cases are constant. If the image obtained with the demosaicing algorithm was modified, both hypotheses will be truth for each sample, but with different probabilities. This allows to represent the feature $L(k,l)$ as a mixture of two Gaussian distributions with mean $\mu_1 > 0$ in the regions where the artifacts are present, – in the intrinsic regions, and with mean $\mu_2 = 0$ in the regions where CFA artifacts are absent, – in the forged ones.

To estimate the distribution parameters of the feature $L(k,l)$: μ_1, σ_1^2, σ_2^2, we use the EM algorithm. It is an iterative algorithm consisting of two steps at each iteration. It allows dividing the mixture of several distributions and determining their latent variables by maximizing the likelihood ratio. Knowing the parameters for each sample, the posterior probabilities of each of the hypotheses $P\{M_1|L(k,l)\}$ and $P\{M_2|L(k,l)\}$ can be determined.

At the E-step of the algorithm, the probabilities of belonging samples to each of the models are calculated. In this case, we will consider the a priori probabilities of each of the hypotheses as equals: $P\{M_1\} = P\{M_2\} = \frac{1}{2}$. Then, the probability that the block was not changed and CFA artifacts are present in it, i.e. the probability of the hypothesis M_1 is determined by the Bayes rule (14):

$$P\{M_1|L(k,l)\} = \frac{P\{L(k,l)|M_1\}}{P\{L(k,l)|M_1\} + P\{L(k,l)|M_2\}}, \quad (14)$$

where $P\{L(k,l)|M_1\}$ – the probability of the $L(k,l)$ with the truth of the hypothesis M_1; $P\{L(k,l)|M_2\}$ – the probability of the $L(k,l)$ with the truth of the hypothesis M_2.

Similarly, according to the Bayes rule, the probability of the hypothesis M_2 – $P\{M_2|L(k,l)\}$ can be calculated. It is the probability that the block was tampered and CFA artifacts are absent.

Using the calculated probabilities, the distribution parameters: μ_1, σ_1^2, σ_2^2 can be estimated. These variables are fixed at the M-step, which makes it possible to calculate the likelihood ratio by the (15):

$$\Lambda(L(k,l)) = \frac{P\{L(k,l)|M_2\}}{P\{L(k,l)|M_1\}}. \quad (15)$$

The parameters providing the maximum value of the likelihood ratio are the required parameters, and the calculated likelihood ratio values represent a tampering probability map in which each sample $\Lambda(L(k,l))$ defines a probability of presence CFA artifacts in block $B_{k,l}$, so a small value is a sign that the block was tampered.

2.5. Model validation

The performance of the method can be measured by the true positive rate R_{TP}, characterizing the rate of correctly detected tampered blocks according to the formula (16), and false positive rate R_{FP}, characterizing the rate of falsely detected blocks according to the formula (17) [14].

$$R_{TP} = \frac{N_{m_{R2}}}{N_{R_2}}, \quad (16)$$

where R_2 – the forgered region of the image; N_{R_2} – the total amount of blocks in the forgered region R_2; $N_{m_{R2}}$ – the amount of blocks detected as tampered in the region R_2.

$$R_{FP} = \frac{N_{m_{R1}}}{N_{R_1}}, \tag{17}$$

where R_1 – the untampered region of the image; N_{R_1} – the total amount of blocks in the untampered region R_1; $N_{m_{R1}}$ – the amount of blocks detected as tampered in region R_1.

Further R_{TP} and R_{FP} are used to estimate the quality of detection of tampered areas.

3. Experimental research

For the experiments, four RAW files were taken from the database [15]. The selected images were obtained with four different cameras using the Bayer filter, namely: Canon EOS 450D, Nikon D50, Nikon D90, Nikon D7000. Type of a filter used in a camera can be learned from its technical characteristics. To obtain a three-channel TIFF image from a RAW file, we used the dcraw application [16]. Two types of images were pasted into the images to form forgeries: artificially created images, whose samples are not correlated with each other and images taken from sources [17, 18], which differ in interpolation properties from the source images.

First of all, in the course of experiments, the ability of algorithm to detect artificially generated built-in regions of various nature and shape was verified. Figure 3 shows an example of an image with a forged region of an arbitrary shape and the corresponding tampering probability map computed by 8 × 8 blocks. The pasted region was obtained by the camera. It should be noted that the algorithm makes it possible to detect tampering of very small sizes, so the tampering map can be calculated by blocks with a minimum size of 2 × 2.

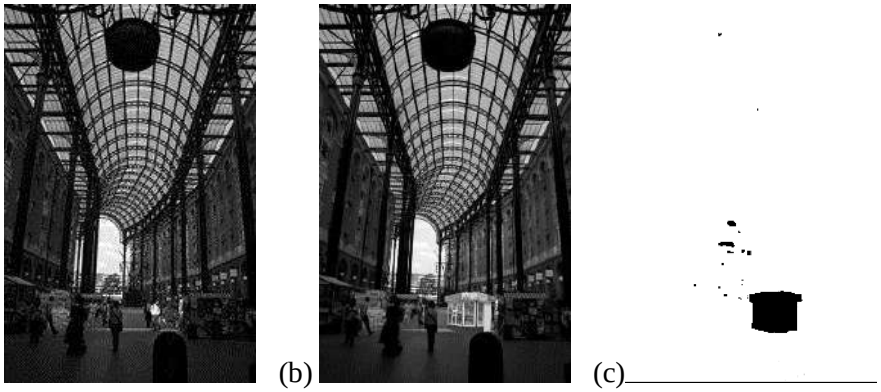


Figure 3. Examples of the algorithm performance with the size of the processed block 8 × 8: a) the source image, b) the image with a forged region of an arbitrary shape, c) tampering probability map.

Let the size of the embedded region – $M \times M$. The graphs of the dependency of rates $R_{TP}(M)$ and $R_{FP}(M)$ on the size of the built-in area are shown in Figure 4.

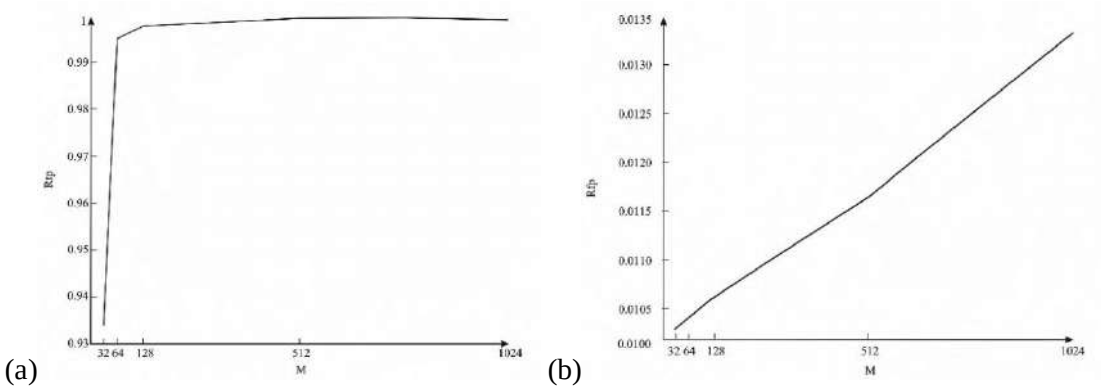


Figure 4. The dependency of quality rates on the size of the embedded region a) $R_{TP}(M)$, b) $R_{FP}(M)$.

The results of the experiment showed that with the increase in the size of the pasted area, the detection quality improves and reaches a maximum value – 1 at a size of 512 × 512. Note that with a minimum explored size of pasted region – 32 × 32 $R_{TP} = 0{,}93$, that characterizes the high quality of detection ability. In this case, the number of falsely detected unforged image blocks grows insignificantly and at a size of pasted area of 1024 × 1024 $R_{FP} = 0{,}0133$.

3.1. Investigation of method robustness against different types of distortions

To study the stability of the algorithm to various types of distortions, we used previously obtained 40 test images with a fixed size of the built-in area of 128 × 128.

As a part of the research, additive Gaussian noise with the signal-to-noise ratio (SNR (dB)): 30, 35, 40, 45, 50 was added to each of them. An example of the experiment is shown in Figure 5.

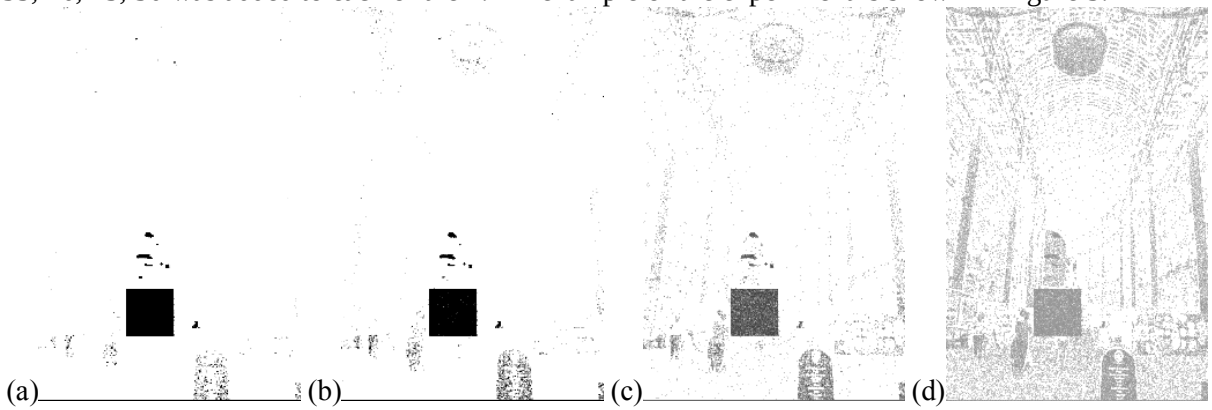


Figure 5. Tampering probability map of the forged image after adding additive Gaussian noise with SNR (dB): a) SNR = 50, b) SNR = 45, c) SNR = 40, d) SNR = 35.

To improve the visual perception of the obtained results, contrast enhancement was applied to all of the tampering probability maps.

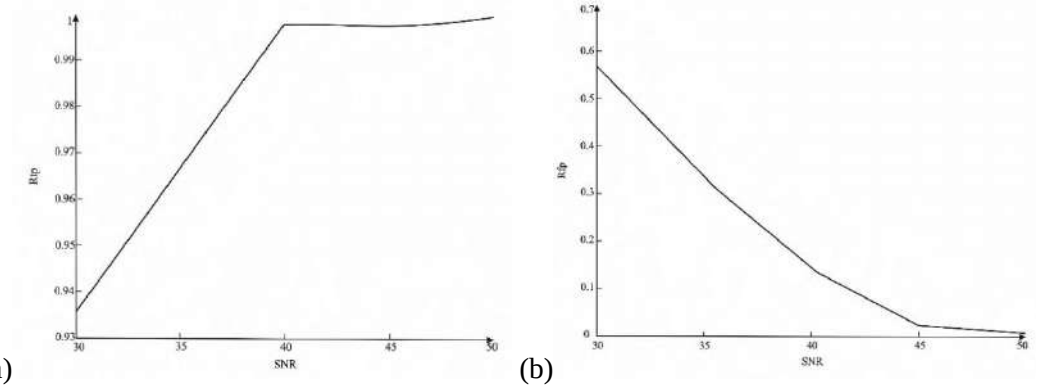


Figure 6. The dependency of quality rates on the SNR (dB): a) $R_{TP}(SNR)$, b) $R_{FP}(SNR)$.

Figure 6 shows the dependency of the quality rates of detection $R_{TP}(SNR)$ and $R_{FP}(SNR)$ on the values of SNR. The results of the experiment showed that the number of correctly detected forged blocks in the image is large for a given range of parameters, but at SNR = 35 dB or less, the number of false alarms of the algorithm increases, so we can say that the method works at values of SNR = 35 dB and above.

Further, as a part of research, JPEG compression with different values of the quality parameter Q, varying from 0 to 100, was applied to the same set of images. Example of the algorithm performance with the quality parameter values Q = 100, 98, 96, 94 is shown in Figure 7.

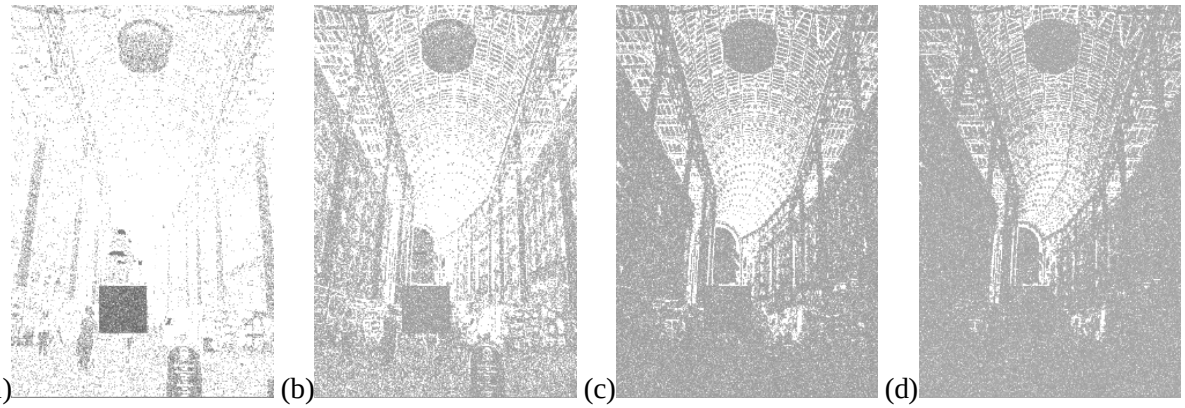


Figure 7. Tampering map of the forged region after applying JPEG compression noise with the quality parameter values Q : a) Q = 100, b) Q = 98, c) Q = 96, d) Q = 94.

Figure 8 shows the dependency of the quality rates of the detection on the value of the compression quality parameter JPEG Q: $R_{TP}(Q)$ and $R_{FP}(Q)$. From the obtained results, it can be seen that the method does not have the resistance to JPEG compression – even at high quality parameter values the number of false detections is large and already at Q = 92 $R_{FP}(Q) = 0{,}803$. Such a result can be considered a confirmation of the obvious assumptions. The use of the JPEG algorithm destroys the interpolation properties in the image, which leads to a sharp increase in the falsely detectable image fragments.

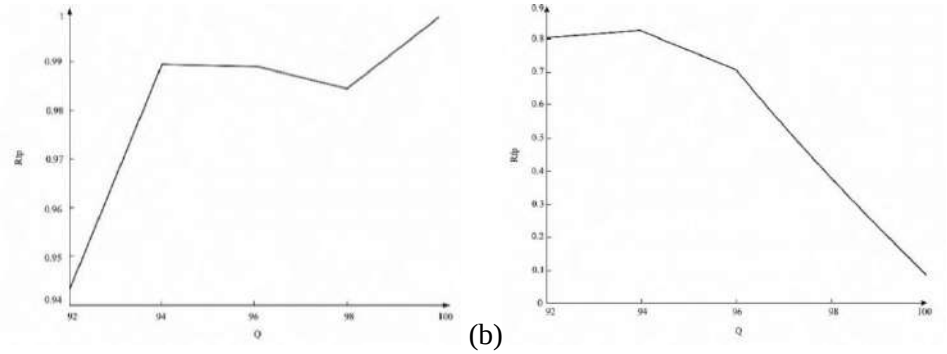


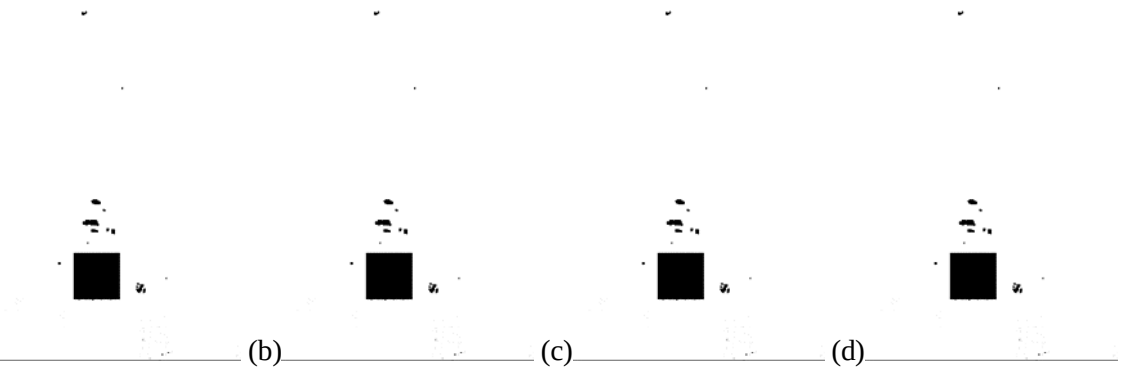
Figure 8. The dependency of quality rates on the values of Q : a) $R_{TP}(Q)$, b) $R_{FP}(Q)$.

As a part of experiments, an investigation of the stability of the method in case the JPEG compression was applied only to a distorted image region was made. It did not affect the detection result, which proves the fact that the algorithm allows to detect pasted regions of a different nature.

As a part of research, we also investigated the robustness of the method against linear contrast. The images entered into the computer are often low contrast, i.e. the variances of brightness function are small in comparison to its average value. Real dynamic range of brightness $[f_{\min}, f_{\max}]$ for such images is much less than the permissible range of brightness scale. The task of contrasting is to "stretch" the real dynamic range to the entire scale. The contrasting was implemented with the help of linear element-by-element conversion: $g = af + b$, where a, b – conversion parameters.

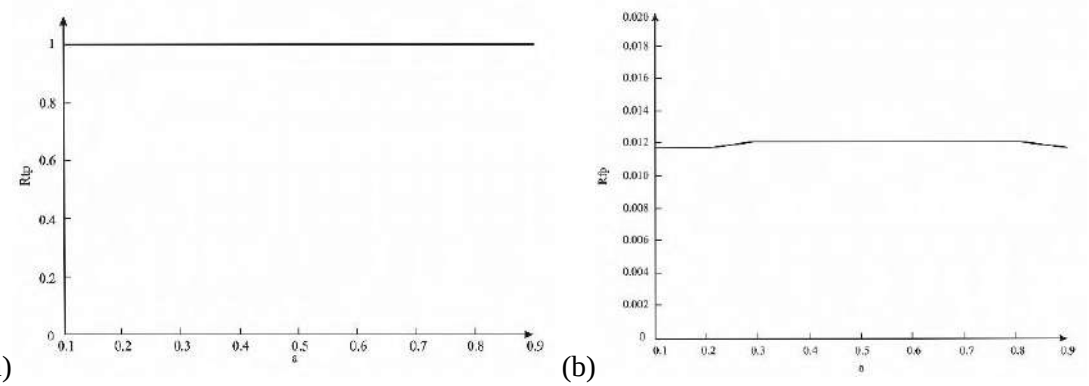
Figure 9 shows an example of how a method works if a linear contrast was applied to the forged image with different values of the transformation parameters. In the presented figure, value of the parameter b was fixed: b = 20. Figure 10 shows the dependency of the quality rates of detection on the value of the transformation parameter a: $R_{TP}(a)$ and $R_{FP}(a)$.

Similarly, an experiment was conducted in which the value of the parameter a was fixed and the parameter values b changed. The results showed that the algorithm is robust against linear contrast and the result of detection of the built-in region does not depend on the values of the parameters of linear contrast.



(a) (b) (c) (d)

Figure 9. Tampering probably map of the forged image after applying linear contrast with different values of the transformation parameter a : a) a = 0,2; b) a = 0,4; c) a = 0,6; d) a = 0,8.



(a) (b)

Figure 10. The dependency of quality rates on the values of a : a) $R_{TP}(a)$, b) $R_{FP}(a)$.

4. Conclusion

In this paper we consider the method of photomontage detection. It was established that the method allows detecting the built-in areas of various nature and form in images. As the size of the pasted area increases, the quality of detection increases, but the number of false positives increases slightly. The minimum size of the built-in area that can be detected is 2×2.

Experimental studies also showed that the algorithm is robust against such distortions as additive white Gaussian noise at values above 35 dB and linear contrast for any values of the transformation parameters. However, the method proved to be unstable to JPEG compression. Even at high values of the quality parameter, the number of false positives is large.

The method can be used to verify the authenticity of images. It allows to find pasted areas of even very small sizes, but its use is limited (it does not work for detecting pasted regions in compressed images).

5. References

[1] How to deal with fake photo reports (Access mode: https://club.esetnod32.ru/articles/analitika/kak-borotsya-s-poddelkami-fotootchetov/)

[2] Choi C, Lee H and Lee H 2013 Estimation of color modification in digital images by CFA pattern change *Forensic Science International* **226** 94-105

[3] Chakraverti A K and Dhir V 2017 Review on Image Forgery & its Detection Procedure *Journal of Advanced Research in Computer Science* **8(4)** 440-443

[4] Evdokimova N I and Kuznetsov A V 2017 Local patterns in the copy-move detection problem solution *Computer Optics* **41(1)** 79-87 DOI: 10.18287/2412-6179-2017-41-1-79-81

[5] Glumov N I, Kuznetsov A V and Myasnikov V V 2013 The algorithm for copy-move detection on digital images *Computer Optics* **37(3)** 360-367

[6] Burvin P S and Esther J M 2014 Analysis of Digital Image Splicing Detection *Journal of Computer Engineering (IOSR-JCE)* **16(2)** 10-13

[7] Snigdha K M and Ajay A G 2015 Image Forgery Types and Their Detection *Advanced Research in Computer Science and Software Engineering* **5(4)** 174-178
[8] Ferrara P, Bianchi T, Rosa A and Piva A 2012 Image Forgery Localization via Fine-Grained Analysis of CFA Artifacts *IEEE Transactions on Information Forensics and Security* **7(5)** 1566-1577
[9] Popescu A and Farid H 2005 Exposing Digital Forgeries in Color Filter Array Interpolated Images *IEEE Transactions on Signal Processing* **53(10)** 3948-3959
[10] Gallagher A and Chen T 2008 Image authentication by detecting traces of demosaicing *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* 1-8 DOI: 10.1109/CVPRW.2008.4562984
[11] Li L, Hue J, Wang X and Tian L 2015 A robust approach to detect digital forgeries by exploring correlation patterns *Pattern Analysis and Applications* **18(2)** 351-365 DOI: 10.1007/s10044-013-0319-9
[12] Bayram S, Sencar H, Memon N and Avcibas I 2005 Source camera identification based on CFA interpolation *IEEE Image Processing* **3** 63-72
[13] Bishop C M 2006 *Pattern Recognition and Machine Learning* (Springer Verlag)
[14] Fawcett T 2006 An introduction to ROC analysis *Pattern Recognition Letters* **27** 861-874 DOI: 10.1016/j.patrec.2005.10.010.
[15] *The original RAW-Samples* (Access mode: http://rawsamples.ch)
[16] *Dcraw* (Access mode: http://www.centrostudiprogressofotografico.it/en/dcraw/)
[17] *Photo database. Zermatt Matterhorn* (Access mode: http://www.zermatt.ch/ru/Media/Media-corner/Photo-database) (30.08.2017)
[18] *Columbia University Image Library (COIL-100)* (Access mode: http://www.cs.columbia.edu/CAVE/software/softlib/coil-100.php)

Acknowledgments
This work was supported by the Federal Agency of scientific organization (Agreement 007-GZ/43363/26) in part "The proposed forgery detection method" and by the Russian Foundation for Basic Research (#17-29-03190 - ofi_m) in parts "Experimental results".

Application of mixed models of random fields for the segmentation of satellite images

N A Andriyanov[1] and V E Dement'ev[1]

[1]Ulyanovsk State Technical University, Severny Venets str. 32, Ulyanovsk, Russia, 432027

Abstract. The problem of images segmentation is considered in the article. A brief overview of the existing segmentation methods is provided. We suggested to use estimates obtained in the course of nonlinear recurrent filtering for segmentation of inhomogeneous images. The proposed segmentation algorithm was investigated when working with generated images and real ones. It is shown that effective estimation of model parameters can provide the best quality of segmentation in comparison with the ISODATA algorithm. In addition, it is shown how it is possible to modify the segmentation model used to find the boundaries between objects.

1. Introduction

Recently, the problems associated with the development and research of image processing algorithms and video sequences in various machine vision systems have become especially topical. This is due to the constantly increasing volume of stored and processed digital images and the growth of the capabilities of modern computer technology. A typical example of such systems is a variety of space complexes that provide data for remote sensing of the Earth (RS). Satellite imagery is widely used for monitoring the state of the atmosphere, the surface of the oceans, polar territories, agricultural lands, urban areas, deserts and forests.

An important obstacle to the wide use of high-resolution data is the limitations in the tools used, which provide an automated analysis and interpretation of such data. One of the fundamental stages in the processing of images is their segmentation, which is carried out to divide the image into segments containing pixels similar in their visual characteristics. Each pixel is assigned a certain label (the number of the segment to which it is assigned), followed by the formation of a segment map. Such processing allows, for example, to single out on a satellite image homogeneous areas (forest, field, urban development, etc.), the subsequent analysis of which is much simpler in comparison with the study of the original heterogeneous satellite image.

2. Basic groups of segmentation algorithms

Most existing and used image segmentation techniques can be divided into the following main groups [1].

1. First group is segmentation algorithms based on visual homogeneity of the area. The methods of this group use the homogeneity criterion for obtaining connected image areas [1,2,7,8].

2. Second group includes segmentation algorithms based on the delineation of boundaries [2,3,6,10]. The methods of this group are based on the hypothesis of the discontinuity of the brightness properties of the image during the transition from one homogeneous region to another, i.e. on the

existence of the edges of regions. The edges belong to the boundaries of the regions corresponding to the segmentation result. Methods for selecting edges are divided into local and global.

3. The third group is segmentation methods based on histogram analysis [1,2,4,11]. These methods use the construction of one or more histograms for a given color image, finding histogram peaks, determining the intervals containing these peaks, and using these intervals to classify the pixels.

4. The fourth group may be considered as fuzzy segmentation. Fuzzy clustering in color or multispectral space [4,5] using, among other things, neural network solutions. The methods of this group are focused on finding the preliminary number and characteristics of homogeneous regions, for example, by fuzzy analysis of one-dimensional histograms calculated for each base color. This analysis allows you to detect peaks of histograms and at the same time determines the intervals around these peaks.

5. Segmentation algorithms based on physical properties of the image [1,2,9] is the main component of fifth group. A special feature of the methods of this group is the orientation to the selection of the area corresponding to the image of the real physical object.

3. Estimation of parameters and segmentation of simulated images

In this paper, we investigate the possibility of using not the brightness values of individual pixels, but the correlation characteristics of these images for image segmentation. To obtain such correlation characteristics, let us use the description of a segmented image using a doubly stochastic model [12-16].

So the image is considered as random field (RF) given on a rectangular multidimensional grid Ω. The method of such a representation implies that its values $x_{\bar{i}} = F(x_{\bar{j}}, \alpha_{\bar{i}}, \xi_{\bar{i}})$, where $\bar{i}, \bar{j} \in \Omega$, $\bar{j} \in D_{\bar{i}}$; $D_{\bar{i}} \in \Omega$ is model definition domain at $\bar{i}$; $F()$ is some transformation; $\alpha_{\bar{i}}$ are model parameters, which are RF independent of $\xi_{\bar{i}}$. A simple example of a doubly stochastic model is the following construction defined on a two-dimensional grid $\Omega = \{i = 1,2,..M_1; j = 1,2,..M_2\}$. Such a model uses a combination of autoregressive models with multiple roots of the characteristic equations of multiplicity 2 and 1 [17]. Thus, we can write the particular doubly stochastic model as following

$$\begin{aligned} x_{ij} &= 2\rho_{xij} x_{i-1,j} + 2\rho_{yij} x_{i,j-1} - 4\rho_{xij}\rho_{yij} x_{i-1,j-1} - \rho_{xij}^2 x_{i-2,j} - \rho_{yij}^2 x_{i,j-2} + \\ &+ 2\rho_{xij}^2 \rho_{yij} x_{i-2,j-1} + 2\rho_{yij}^2 \rho_{xij} x_{i-1,j-2} - \rho_{xij}^2 \rho_{yij}^2 x_{i-2,j-2} + \xi_{ij}, \end{aligned} \quad (1)$$

where ξ_{ij} is independent random variable (RV) with Gaussian distribution; $M(\xi_{ij}) = 0$; $M(\xi_{ij}^2) = \sigma_\xi^2$, and $\{\rho_{1ij}, i = 1,2,\ldots,M_1, j = 1,2,\ldots,M_2\}$ and $\{\rho_{2ij}, i = 1,2,\ldots,M_1, j = 1,2,\ldots,M_2\}$ are a set of correlation parameters that obey the following relations:

$$\begin{aligned} \rho_{1ij} &= r_{11}\rho_{1(i-1)j} + r_{12}\rho_{1i(j-1)} - r_{11}r_{12}\rho_{1(i-1)(j-1)} + \xi_{1ij}, \\ \rho_{2ij} &= r_{21}\rho_{2(i-1)j} + r_{22}\rho_{2i(j-1)} - r_{21}r_{22}\rho_{2(i-1)(j-1)} + \xi_{2ij}, \end{aligned} \quad (2)$$

where $\{\xi_{1ij}\}$ and $\{\xi_{2ij}\}$ are two-dimensional RFs of independent Gaussian RV with zero means and variances $M\{\xi_{1ij}^2\} = (1-r_{11}^2)(1-r_{12}^2)\sigma_{\rho 1}^2$, $M\{\xi_{2ij}^2\} = (1-r_{21}^2)(1-r_{22}^2)\sigma_{\rho 2}^2$; $\sigma_{\rho 1}^2 = M\{\rho_{1ij}^2\}$, $\sigma_{\rho 2}^2 = M\{\rho_{2ij}^2\}$. For convenience, we can denote r_{11} and r_{12} as r_{x1} and r_{x2} respectively, and r_{21} and r_{22} as r_{y1} and r_{y2}.

It should be noted that for similar autoregressive doubly stochastic models, it is possible to construct recurrent nonlinear filtration procedures that allow the estimation of both brightness properties ($\{x_{ij} : i = 1,2,..M_1; j = 1,2,..M_2\}$) and correlation properties ($\{\rho_{1ij}, \rho_{2ij} : i = 1,2,..M_1; j = 1,2,..M_2\}$).

For this we compose the following vector length $4M_1 + 5$ elements:

$$\bar{x}_{ij} = \begin{pmatrix} \bar{x}_{xij} & \bar{\rho}_{xij} & \bar{\rho}_{yij} \end{pmatrix}^T,$$

$$\bar{x}_{xij} = \begin{pmatrix} x_{i-1M_1}\, x_{ij} & x_{ij-1} & \ldots & x_{i1} & x_{i-1M_1} & \ldots & x_{i-11} & x_{i-2M} & \ldots & x_{i-2j-2} \end{pmatrix}^T,$$

$$\bar{\rho}_{xij} = \begin{pmatrix} \rho_{xij} & \rho_{xij-1} & \ldots & \rho_{xi1} & \rho_{xi-1M_1} & \ldots & \rho_{xi-1j} \end{pmatrix}^T, \bar{\rho}_{yij} = \begin{pmatrix} \rho_{yij} & \rho_{yij-1} & \ldots & \rho_{yi1} & \rho_{yi-1M_1} & \ldots & \rho_{yi-1j} \end{pmatrix}^T.$$

Then the RF model will be written in the form

$$\overline{x}_{ij} = \wp_{ij}\overline{x}_{ij-1} + \overline{\xi}_{ij},$$

where $\wp_{ij} = \begin{pmatrix} \wp_{ijx} & 0 & 0 \\ 0 & \wp_{ij\rho x} & 0 \\ 0 & 0 & \wp_{ij\rho y} \end{pmatrix}$ is matrix having size $(4M_1+5)\times(4M_1+5)$.

The first row of the matrix $\wp_{ijx}$ is equal to $\wp_{ijx}(1) = \left(2\rho_{xij-1} \;\; -\rho^2_{xij-1} \;\; 0 \;\; .. \;\; 2\rho_{yi-1j} \;\; -4\rho_{xij-1}\rho_{yi-1j} \;\; 2\rho_{xij-1}\rho^2_{yi-1j} \;\; 0 \;\; .. \;\; -\rho^2_{yij-1} \;\; 2\rho^2_{xij-1}\rho_{yi-1j} \;\; -\rho^2_{xij-1}\rho^2_{yi-1j}\right)$. The remaining rows are composed by attaching a zero column to the identity matrix.

$$\wp_{ijx} = \begin{pmatrix} 2\rho_{xij-1} & -\rho^2_{xij-1} & 0 & .. & 2\rho_{yi-1j} & -4\rho_{xij-1}\rho_{yi-1j} & 2\rho_{xij-1}\rho^2_{yi-1j} & 0 & .. & -\rho^2_{yij-1} & 2\rho^2_{xij-1}\rho_{yi-1j} & -\rho^2_{xij-1}\rho^2_{yi-1j} \\ 1 & 0 & 0 & .. & 0 & 0 & 0 & 0 & .. & 0 & 0 & 0 \\ 0 & 1 & 0 & .. & 0 & 0 & 0 & 0 & .. & 0 & 0 & 0 \\ 0 & 0 & 1 & .. & 0 & 0 & 0 & 0 & .. & 0 & 0 & 0 \\ 0 & 0 & 0 & .. & 0 & 0 & 0 & 0 & .. & 0 & 0 & 0 \\ 0 & 0 & 0 & .. & 1 & 0 & 0 & 0 & .. & 0 & 0 & 0 \\ 0 & 0 & 0 & .. & 0 & 1 & 0 & 0 & .. & 0 & 0 & 0 \\ 0 & 0 & 0 & .. & 0 & 0 & 1 & 0 & .. & 0 & 0 & 0 \\ 0 & 0 & 0 & .. & 0 & 0 & 0 & 1 & .. & 0 & 0 & 0 \\ 0 & 0 & 0 & .. & 0 & 0 & 0 & 0 & .. & 0 & 0 & 0 \\ 0 & 0 & 0 & .. & 0 & 0 & 0 & 0 & .. & 1 & 0 & 0 \\ 0 & 0 & 0 & .. & 0 & 0 & 0 & 0 & .. & 0 & 1 & 0 \end{pmatrix}$$

$$\wp_{ij\rho x} = \begin{pmatrix} r_{x1} & ... & r_{x2} & -r_{x1}r_{x2} \\ 1 & ... & 0 & 0 \\ ... & ... & ... & ... \\ 0 & ... & 1 & 0 \end{pmatrix}; \quad \wp_{ij\rho y} = \begin{pmatrix} r_{y1} & ... & r_{y2} & -r_{y1}r_{y2} \\ 1 & ... & 0 & 0 \\ ... & ... & ... & ... \\ 0 & ... & 1 & 0 \end{pmatrix}.$$

We represent these matrix relations in the form of the following formula $\overline{x}_{ij} = \varphi(\overline{x}_{ij-1}) + \overline{\xi}_{ij}$. We introduce the extrapolated estimate $\hat{\overline{x}}_{эij} = \varphi(\overline{x}_{ij-1})$ and we find the matrix $\varphi'(\overline{x}_{ij-1}) = \dfrac{\partial\varphi(\overline{x}_{ij-1})}{\partial\overline{x}_{ij-1}}$. Direct calculations show that it will be identical to the matrix $\wp_{ij}$, except for the first line, which will be equal to $\varphi'_1 = \left(A_1 \quad A_2 \quad A_3\right)$, where

$A_1 = \left(2\rho_{xij-1} \;\; -\rho^2_{xij-1} \;\; 0 \;\; .. \;\; 2\rho_{yi-1j} \;\; -4\rho_{xij-1}\rho_{yi-1j} \;\; 2\rho_{xij-1}\rho^2_{yi-1j} \;\; 0 \;\; .. \;\; -\rho^2_{yij-1} \;\; 2\rho^2_{xij-1}\rho_{yi-1j} \;\; -\rho^2_{xij-1}\rho^2_{yi-1j}\right)$,

$A_2 = \left(2x_{i-1j} - 4\rho_{yi-1j}x_{i-1j-1} - 2\rho_{xij-1}x_{i-2,j} + 4\rho_{xij-1}\rho_{yi-1j}x_{i-2,j-1} + 2\rho^2_{yi-1j}x_{i-1,j-2} - 2\rho_{xij-1}\rho^2_{yi-1j}x_{i-2,j-2} \;\; ... \;\; 0 \;\; 0\right)$,

$A_3 = \left(0 \;\; ... \;\; 2x_{ij-1} - 4\rho_{xij-1}x_{i-1j-1} - 2\rho_{xij-1}x_{i,j-2} + 4\rho_{xij-1}\rho_{yi-1j}x_{i-1,j-2} + 2\rho^2_{xij-1}x_{i-2,j-1} - 2\rho^2_{xij-1}\rho_{yi-1j}x_{i-2,j-2} \;\; 0\right)$.

A_1, A_2, A_3 are rows consisting of M_1+1 elements.

Using these relations and the method of recurrent vector filtration, we can write the following two-dimensional nonlinear filter:

$$\hat{\overline{x}}_{ij} = \hat{\overline{x}}_{эij} + B_{ij}(z_{ij} - \hat{x}_{эij}), \tag{3}$$

where $\hat{x}_{эij}$ is the first element of the vector $\hat{\overline{x}}_{эij}$; $B_{ij} = P_{эij}C^T D_{ij}^{-1}$; $C = (1,0,0,...,0)$; $D_{ij} = CP_{эij}C^T + \sigma_n^2$.

It is important that the estimated brightness and correlation characteristics are constructed as a result of a consistent evaluation of the entire image. This allows for a higher quality of evaluation than, for example, in a sliding window.

Arrays of estimates $\{\rho_{1ij}, \rho_{2ij} : i = 1,2,..M_1; j = 1,2,..M_2\}$ obtained during nonlinear filtering can be considered as two-dimensional arrays characterizing the correlation properties of the original image.

Accordingly, various processing algorithms can be applied to them, including the segmentation procedures.

Figure 1(a) shows simulated image obtained using the model (2). In this image there are two types of objects, close in brightness characteristics, but differing in correlation properties. Figure 1(b) shows field of auxiliary correlation parameters for the original image after filtering (3), and Figure 1(c) shows histogram for these correlation parameters. It shows two characteristic peaks separated by a local extremum. Using this extremum as a boundary, it is possible to perform a simple partition of the correlation parameter field and the original image into two disjoint regions. Figure 1(d) shows the results of such segmentation.

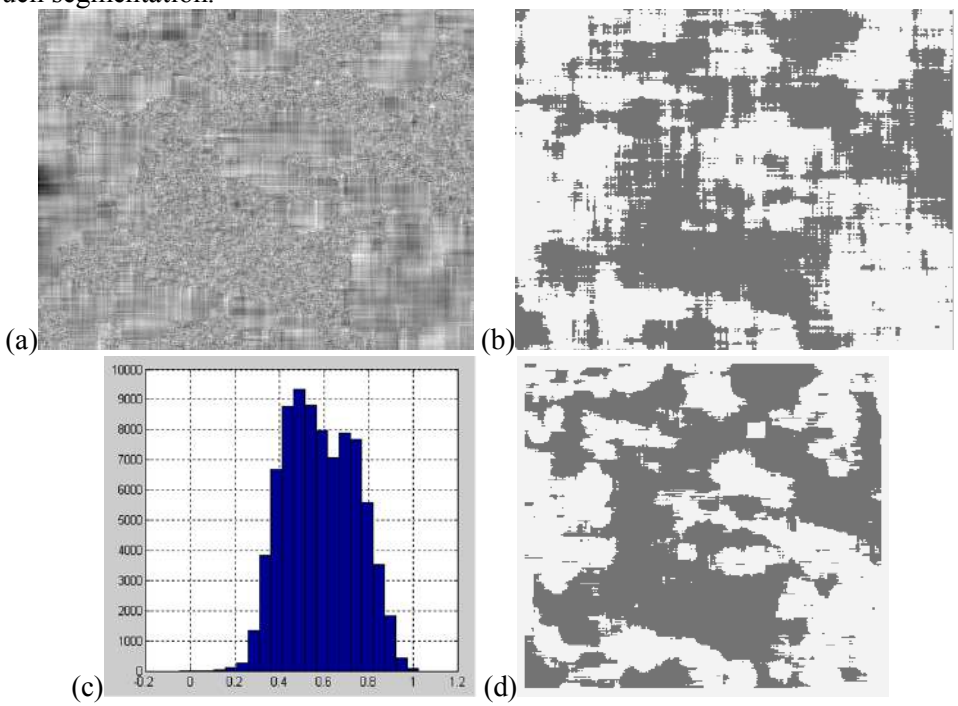


Figure 1. Segmentation of an image with varying correlation properties.

The analysis of the obtained results testifies to the high quality of the segmentation performed. About 89% of the original image points were segmented correctly. However, it should be noted that the peaks of the histogram of the correlation parameters in this case corresponded to the values of the correlation coefficients 0.5 and 0.7, respectively, i.e. which differ by approximately 27%.

Thus, the main advantage of the proposed model in comparison with the known ones is the possibility of taking into account the internal connections between the pixels during segmentation, in addition to methods based only on the brightness of specific pixels. However, the use of preprocessing requires more computational complexity than the simple application of known algorithms, such as ISODATA, k-means, MRF-segmentation.

Table 1 shows the time taken for image segmentation performed on a PC AMD-FX 4350 Quad-Core 4.2 GHz, 8 Gb RAM. The image size is 300x300.

Table 1. The complexity of segmentation.

№	Algorithm	Time, sec
1	k-means	0,5
2	ISODATA	0,78
3	Nonlinear filtering + k-means	11,2
4	Nonlinear filtering + ISODATA	11,5

Analysis of the data in Table 1 shows that the proposed algorithm reduces the processing speed, but in some cases it allows to achieve an acceptable segmentation in comparison with the known methods. Increase the speed can be using window estimation of parameters, instead of a nonlinear filter.

4. Results of segmentation of real images

Nevertheless, the result obtained for simulated images allows us to reasonably hope that in the case of applying more complex segmentation procedures in processing the correlation parameter field, the segmentation algorithm found can be applied to real images.

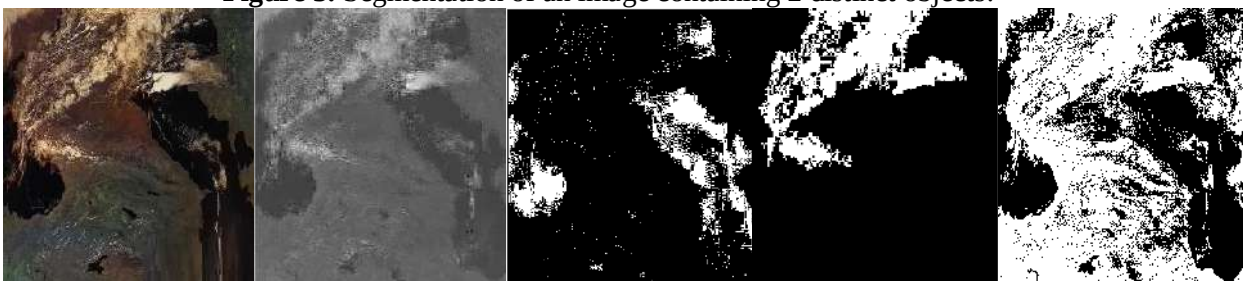
Indeed, Figures 2-7 presents the results of segmentation (binarization) of some typical images by applying a combination of the proposed algorithm and the ISODATA algorithm applied to the correlation parameter field. Figures 2(a)-7(a) show the original images, Figures 2(b)-7(b) show corresponding fields of correlation parameters, Figures 2(c)-7(c) shows results of segmentation based on correlation properties, Figures 2(d)-7(d) show segmentation results with intermediate subsampling. It means, that the estimate of the parameters was averaged over some small neighborhood. Finally, Figures 2(e)-7(e) shows the results of applying the ISODATA algorithm to the original images.

(a) (b) (c) (d) (e)

Figure 2. Segmentation of a complex color image.

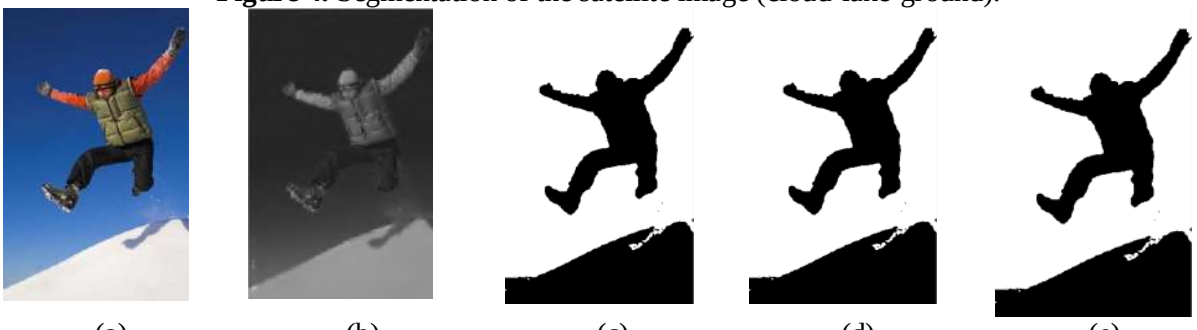
(a) (b) (c) (d) (e)

Figure 3. Segmentation of an image containing 2 distinct objects.



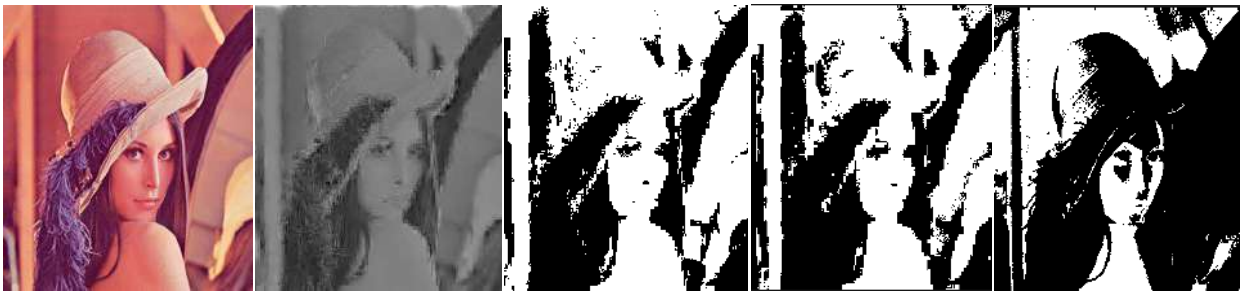
(a) (b) (c) (d) (e)

Figure 4. Segmentation of the satellite image (cloud-lake-ground).



(a) (b) (c) (d) (e)

Figure 5. Segmentation of an image with pronounced luminance characteristics of objects.



(a) (b) (c) (d) (e)

Figure 6. Segmentation of the test image Lena.

(a) (b) (c) (d) (e)

Figure 7. Segmentation of satellite image (ground-water).

The analysis of the given images, as well as the direct calculation of correctly assigned pixels, allows us to draw the following conclusions. First, the field of correlation parameters allows you to visually distinguish the objects existing on the source images. This gives base for using such a field for further processing, in particular segmentation. Secondly, in most cases (5 of 6) preliminary nonlinear filtering allowed to increase the quality of segmentation by an average of 8%. In this case, the gain is greater, the more noticeable is the difference between the correlation properties of objects in comparison with the luminance ones.

Finally, Figures 8 and 9 show the results of automatic segmentation of image data and segmentation, taking into account the selection of 2 objects based on the MRF-segmentation algorithm.

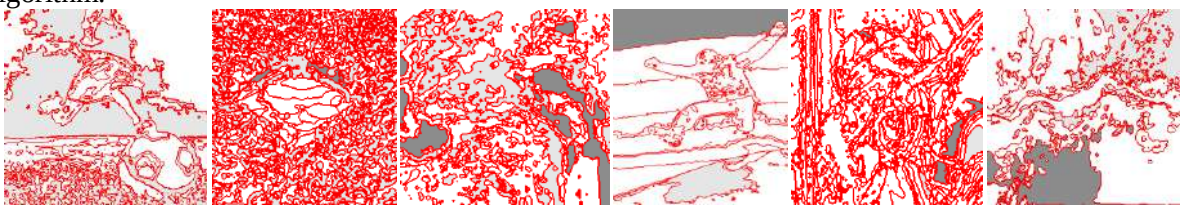


Figure 8. MRF-segmentation with auto-determination.

Figure 9. MRF-segmentation of 2 objects.

Obviously, such segmentation based on auto-determination of the number of objects is inefficient. When the algorithm indicates the number of objects of interest, then for some images, adequate segmentation is obtained, but an error remains about of 10-12% level. And in more complex images, segmentation remains unsatisfactory. Preliminary evaluation of the relationships between pixels helps either to eliminate errors in simple images, or to perform adequate segmentation (about 90%) for complex images.

5. Boundaries detection

Note that the use of the recurrent filter as a tool for obtaining and brightness and correlation characteristics of the image allows, among other things, to solve the problem of directly determining the boundaries between objects in images. To do this, in many cases it is sufficient to filter the image in the forward and backward directions $\left\{\bar{\hat{x}}_{\bar{i}} = (\hat{x}_{\bar{i}}, \hat{\alpha}_{\bar{i}}, \hat{\rho}_{1\bar{i}}, \ldots, \hat{\rho}_{N\bar{i}}, \hat{\beta}_{\bar{i}})\right\}$ and $\left\{\bar{\bar{\hat{x}}}_{\bar{i}} = (\hat{\hat{x}}_{\bar{i}}, \hat{\hat{\alpha}}_{\bar{i}}, \hat{\hat{\rho}}_{1\bar{i}}, \ldots, \hat{\hat{\rho}}_{N\bar{i}}, \hat{\hat{\beta}}_{\bar{i}})\right\}$ Then it is necessary to determine statistics L for neighboring points $(\bar{i})$ and $(\bar{j})$. It can be calculated as

$$L = \sum_{\bar{l} \in D\bar{i}} K_1(\bar{l})(\bar{\hat{x}}_{\bar{i}} - \bar{\bar{\hat{x}}}_{\bar{i}})^T + \sum_{\bar{l} \in D\bar{j}} K_2(\bar{l})(\bar{\hat{x}}_{\bar{i}} - \bar{\bar{\hat{x}}}_{\bar{i}})^T ,$$

where $K_1(\bar{l})$ and $K_2(\bar{l})$ are vector coefficients of statistics L.

In the case where L is greater than the threshold value L_0, it is decided that there is a boundary between the points $(\bar{i})$ and $(\bar{j})$. Based on the modified likelihood ratio, one can show the validity of this decision rule. The sense of the detector is related to the fact that the doubly stochastic filter, when passing the explicit boundary between two objects, each of which is described by its doubly stochastic model implementation, demonstrates a short-term increase in the variance of the estimation error. This increase is the more, then the difference between these objects is more obvious. You can detect this jump in the variance by comparing the estimates of the forward and backward filters.

Figure 10 shows artificial doubly stochastic image (see Figure 10(a)), calculated statistics L (see Figure 10(b)), fragment of a real satellite image (see Figure 10(c)) and the calculated statistics L for real image (see Figure 10(d)).

Figure 10. Selecting borders on an image using a two-pass doubly stochastic filter.

Thus, based on the proposed models, it is possible to detect boundaries between objects, which also makes it possible to improve the quality of segmentation.

6. Conclusion

Thus, the results of the conducted studies confirmed the possibility of using nonlinear recurrent filtering as an auxiliary tool that allows to improve the quality of segmentation of images of various types. This allows us to recommend this type of treatment for real machine vision systems.

7. References

[1] Gonzalez R and Woods R 2006 *Digital image processing* (Moscow: Technosphere) p 812
[2] Bakut P A and Kolmogorov G S 1987 Image Segmentation: Methods for delimiting the boundaries of regions *Foreign Radio Electronics* **10** 16-23
[3] Malik J, Belongie S, Leung T and Shi J 2001 Contour and texture analysis for image segmentation *International Journal of Computer Vision* **43(1)** 7-27 DOI: 10.1023/A:1011174803800
[4] Li S and Wu D O 2015 Modularity-based image segmentation *IEEE Transactions on Circuits and Systems for Video Technology* **25(4)** 570-581 DOI: 10.1109/TCSVT.2014.2360028
[5] Long J, Shelhamer E and Darrell T 2015 Fully convolutional networks for semantic

segmentation *IEEE Conference on Computer Vision and Pattern Recognition* 3431-3440 DOI: 10.1109/CVPR.2015.7298965
[6] Cigla C and Alatan A A 2010 Efficient graph-based image segmentation via speeded-up turbo pixels *ICIP* 3013-3016 DOI: 10.1109/ICIP.2010.5653963
[7] Fursov V A, Bibikov S A and Baida O A 2014 A thematic classification of hyperspectral images in terms of the conjugacy index *Computer Optics* **38(1)** 154-158
[8] Zimichev E A, Kazansky N L and Serafimovich P G 2014 Spatial classification of hyperspectral images using the k-means ++ clustering method *Computer Optics* **38(2)** 281-286
[9] Blokhinov Yu B, Gorbachev V A, Rakutin Yu O and Nikitin A D 2018 Development of an algorithm for the semantic segmentation of real-time aerial photographs *Computer Optics* **42(1)** 141-148 DOI: 10.18287/2412-6179-2018-42-1-141-148
[10] Belim S V and Larionov S B 2016 Algorithm of image segmentation, based on the search for communities on graphs *Computer Optics* **40(6)** 904-910
[11] Yeruslanov R V, Orekhova M N and Dubrovin V N 2015 Segmentation of images of organs of the retroperitoneal space using computed tomographic images based on the level function *Computer Optics* **39(4)** 592-599
[12] Vasil'ev K K, Dement'ev V E and Andriyanov N A 2015 Doubly stochastic models of images *Pattern Recognition and Image Analysis* **25(1)** 105-110
[13] Vasiliev K K, Dementiev V E and Andriyanov N A 2017 Filtration and restoration of satellite images using doubly stochastic random fields *CEUR Workshop Proceedings* **1814** 10-20
[14] Azanov P, Danilov A and Andriyanov N 2017 Development of software system for analysis and optimization of taxi services efficiency by statistical modeling methods *CEUR Workshop Proceedings* **1904** 232-238
[15] Vasiliev K K and Andriyanov N A 2017 Synthesis and analysis of doubly stochastic models of images *CEUR Workshop Proceedings* **2005** 145-154
[16] Andriyanov N A and Gavrilina Yu N 2018 Image Models and Segmentation Algorithms Based on Discrete Doubly Stochastic Autoregressions with Multiple Roots of Characteristic Equations *CEUR Workshop Proceedings* **2076** 19-29

Acknowledgments
The study was supported by RFBR, project № 18-31-00056.

Video-based age and gender recognition in mobile applications

A S Kharchevnikova[1] and A V Savchenko[1]

[1]National Research University Higher School of Economics, Myasnitskaya str. 20, Moscow, Russia, 101000

Abstract. In this paper we develop the age and gender recognition mobile system using deep convolutional neural networks for mobile applications. The brief literature survey on the age/gender problem in retail applications is presented. The comparative analysis of classifier fusion algorithms to aggregate decisions for individual frames is provided. In order to improve the age and gender identification accuracy we implement the video-based recognition system with several aggregation methods. We provide the experimental comparison for IJB-A, Indian Movies, Kinect and EmotiW2018 datasets. It is demonstrated that the most accurate decisions are obtained using the geometric mean and mathematical expectation of the outputs at softmax layers of the convolutional neural networks for gender recognition and age prediction, respectively. As a result, the off-line application of the proposed system is implemented on the Android platform.

1. Introduction

Due to the rapid growth of interest in video processing, the modern face analysis technologies are oriented to identify various properties of an observed person. In particular, age and gender characteristics can be applied in retail for contextual advertising for particular group of customers [1]. Often, people ignore the advertisements because the information is irrelevant, uninteresting for them at the current moment. Consequently, companies incur huge losses from investing in contextual advertising, which turns out to be inefficient and ineffective. Therefore, one of the key tasks of video analytics in retail is to provide relevant information that meets the interests of a specific target audience. For instance, depending on the automatically detected customer data, the application could provide relevant information that corresponds to a specific target audience. Consequently, the video-based age and gender recognition would improve the efficiency of the contextual advertising and increase sales. The necessary video frames for the following recognition can be obtained from digital screens or interactive panels in the shops. Such applications, running in real time, should perform the recognition task at the required speed on platforms that are limited by power and memory resources. Therefore the advanced decision for mobile platforms is required. Despite the fact that over the past few years a large number of different algorithms for age and gender recognition have appeared [2,3], the reliability of existing solutions remains insufficient for practical application [4].

Unlike traditional single-image processing systems, the video analysis lets us use addition information. For rather fast recognition algorithms one can obtain more than 100 frames of the classified object in the dynamics within a few seconds from the video stream [1,5]. It is sufficient to guarantee that at least several frames belong to the same class from the reference base. The intuition is that if each classifier makes different errors, then the total errors can be reduced by an appropriate

combination of these classifiers. Thus, this research work is intended to consider the video-based age and gender recognition task as the problem of choosing the most reliable solution using the classifier fusion (or ensemble) methods [6,7,8]. After a set of solutions for each classifier is obtained, it is necessary to implement a combining function to make a single decision. The most obvious strategy is a simple vote, in which the decision is made in favor of the class with the maximum number of predictions. This paper compares the classifier fusion obtained by traditional averaging of individual decision rules [9] with solutions based on the principle of maximum a posteriori probability [10,11,12].

The rest of the paper is organized as follows. In Section 2 the brief literature survey on age and gender image recognition is presented. In Section 3 we provide classifier fusion solutions and describe the proposed recognition scheme. Experimental results and concluding comments are presented in Section 4 and Section 5 respectively.

2. Literature Survey

The task of classifying a video image of a person is as follows. Initially, each frame $\{X(t)\}, X(t) = \|x_{uv}(t)\|, t = \overline{1,T}$ is assigned to one of the L classes by feeding the RGB matrix of pixels of a facial image $X(t)$ to the CNN [4,13]. This deep neural network should be preliminarily trained using the very large dataset of facial images with known age and gender labels. For simplicity, we assume that the video contains only one classified person with a previously selected face area on the frame, so on each image $\{X(t)\}$ the face area is detected and left. Based on this, the task of recognizing gender is a typical example of a binary classification [10]. Despite the age prediction is an example of a regression problem, in practice the highest accuracy is achieved when it is assigned to the classification problem with the definition of several age categories (L = 8 in [14]).

The challenge of automatically extracting age and gender related attributes from facial images has received increasing attention in recent few years and the huge number of recognition algorithms has been proposed. Early age estimation methods are based on the calculation of the relationships between different dimensions of facial features. A detailed survey of such algorithms is presented by Kwon [15]. Since this solution requires an accurate calculation of the facial features location, that is a fairly complex problem, they are unsuitable for raw images, video frames. Geng [16] proposes a method for automatic age recognition - AGing pattErn Subspace (AGES), the concept of which is creating an aging pattern. However, the requirements of front alignment of images impose significant restrictions on the set of input parameters. The frequency-based approach is also known among the age identification algorithms. For instance, a combination of biological features of the image is studied by Guo et al. [17] (BIF - Biologically Inspired Features).

Gender recognition for a facial image is a much more simple task, because it includes only L = 2 classes. Hence, traditionally binary classifiers can be applied. Among them, such methods as SVM [20], boosting-based algorithms, and neural networks are widely used.

Unfortunately, the accuracy of traditional methods of computer vision and pattern recognition does not meet the requirements of practical application. With regard to the effectiveness of the convolution neural networks (CNN) implementation, in particular, to classification challenges, Levi [14] provides new insights into the process of solving age and gender recognition problems by applying this method. After that, several other papers have proved the efficiency of deep CNNs in these tasks [21,24,25]. Specifically, deep VGG-16 [22], trained to recognize gender and age by image, is described in [21]. Hence, we will use this deep learning approach in order to recognize age and gender for video data.

3. Proposed Algorithm

The output of the CNN is usually obtained in the Softmax layer that provides the estimation of posterior probabilities $P(l|X(t))$ for the *t-th* frame belonging to the *l-th* class label from the reference base [27]:

$$P(l|X(t)) = \text{softmax}\, z_l(t) = \frac{\exp(z_l(t))}{\sum_{j=1}^{L}\exp(z_j(t))}, l = 1,2,...,L \quad (1)$$

where $z_l(t)$ is the output of the l-th neuron in the last (usually fully connected) layer of the neural network. The decision is made in favor of a class with a maximum a posteriori probability (MAP) (1).

Due to the influence of diverse factors such as unknown illumination, quick change of camera angle, low resolution of video camera, etc., making a decision based on the MAP approach for every frame is usually inaccurate. Therefore, we will use the fusion of decisions for individual frames to increase recognition accuracy. The review and analysis of publications in the field of data processing shows that the synthesis of classifier fusion is one of the most effective approaches to increasing the accuracy and stability of classification [24,26,27]. According to aggregation algorithms, several criteria are used, each of which is able to assign a class label after that general classification result is formed on the basis of some principle [8]. In the task of video recognition, firstly the traditional problem of automatic image recognition with CNN is solved for each incoming $X(t)$ frame and then all individual solutions are combined into one common decision for a specific video recording. The most obvious approach is to use more complex algorithms for constructing classifier fusion based on algebraic methods [8, 26]. Most of these algorithms (such as weighted majority committee, bagging and boosting [10,28]) require a sufficient representative training sample. Unfortunately, in many image recognition cases, the existing database contains an insufficient number of standards for each class. In the present paper it is proposed to use known statistical methods of synthesis solutions [8] that do not require the test sample. So, we examine the following criteria [29, 11]:

1. *Simple voting*, in which each classifier votes on the class it predicts, and the class receiving the largest number of votes is the ensemble decision., in which the final decision is made in favor of the class [6, 11]:

$$l^* = \underset{l=\overline{1,L}}{\text{argmax}} \sum_{t=1}^{T}\delta(l^*(t) - l) \quad (2)$$

2. *Arithmetical mean* of posterior probability estimates (1), i.e. the sum rule [6]:

$$l^* = \underset{l=\overline{1,L}}{\text{argmax}} \frac{1}{T}\sum_{t=1}^{T} P(l|X(t)) \quad (3)$$

3. If we follow the "naive" assumption about the independence of all frames [29], then the decision should be taken according to the *geometric mean* of posterior probabilities, or the product rule [6]:

$$l^* = \underset{l=\overline{1,L}}{\text{argmax}} \prod_{t=1}^{T} P(l|X(t)) = \underset{l=\overline{1,L}}{\text{argmax}} \sum_{t=1}^{T}\log P(l|X(t)) \quad (4)$$

In addition, we recall that the age prediction task is an essential regression problem. Hence, in this case it is possible to compute an *expected value* (mathematical expectation):

$$l^* = \sum_{l=1}^{L} P(l|X(t))\cdot l \quad (5)$$

The general data flow in the proposed video-based age and gender recognition system is presented in Fig. 1.

The first step implies supplying images from video camera to the input of the system. Isolated frames are selected from the video stream with a fixed frequency (about 10-20 times per second) in the frame selection block. Then it is important to fix and leave only the face area that is performed in the corresponding block. Face detection is conducted using the cascade method of Viola-Jones and the Haar features from the OpenCV library [31]. To speed up the work, known procedures for tracking a person identified in previous frames can be used [30,35].

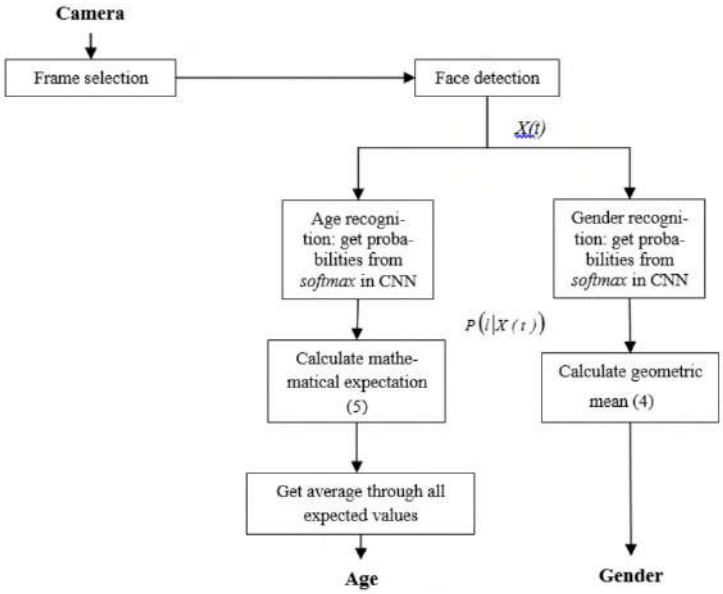


Figure 1. Proposed data flow.

At the next stage all the received images of persons on one frame are reduced to a single scale. In addition, often the subtraction of the mean image (*Image mean subtraction*) from each face image is applied [14]. The next step is supposed to provide the recognition of each frame, where the CNN is used. As the result, the estimates of posterior probabilities are obtained from the *softmax* layer (1). Recognition is performed using the model the Tensorflow library functionality. Based on the data that is the output of the classifier fusion block (2)-(4), a final recognition solution is implemented in favor of the corresponding class.

4. Experimental results

The experimental study of the proposed age/gender recognition algorithm scheme (Fig. 1) implementation is carried out in IDE Pycharm using "Python 3.6". The characteristics of the machine: Intel Core i5-2400 CPU, 64-bit operating system Windows 7, with video card NVIDIA GeForce GT 440. The described approach taking into account the experimental results is also implemented on Android platform. The user interface of this application is presented in (Fig. 2).

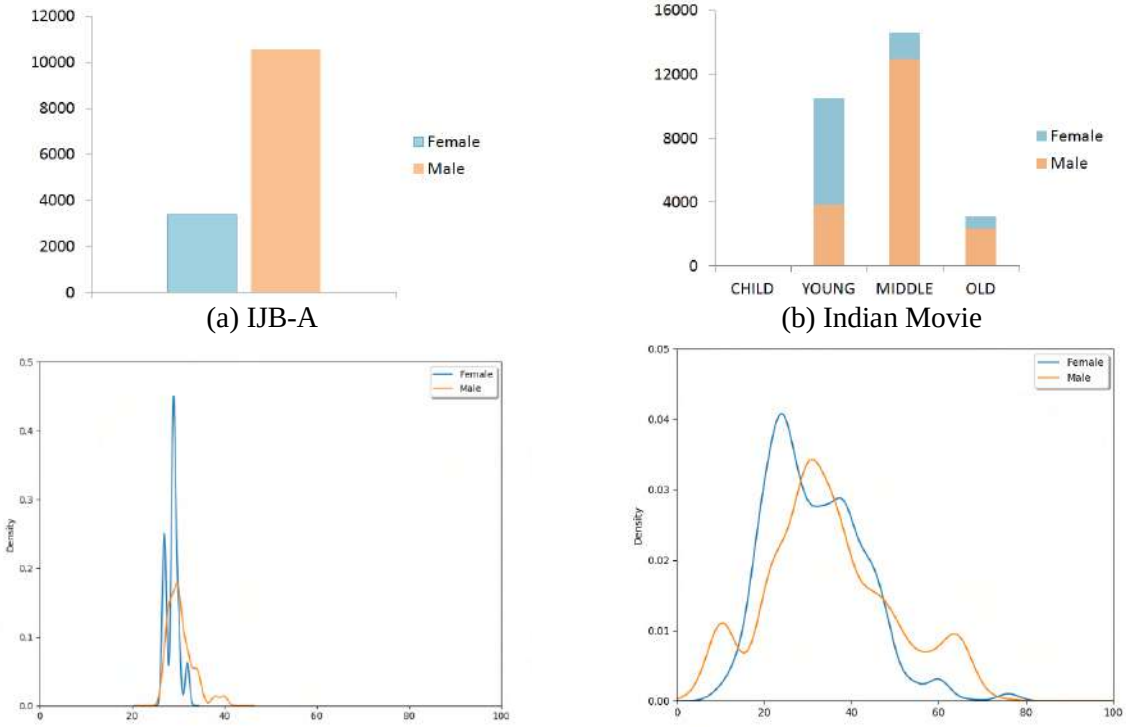
Figure 2. GUI of the proposed system.

According to the provided GUI the main part of the screen is designed to display video. There also is a text field to demonstrate the recognition and aggregation solution results in real time mode. By

default it is defined as "No face detected", this indicates that the face has not been detected. Otherwise, the gender and age of the person on the frame are displayed.

The choice of datasets for recognition accuracy and performance testing is an inherently challenging problem. The reason for this is that just the limited number of databases provides such personal information as age, gender or both about a person on an image. Furthermore the video-based approach is considered in this work. In this case the databases that are used to train and test described CNNs architectures could not be applied. Hence, in this research testing the accuracy of recognition is conducted using the facial datasets IARPA Janus Benchmark A (IJB-A) [31], Indian Movie [32], EURECOM Kinect [33] and EmotiW2018[30], which gender and age information is available for, and the video frames of a single track are stored. The first dataset consists of 2,043 videos, where only the gender information is available. The data distribution for this dataset is presented in Fig 3a.

The next database is a collection of video frames assembled from Indian films. In total there are about 332 different videos and 34,512 frames of one hundred Indian actors, whose age is divided into four categories:"Child", "Young", "Middle" and "Old". In this example, the verbal description of age is replaced by provided specific age intervals: 1-12, 13-30, 31-50, 50+ respectively. In the following experiments the intersections of the recognition results at the given intervals will be estimated. The data distribution for the dataset can be found in Fig. 3b. The Kinect dataset contains 104 videos with 52 people (14 women and 38 men). The database provides information about the gender and the year of birth that simplifies the estimation of age (Fig. 3c). The EmotiW2018 database is a collection of videos taken from various films and serials. When implementing the algorithm, age is considered in the range with addition and subtraction of 5 years, since it is necessary to identify the accuracy of the intersection with the recognized age interval. The gender of the actor on the video is provided, as well as his age. In total, the database consists of 1,165 videos. Since this dataset does not contain the final video images but video files, it became necessary to split the video into frames and subsequently detect the face area. Information about data in EmotiW2018 is presented in Fig. 3d.



(a) IJB-A (b) Indian Movie

(c) Kinect (d) Emotiw2018

Figure 3. Data distribution in experimental datasets.

We compare two publicly available CNNs architectures: Age net and Gender net models [14] and deep VGG-16 [22] neural network trained for age/gender prediction [21]. Moreover, we implement

image normalization techniques, namely, mean image subtraction to cope with illumination effects, low camera resolution, etc.

First of all, average inference time of individual CNNs for testing machine on Android platform is presented in Table 1. The best results are shown in bold.

Table 1. Average inference time and models' size.

	Time (sec)	Size (MB)
Gender/Age net	**1.076**	**43.55**
VGG-16	273	513.75

As expected, the most resource-intensive model is deep vgg16, which occupies almost 10 times more memory space than the Gender/Age nets. This fact imposes significant restrictions on the storage of such architecture, as well as on the recognition speed. It took about four minutes to make a decision for one frame using the vgg16 model.

Evaluation of the CNN quality and aggregation algorithms is carried out with the help of accuracy metric, since the problem of gender determination is a classic example of binary classification, and age recognition is considered to the investigation of several intervals-the multiclass classification. Accuracy is the proportion of correct algorithm responses.

In this paper, all proposed classifier solution methods are compared with the traditional recognition solution for each frame (2). Next, to visualize the results, we introduce the following abbreviations for discussed aggregation techniques:

FBF – *frame by frame* (2).
SV – *simple voting* (3).
SR – *sum rule* (4).
PR – *product rule* (5).
ME – *mathematical expectation* (6).

The comparison of CNN architectures and classifier fusion algorithms for gender task is presented in Table 2. The best results are in bold**.**

Table 2. Gender recognition accuracy (%).

	FBF	**SV**	**SR**	**PR**
	IJB-a			
Gender_net	51	**60**	**59**	**59**
VGG-16	72	81	81	**82**
	Kinect			
Gender_net	55	73	75	**77**
VGG-16	69	**84**	**84**	**84**
	Indian Movie			
Gender_net	61	71	72	**75**
VGG-16	75	81	87	**88**
	EmotiW2018			
Gender_net	72	**75**	**75**	**75**
VGG-16	71	78	80	**81**

Thus, based on the conducted experiments, it can be concluded that the classifier fusion implementation increases the accuracy of the gender recognition in comparison with the traditional approach. The difference is 3-10%. The product rule shows its efficiency among the aggregation algorithms almost in all cases. The deep Vgg16 is more accurate than Gender_net. For instance, the difference between models is considered to be about 9% for Kinect dataset.

Age recognition results are provided in Table 3.

Table 3. Age recognition accuracy (%).

	FBF	SV	SR	PR	ME
	Kinect				
Age_net	52	41	43	45	69
VGG-16	58	60	66	**71**	**71**
	Indian Movie				
Age_net	56	68	45	48	32
VGG-16	38	29	29	29	**54**
	EmotiW2018				
Age_net	26	27	27	27	**30**
VGG-16	47	47	48	48	**52**

The estimation of the mathematical expectation (5) has shown the effectiveness in determining the age in most cases. Thus, it could be noticed that the VGG-16 architecture is ahead of Gender net and Age net models for the age accuracy. Here we have a general trade-off between performance and accuracy. The low accuracy of age recognition can be due to the complexity of the problem as a whole, since this biometric characteristic depends on many factors and cannot always be uniquely determined.

5. Conclusion

The video-based age and gender recognition algorithm with the implementation of the classifier committees is proposed in this work. The experimental results have demonstrated the increase of recognition accuracy of the proposed algorithm when compared to traditional simple voting decision. The method of finding the geometric mean (product rule) with normalization of the input video images is the most accurate in gender classification task. At the same time, the most accurate age prediction is achieved with the computation of the expected value. We have presented the results of comparing the following CNN architectures: Age net and Gender net [14] and VGG-16 [22] trained for age and gender prediction. Eventually, the accuracy of the VGG-16 architecture is about 10-20% higher for the gender recognition and age prediction than Age and Gender net models. However, the inference time of the VGG-16 is 4-9 times lower. A limiting factor of VGG-16 practical usage has been overcome with optimization techniques [1,26,33,34]. As a result, a prototype of the age and gender recognition system (Fig. 1) for retail needs has been implemented in the Android application (Fig. 2). The intuition is that this application can improve the efficiency of contextual advertising.

6. References

[1] Savchenko A 2016 *Search techniques in intelligent classification systems* (Springer International Publishing)
[2] Chao W, Liu J and Ding J 2013 Facial age estimation based on label-sensitive learning and age-oriented regression *Pattern Recognition* **46** 628-641
[3] Rybintsev A V, Konushin V S and Konushin A S 2015 Consecutive gender and age classification from facial images based on ranked local binary patterns *Computer Optics* **39(5)** 762-769 DOI: 10.18287/0134-2452-2015-39-5-762-769
[4] Wang H, Wang Y and Cao Y, 2009 Video-based face recognition: A survey *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering* **3(12)** 2809-2818
[5] Kalinovskii I A and Spitsyn V G 2016 Review and testing of frontal face detectors *Computer Optics* **40(1)** 99-111 DOI: 10.18287/2412-6179-2016-40-1-99-111
[6] Kittler J and Alkoot F 2003 Sum versus vote fusion in multiple classifier systems *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25** 110-115
[7] Tresp V 2001 *Committee machines. Handbook for Neural Network Signal Processing* 135-151
[8] Rudakov K 1999 On methods of optimization and monotonic correction in the algebraic approach to the problem of recognition *RAS Papers* 314-317 (in Russia)

[9] Mazurov V 1990 *Method of committees in problems of optimization and classification* (Moscow: Science) p 248
[10] Theodoridis S, Koutroumbas C 2009 *Pattern Recognition* (Elsevier Inc.) p 840
[11] Savchenko A 2012 The choice of the parameters of the image recognition algorithm on the basis of the collective of decision rules and the principle of maximum a posteriori probability *Computer Optics* **36(1)** 117-124
[12] Savchenko A 2012 Adaptive Video Image Recognition System Using a Committee Machine, *Optical Memory and Neural Networks (Information Optics)* **21** 219-226
[13] Savchenko A V 2018 Trigonometric series in orthogonal expansions for density estimates of deep image features *Computer Optics* **42(1)** 149-158
[14] Levi G and Hassner T 2015 Age and gender classification using convolutional neural networks *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* 34-42
[15] Kwon Y and da Vitoria Lobo N 1994 Age classification from facial images *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 762-767
[16] Geng X 2006 Learning from facial aging patterns for automatic age estimation *Proceedings of the 14th ACM International Conference on Multimedia* 307-316
[17] Guo G, Mu G and Fu Y 2009 Human age estimation using bio-inspired features *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 112-119
[18] Choi S 2011 Age estimation using a hierarchical classifier based on global and local facial features *Pattern Recognition* **44(6)** 1262-1281
[19] Makinen E, Raisamo R 2008 Evaluation of gender classification methods with automatically detected and aligned faces *IEEE Transactions on Pattern Analysis and Machines Intelligence* **30(3)** 541-547
[20] Shan C 2012 Learning local binary patterns for gender classification on real-world face images *Pattern Recognition Letters* **33(4)** 431-437
[21] Rothe R, Timofte R and Van L 2015 Deep expectation of apparent age from a single image *Proceedings of the IEEE International Conference on Computer Vision Workshops* 10-15
[22] Simonyan K and Zisserman A 2014 Very deep convolutional networks for large-scale image recognition *Preprint arXiv:1409.1556*
[23] Szegedy C 2015 Going deeper with convolutions *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 1-9
[24] Krizhevsky A, Sutskever I and Hinton G 2012 ImageNet classification with deep convolutional neural networks *Advances in neural information processing systems* 1097-1105
[25] Esmaeili M 2007 Creating of Multiple Classifier Systems by Fuzzy Decision Making in Human-Computer Interface Systems *Proceedings of the IEEE Conference on Fuzzy Systems* 1-7
[26] Savchenko A 2002 Adaptive Video Image Recognition System Using a Committee Machine *Optical Memory and Neural Networks (Information Optics)* **21(4)** 219-226
[27] Shan C 2010 Face recognition and retrieval in video *Video Search and Mining* 235-260
[28] Lienhart R and Maydt J 2002 An extended set of Haar-like features for rapid object detection *Proceedings of the IEEE Conference on Image Processing* **1**
[29] Dhall A, Joshi J, Sikka K, Goecke R and Sebe N 2015 The more the merrier: Analysing the affect of a group of people in images *Proceedings of the 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition*
[30] Klare B 2015 Pushing the frontiers of unconstrained face detection and recognition: IARPA Janus Benchmark A *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*
[31] *IMFDB dataset* (Access mode: http://cvit.iiit.ac.in/projects/IMFDB/)
[32] *Eurecom Kinect dataset* (Access mode: http://rgb-d.eurecom.fr/)
[33] Savchenko A 2017 Maximum-likelihood dissimilarities in image recognition with deep neural networks *Computer Optics* **41(3)** 422-430 DOI: 10.18287/2412-6179-2017-41-3-422-430
[34] Rassadin A and Savchenko A 2017 Compressing deep convolutional neural networks in visual emotion recognition *CEUR Workshop Proceedings* **1901** 207-213

[35] Nikitin M Y, Konushin V S and Konushin A S 2017 Neural network model for video-based face recognition with frames quality assessment *Computer Optics* **41(5)** 732-742 DOI: 10.18287/2412-6179-2017-41-5-732-742

Acknowledgements
The work was conducted at Laboratory of Algorithms and Technologies for Network Analysis, National Research University Higher School of Economics and supported by RSF (Russian Science Foundation) grant 14-41- 00039.

Point clouds registration based on the point-to-plane approach for orthogonal transformations

A Makovetskii[1], S Voronin[1], V Kober[2], A Voronin[1] and D Tihonkih[1]

[1]Chelyabinsk State University, Bratiev Kashirinykh str. 129, Chelyabinsk, Russia, 454001
[2] Department of Computer Science, CICESE, Carretera Ensenada-Tijuana 3918, Ensenada, B.C., Mexico, 22860

Abstract. The most popular algorithm for aligning of 3D point data is the Iterative Closest Point (ICP). This paper proposes a new algorithm for orthogonal registration of point clouds based on the point-to-plane ICP algorithm for affine transformation. At each iterative step of the algorithm, an approximation of the closed-form solution for the orthogonal transformation is derived.

1. Introduction

The Iterative Closest Point (ICP) algorithm [1-5] has become the dominant method for aligning three dimensional models based purely on the geometry. For alignment it is necessary to find a geometric transformation that connects two point clouds in $\mathbb{R}^3$ by the best way with respect to the L_2 norm. The ICP algorithm consists of two main stages:

1. Searching of corresponding points (pairs) in two clouds;
2. Minimizing the error metric (variational subproblem of the ICP).

There are two basic approaches to choosing the error metric for pairs of points. Within the point-to-point approach [1], the distance between the elements of the pair in $\mathbb{R}^3$ is used. Within the point-to-plane approach [2] the distance between the point of the first cloud and the tangent plane to the corresponding point of the second cloud is used.

The key point [6] of the ICP algorithm is the search of either an orthogonal or affine transformations, best in the sense of a quadratic metric that combines two point clouds with a given correspondence between points (the variational subproblem of the ICP algorithm).

For the point-to-point metric in the case of orthogonal transformations, the solution in a closed-form was obtained by Horn [7,8]. The solution [7] is based on the use of quaternions, whereas the solution [8] uses orthogonal matrices. The solutions are linear in time with respect to the number of point pairs. The original ICP algorithm is widely used for the rigid objects registration, but it does not work well for the case of the non-rigid objects. An extension of the ICP algorithm is proposed [9], using scaling in addition to rotation and translation. A generalization of this algorithm to the case of an arbitrary affine transformation was done [10,11]. A closed-form solution to the point-to-point problem was derived [12-14].

The above mentioned approaches for solving the variational subproblem of the ICP algorithm are based on the point-to-point metric. The point-to-plane metric has been shown to perform better than the point-point metric in terms of accuracy and convergence rate [15]. A closed-form solution to the point-to-plane case for orthogonal transformations is an open problem. Instead, iterative methods

based on the linear least-squares optimization or closed-form methods for small angles only are often used [12]. Iterative solutions require an initial approximate estimate of the transformation parameters, and the iterations might converge slowly, converge to a local optimum or not converge at all.

In [16,17] a closed-form solution to the point-to-plane problem for an arbitrary affine transformation is proposed. The affine approach works well when the correspondence between point clouds is good. In this case, the affine point-to-plane method precisely reconstructs original geometric transformation for arbitrary affine transformations, in particular for orthogonal transformations [16,17]. When a correspondence between clouds is not sufficiently good, the affine approach cannot reconstructs an original orthogonal transformation.

In this paper, we propose an approximation of a closed-form solution to the point-to-plane problem for orthogonal transformation. The method is based on the closed-form solution for the affine point-to-plane problem [16,17], matrix polar decomposition and the Horn's method for calculating the nearest orthonormal matrix [8]. The proposed method does not require an initial approximate estimate. Computer simulation results are provided to illustrate the performance of the proposed method of solving the minimization problem.

2. Closed-form solution for affine point-to-plane problem

Let $P = \{p^1, \dots, p^n\}$ be a source point cloud, and $Q = \{q^1, \dots, q^n\}$ be a destination point cloud in $\mathbb{R}^3$. Suppose that the relationship between points in P and Q is given in such a manner that for each point p_i exists a corresponding point q_i. The ICP algorithm is commonly considered as a geometrical transformation for rigid objects mapping P to Q

$$Rp_i + t, \tag{1}$$

where R is a rotation matrix, t is a translation vector, $i = 1, \dots, n$.

The group of affine transformations in the dimension of three has 12 generators. It means that the affine transformation in the dimension of three is a function of 12 variables. Let us consider the ICP variational problem for an arbitrary affine transformation in the point-to-plane case. Denote by $S(Q)$ a surface constructed from the cloud Q, by $T(q^i)$ denote a tangent plane of $S(Q)$ at point q^i. Let $J(A,T)$ be the following function:

$$J(A) = \sum_{i=1}^{n} (< A\, p^i - \ q^i, n^i >)^2, \tag{2}$$

where $<\cdot,\cdot>$ denotes the inner product, A is a matrix of an affine transformation in the homogenous coordinates:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & t_1 \\ a_{21} & a_{22} & a_{23} & t_2 \\ a_{31} & a_{32} & a_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{3}$$

p^i is a point from the cloud P, n^i is the unitary normal for $T(q^i)$

$$p^i = \begin{pmatrix} p_1^i \\ p_2^i \\ p_3^i \\ 1 \end{pmatrix}, \qquad n^i = \begin{pmatrix} n_1^i \\ n_2^i \\ n_3^i \\ 0 \end{pmatrix}. \tag{4}$$

The ICP variational problem can be stated as follows:

$$\arg min_A J(A)\,. \tag{5}$$

The solution of the problem (5) is given by the following way [16,17]:

$$MA = C. \tag{6}$$

M is the coefficients matrix 12×12

$$M_j = \begin{pmatrix} m_{11}^j & m_{21}^j & m_{31}^j & m_{41}^j & m_{12}^i & m_{22}^j & m_{32}^j & m_{42}^j & m_{13}^j & m_{23}^j & m_{33}^j & m_{43}^j \end{pmatrix}, \quad j = 1, \dots, 3, \tag{7}$$

$$m_{kl}^j = \sum_{i=1}^{n} (n_j PN)_{kl}^i,\ k, l = 1, \dots, 4, j = 1, \dots, 3, \tag{8}$$

$$(n_j PN)^i = \begin{pmatrix} p_1^i n_1^i n_j^i & p_1^i n_2^i n_j^i & p_1^i n_3^i n_j^i & 0 \\ p_2^i n_1^i n_j^i & p_2^i n_2^i n_j^i & p_2^i n_3^i n_j^i & 0 \\ p_3^i n_1^i n_j^i & p_3^i n_2^i n_j^i & p_3^i n_3^i n_j^i & 0 \\ n_1^i n_j^i & n_2^i n_j^i & n_3^i n_j^i & 0 \end{pmatrix}, i = 1, \dots, n, j = 1, \dots, 3, \quad (9)$$

$$M_{3i+j} = \begin{pmatrix} m_{11}^{ij} & m_{21}^{ij} & m_{31}^{ij} & m_{41}^{ij} & m_{12}^{ii} & m_{22}^{ij} & m_{32}^{ij} & m_{42}^{ij} & m_{13}^{ij} & m_{23}^{ij} & m_{33}^{ij} & m_{43}^{ij} \end{pmatrix},$$
$$i, j = 1, \dots, 3, \quad (10)$$

$$m_{kl}^{ij} = \sum_{i=1}^{n} (p_j n_i PN)_{kl}^i, k, l = 1, \dots, 4, i, j = 1, \dots, 3, \quad (11)$$

$$(p_j n_i PN)^k = \begin{pmatrix} p_1^k n_1^k p_j^k n_i^k & p_1^k n_2^k p_j^k n_i^k & p_1^k n_3^k p_j^k n_i^k & 0 \\ p_2^k n_1^k p_j^k n_i^k & p_2^k n_2^k p_j^k n_i^k & p_2^k n_3^k p_j^k n_i^k & 0 \\ p_3^k n_1^k p_j^k n_i^k & p_3^k n_2^k p_j^k n_i^k & p_3^k n_3^k p_j^k n_i^k & 0 \\ n_1^k p_j^k n_i^k & n_2^k p_j^k n_i^k & n_3^k p_j^k n_i^k & 0 \end{pmatrix}, k = 1, \dots, n, i, j = 1, \dots, 3. \quad (12)$$

C is the coefficients column with 12 elements

$$c_j = \sum_{i=1}^{n} n_j^i \ < q^i, n^i >, j = 1, \dots, 3, \quad (13)$$

$$c_{3i+j} = \sum_{k=1}^{n} p_j^k n_i^k \ < q^k, n^k >, i, j = 1, \dots, 3. \quad (14)$$

A is the column of variables with 12 elements

$$A = (a_{11} \quad a_{12} \quad a_{14} \quad a_{14} = t_1 \quad a_{21} \quad a_{22} \quad a_{23} \quad a_{24} = t_2 \quad a_{31} \quad a_{32} \quad a_{33} \quad a_{34} = t_3)^t. \quad (15)$$

The reconstructed affine transform is done by the following formula:

$$A = M^{-1} C. \quad (16)$$

3. Polar decomposition and orthogonal transformations

A square matrix M can be decomposed into the product of an orthonormal matrix R and a positive semi-definite matrix S [5]. The matrix S is always uniquely determined. The matrix R is uniquely determined when M is nonsingular. When M is nonsingular, we can actually write directly [8]

$$M = RS, \quad (17)$$

$$R = M(M^t M)^{-\frac{1}{2}}. \quad (18)$$

The matrix $M^t M$ is positive semi-definite and symmetric. The orthogonal matrix R in (18) can be computed by the following way [8]:

$$R = MC \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{\lambda_3}} \end{pmatrix} C^t, \quad (19)$$

where C is orthogonal matrix consisting of columns, that are eigenvectors of the matrix $M^t M$. Numbers λ_i, $i = 1, \dots, 3$, are eigenvalues of the matrix $M^t M$. The formula (18) also defines [8] a nearest orthogonal matrix R for the nonsingular matrix M. It means that the formula (18) describes the projection from the group $SL(3)$ to the subgroup $SO(3)$.

4. Projection on $SO(3)$

For approximation of the exact solution of the problem (5) we propose the following method. At each step of the ICP algorithm, we project a top-left submatrix A' (size of 3×3) of a matrix A of an affine transform, computed by the formula (16), to $SO(3)$ by the formula (19). After that it is necessary to recalculate a translation $t = (t_1, t_2, t_3)^t$.

Denote by R a result of projection of a top-left submatrix 3×3 of a matrix A to $SO(3)$. Denote by N the following matrix $n \times 3$:

$$N = \begin{pmatrix} n_1^1 & n_2^1 & n_3^1 \\ & \dots & \\ n_1^n & n_2^n & n_1^n \end{pmatrix}, \quad (20)$$

denote by v the following vector-column $n \times 1$:

$$v_i =< q^i - Rp^i, n^i >. \quad (21)$$

Then the problem

$$\sum_{i=1}^{n}(< R\,p^i + t - \ q^i, n^i >)^2 = \sum_{i=1}^{n}(< t, n^i > - < \ q^i - R\,p^i, n^i >)^2 \to \min_t \quad , \quad (22)$$

is the least squares problem for the equation

$$Nt = v. \quad (23)$$

Thus we have:

$$t = (N^t N)^{-1} N^t v. \quad (24)$$

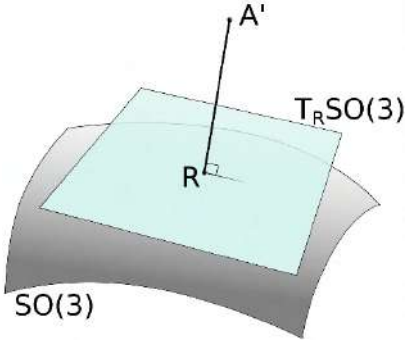


Figure 1. Illustration of projection of the top-left submatrix A' onto SO_3.

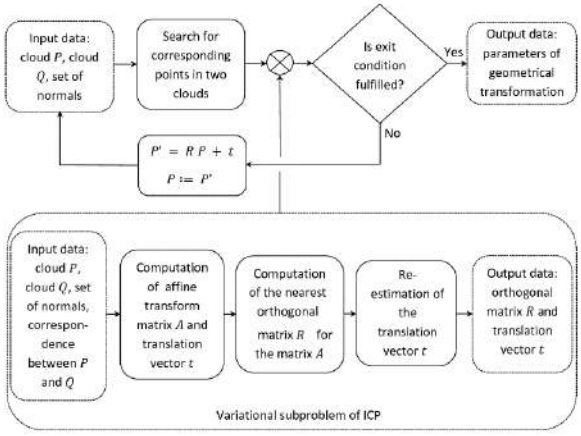


Figure 2. Block-diagram of the proposed algorithm.

Figure 1 illustrates projection of the top-left submatrix A' of the matrix A of the affine transform onto submanifold SO_3. The block-diagram of the proposed algorithm as part of the ICP algorithm is shown in Figure 2.

5. Computer simulation

5.1. We consider two variants of the ICP algorithm here

The first is point-to-point ICP based on Horn algorithm. The second is point-to-plane ICP based on the proposed approximation of an exact solution of the variational problem. Other elements of ICP algorithm are same.

5.1.1. Let P be the cloud consisting of 34817 *points, see figure 1 (blue colour)*

The cloud Q (green colour) is obtained from P by the orthogonal transformation $Q = T_1 * P$, where T_1 is given by

$$T_1 = \begin{pmatrix} 1.00000 & 0.00000 & 0.00000 & 3.10000 \\ 0.00000 & 0.83867 & -0.54464 & 1.13270 \\ 0.00000 & 0.54464 & 0.83867 & 1.92795 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{pmatrix}.$$

Computed by the proposed method transformation M_1 is given as

$$M_1 = \begin{pmatrix} 1.00000 & 0.00000 & 0.00000 & 3.10000 \\ 0.00000 & 0.83867 & -0.54464 & 1.13270 \\ 0.00000 & 0.54464 & 0.83867 & 1.92795 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{pmatrix}.$$

The reconstructed by the point-to-point ICP geometrical transformation has the same matrix. The point-to-point ICP method converges in 31 iterations, processing time 1745 milliseconds. The proposed ICP method converges in 10 iterations, processing time 913 milliseconds.

Figure 3 shows the clouds P (blue) and Q (green), figure 4 shows the clouds $P' = M_1 \cdot P$ (blue) and Q (green) together.

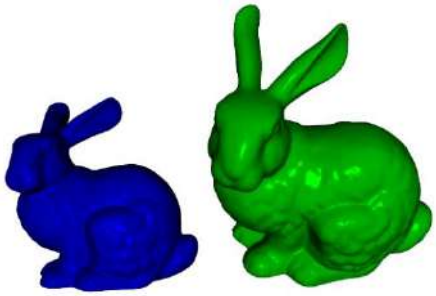


Figure 3. Cloud P (blue), cloud Q (green).

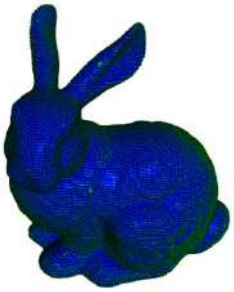


Figure 4. Cloud $P' = M_1 \cdot P$ (blue), cloud Q (green).

5.1.2. Let P be the cloud consisting of 34817 *points, see figure 3 (blue colour)*
The cloud Q (green colour) is obtained from P by the orthogonal transformation $Q = T_2 * P$, where T_1 is given by

$$T_2 = \begin{pmatrix} 0.91015 & -0.36772 & 0.19081 & -0.79646 \\ 0.21782 & 0.81653 & 0.53463 & 2.18083 \\ -0.35240 & -0.44503 & 0.82326 & 2.41239 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{pmatrix}.$$

Computed by the proposed method transformation M_2 is given as

$$M_2 = \begin{pmatrix} 0.91015 & -0.36772 & 0.19081 & -0.79646 \\ 0.21782 & 0.81653 & 0.53463 & 2.18083 \\ -0.35240 & -0.44503 & 0.82326 & 2.41239 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{pmatrix}.$$

The reconstructed by the point-to-point ICP geometrical transformation has the same matrix. The point-to-point ICP method converges in 41 iterations, processing time 2458 milliseconds. The proposed ICP method converges in 16 iterations, processing time 1491 milliseconds.

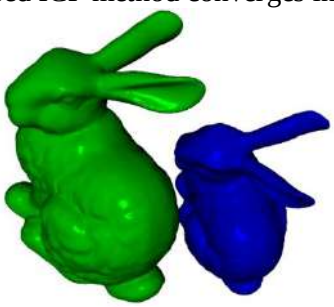


Figure 5. Cloud P (blue), cloud Q (green).

Figure 6. Cloud $P' = M_2 \cdot P$ (blue), cloud Q (green).

Figure 5 shows the clouds P (blue) and Q (green), figure 6 shows the clouds $P' = M_2 \cdot P$ (blue) and Q (green) together.

5.1.3. Let P be the cloud consisting of 34817 *points, see figure 5 (blue colour)*
The cloud Q (green colour) is obtained from P by the orthogonal transformation $Q = T_3 * P$, where T_3 is given by

$$T_3 = \begin{pmatrix} 0.98163 & 0.00000 & -0.19081 & -0.64070 \\ 0.03641 & 0.98163 & 0.18730 & 0.03261 \\ 0.18730 & -0.19081 & 0.96359 & 1.21591 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{pmatrix}.$$

Computed by the proposed method transformation M_1 is given as

$$M_3 = \begin{pmatrix} 0.98163 & 0.00000 & -0.19081 & -0.64070 \\ 0.03641 & 0.98163 & 0.18730 & 0.03261 \\ 0.18730 & -0.19081 & 0.96359 & 1.21591 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{pmatrix}.$$

The reconstructed by the point-to-point ICP geometrical transformation has the same matrix. The point-to-point ICP method converges in 19 iterations, processing time 984 milliseconds. The proposed ICP method converges in 9 iterations, processing time 747 milliseconds.

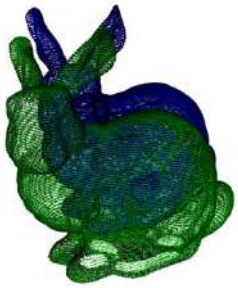


Figure 7. Cloud P (blue), cloud Q (green).

Figure 8. Cloud $P' = M_3 \cdot P$ (blue), cloud Q (green).

Figure 7 shows the clouds P (blue) and Q (green), figure 8 shows the clouds $P' = M_3 \cdot P$ (blue) and Q (green) together.

5.1.4. Let P be the cloud consisting of 106289 points, see figure 7 (blue colour)
The cloud Q (green colour) is obtained from P by the orthogonal transformation $Q = T_4 * P$, where T_4 is given by

$$T_4 = \begin{pmatrix} 0.83867 & 0.54464 & -0.00000 & 1.38331 \\ -0.45677 & 0.70337 & -0.54464 & -0.29804 \\ -0.29663 & 0.45677 & 0.83867 & 0.99881 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{pmatrix}.$$

Computed by the proposed method transformation M_4 is given as

$$M_4 = \begin{pmatrix} 0.83867 & 0.54464 & -0.00000 & 1.38331 \\ -0.45677 & 0.70337 & -0.54464 & -0.29804 \\ -0.29663 & 0.45677 & 0.83867 & 0.99881 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{pmatrix}.$$

The reconstructed by the point-to-point ICP geometrical transformation has the same matrix. The point-to-point ICP method converges in 24 iterations, processing time 6316 milliseconds. The proposed ICP method converges in 16 iterations, processing time 5792 milliseconds.

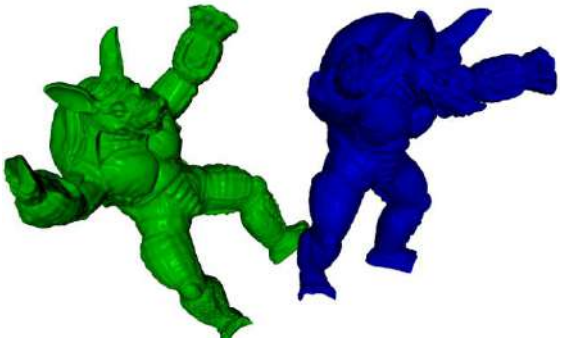


Figure 9. Cloud P (blue), cloud Q (green).

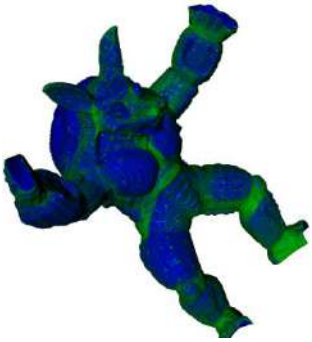


Figure 10. Cloud $P' = M_4 \cdot P$ (blue), cloud Q (green).

Figure 9 shows the clouds P (blue) and Q (green), figure 10 shows the clouds $P' = M_4 \cdot P$ (blue) and Q (green) together.

6. Conclusion

In this paper, we revised error minimizing steps of the ICP algorithm. A new algorithm for orthogonal registration of point clouds based on the point-to-plane ICP algorithm for affine transformation is proposed. At each iterative step of the algorithm, an approximation of the closed-form solution for the orthogonal transformation is derived.

7. References

[1] Besl P and McKay N 1992 A Method for Registration of 3-D Shapes *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14** 239-256
[2] Chen Y and Medioni G 1992 Object Modeling by Registration of Multiple Range Images *Image and Vision Computing* **10** 145-155
[3] Protsenko V I, Kazanskiy N L and Serafimovich P G 2015 Real-time analysis of parameters of multiple object detection systems *Computer Optics* **39(4)** 582-591 DOI: 10.18287/0134-2452-2015-39-4-582-591
[4] Myasnikov V V 2015 A local order transform of digital images *Computer Optics* **39(3)** 397-405 DOI: 10.18287/0134-2452-2015-39-3-397-405
[5] Goshin Ye V and Fursov V A 2015 3D scene reconstruction from stereo images with unknown extrinsic parameters *Computer Optics* **39(5)** 770-776 DOI: 10.18287/0134-2452-2015-39-5-770-775
[6] Turk G and Levoy M 1994 Zippered Polygon Meshes from Range Images *Computer Graphics Proceedings, Annual Conference Series* 311-318
[7] Horn B 1987 Closed-Form Solution of Absolute Orientation Using Unit Quaternions *Journal of the Optical Society of America A* **4** 629-642
[8] Horn B, Hilden H and Negahdaripour S 1988 Closed-form Solution of Absolute Orientation Using Orthonormal Matrices *Journal of the Optical Society of America Series A* **5** 1127-1135
[9] Du S, Zheng N, Ying S, You Q and Wu Y 2007 An extension of the ICP algorithm considering scale factor *Proc. 14th IEEE Internat. Conf. on Image Processing (ICIP)* 193-196
[10] Du S, Zheng N, Ying S and Liu J 2010 Affine iterative closest point algorithm for point set registration *Pattern Recognition Letters* **31** 791-799
[11] Du S, Zheng N, Meng G and Yuan Z 2008 Affine Registration of Point Sets Using ICP and ICA *IEEE Signal Processing Letters* **15** 689-692
[12] Tihonkih D, Makovetskii A and Kuznetsov V 2016 A modified iterative closest point algorithm for shape registration *Proc. SPIE Appl. of Digital Image* **9971** 99712D
[13] Tihonkih D, Makovetskii and Kuznetsov V 2016 The iterative closest points algorithm and affine transformations *Proc. Int. Conference of Analysis of Images, Social Networks, and Texts* 351-359
[14] Vokhmintsev A, Makovetskii A, Kober V, Sochenkov I and Kuznetsov V 2015 A fusion algorithm for building three-dimensional maps *Proc. SPIE's 60 Annual Meeting: Applications of Digital Image Processing XXXVIII* **9599** 959929-1
[15] Rusinkiewicz S and Levoy M 2001 Efficient Variants of the ICP Algorithm *Proceedings of the International Conference on 3-D Digital Imaging and Modeling* 145-152
[16] Makovetskii A, Voronin S, Kober V and Tihonkih D 2017 An efficient point-to-plane registration algorithm for affine transformations *Applications of Digital Image Processing* **10396** 103962J
[17] Makovetskii A, Voronin S, Kober V and Tihonkih D 2017 Affine registration of point clouds based on point-to-plane approach *Procedia Engineering* **201** 322-330

Acknowledgments

The work was supported by the Ministry of Education and Science of Russian Federation (grant № 2.1743.2017) and by the RFBR (grant № 18-07-00963).

Data organization in video surveillance systems using deep learning

A D Sokolova[1], A V Savchenko[1]

[1]National Research University Higher School of Economics, Bolshaya Pecherskaya str. 25/12, Nizhny Novgorod, Russia, 603155

Abstract. In this paper we propose to organize information in video surveillance systems by grouping the video tracks, which contain identical faces. Aggregation of the features of individual frames extracted using deep convolutional neural networks are used in order to obtain a descriptor of video track. The tracks with identical faces are grouped using the known face verification algorithms and clustering methods. We experimentally compare frame aggregation methods using the YouTubeFaces dataset and contemporary neural networks (VGGFace, VGGFace2, LightenedCNN). It is shown that the most accurate video-based face verification is achieved with the L2-normalization of average unnormalized features of individual frames of each video track. Finally, we demonstrate that the best video grouping is obtained by sequential and rank-order clustering methods.

1. Introduction

Nowadays, the automatic organization of visual information is attracting increasing attention due to the growth of the multimedia data volume. The multimedia data organization systems are required not only for a particular user who has an archive of photographs, but also for the field of public security, where video surveillance technologies are used [1, 2]. The limited functionality and the fact that the operator is physically unable to monitor the situation in real time with growth of the number of video cameras indicate the need to improve procedures and techniques, to increase the requirements for the training of operators. One approach is the use of intelligent video analytics systems to automatically process video streams.

Dramatic increase of size of collected media data leads to necessity of their grouping [3]. Existing solutions, e.g. Google Photos or Apple iPhoto, are designed to search, organize and display images of the person. However, they were not developed to process in real-time such large amount of data from video surveillance systems [4]. For example, they collect thousands of images (frames) every second [5, 6, 7]. Consequently, there is a challenge of ordering the visitors, whose faces were observed by a surveillance system [8].

Nowadays, the state-of-the-art results in image processing, object detection or feature extraction are obtained with deep convolutional neural networks (CNNs) [9, 10]. In this paper we use the clustering techniques in order to achieve automated organizing of video data where only one person is shown. As the quality of clustering mainly depends on the correctness of measuring the closeness of examined objects, in this paper we primarily focus on choosing the most appropriate representation of videos by aggregation of features obtained from each frame of video track.

The paper is organized as follows: in Section 2, we discuss the frame aggregation techniques. In Section 3, we present the proposed approach of video data organizing and our software prototype . In Section 4, the experimental results for the YouTubeFaces (YTF) dataset [9] are presented. In Section 5, the concluding comments and future plans are given.

2. Video frame aggregation techniques

The task of paper is to divide the input video sequence of $T>1$ frames into $M < T$ subsequent tracks $\{X(m)\}$, $m = 1,2,...,M$ contained face images of one person and cluster similar tracks. Each m-th track is characterized by the indices of its start $t_1(m)$ and end frame $t_2(m)$. We denote the number of frames in the m-th track as $\Delta t(m)= t_2(m)- t_1(m)+1$.

In order to group tracks contained images of one person clustering methods were used [11, 12]. To utilize them it is necessary to extract face features in each frame [13, 14], aggregate features of separate frame in descriptor for whole track [15] and then compare these descriptors. The output of the CNN's last (bottleneck) layer of the facial image in the t-th frame is stored in the D-dimensional feature vector $\mathbf{x}(t)$. These features are usually matched with the Euclidean (L_2) metric $\rho(\mathbf{x}(t_1), \mathbf{x}(t_2))$ [16]. However, when the video sequences are grouped, it is required to compute the distance $\rho(X(m_1), X(m_1))$ between tracks (subsequences of frames) $X(m_1)$ and $X(m_2)$. The most obvious way to define this distance is the computation of the mean pairwise distances between all frames:

$$\rho\left(X(m_1), X(m_2)\right) = \frac{1}{\Delta t(m_1)\Delta t(m_2)} \sum_{t=t_1(m_1)}^{t_2(m_1)} \sum_{t'=t_1(m_2)}^{t_2(m_2)} \rho\left(\mathbf{x}(t), \mathbf{x}(t')\right) \tag{1}$$

However, the run-time complexity is rather high due to the pair-wise matching of all frames in these tracks causing the computation of $\Delta t(m_1)\Delta t(m_2)$ distances between high-dimensional features. Therefore, we used the following methods to match single representations of the whole tracks.

1. The distance between their medoids:

$$\rho\left(X(m_1), X(m_2)\right) = \rho\left(\mathbf{x}^*(m_1), \mathbf{x}^*(m_2)\right),$$
$$\mathbf{x}^*(m_i) = \underset{\mathbf{x}(t), t\in[t_1(m_i), t_2(m_i)]}{\operatorname{argmin}} \sum_{t'=t_1(m_i)}^{t_2(m_i)} \rho\left(\mathbf{x}(t), \mathbf{x}(t')\right) \tag{2}$$

2. Average features vectors of each track are matched:

$$\rho\left(X(m_1), X(m_2)\right) = \rho\left(\bar{\mathbf{x}}(m_1), \bar{\mathbf{x}}(m_2)\right),$$
$$\bar{\mathbf{x}}(m_i) = \frac{1}{\Delta t(m_i)} \sum_{t'=t_1(m_i)}^{t_2(m_i)} \mathbf{x}(t) \tag{3}$$

3. Comparison of the median features $\mathbf{x}'(m_i)$ of each track:

$$\rho\left(X(m_1), X(m_2)\right) = \rho\left(\mathbf{x}'(m_1), \mathbf{x}'(m_2)\right). \tag{4}$$

4. Learnable pooling of video features of the m-th track using the neural aggregation network [17], which includes special attention blocks to process features of all frames. The goal of this block is to learn the D-dimensional vector $\mathbf{q}$, which is used to assign each t-th frame with its weight $a(t)$ using the softmax function:

$$a(t) = \frac{\exp(\mathbf{q}^T \mathbf{x}(t))}{\sum_{t'=t_1(m)}^{t_2(m)} \exp(\mathbf{q}^T \mathbf{x}(t'))} \tag{5}$$

The final representation of the m-th track is computed as the weighted average of all frames with weights (5):

$$\mathbf{r}(m) = \sum_{t=t_1(m)}^{t_2(m)} a(t)\mathbf{x}(t) \tag{6}$$

In order to improve the quality of such learnable pooling it is significant to use two sequential attention blocks [17]. Let $\mathbf{q}^{(0)}$ be the first block weights and $\mathbf{r}^{(0)}$ be the first aggregating features using $\mathbf{q}^{(0)}$ weights (5), (6). The weights in the sequential block are computed as follows:

$$\mathbf{q}^{(1)} = \tanh(W\mathbf{r}^{(0)} + b), \quad (7)$$

where W and b are the learnable weight matrix and bias vector of the neurons respectively. The feature vector $\mathbf{r}^{(1)}$ generated by $\mathbf{q}^{(1)}$ using (6) will be the final descriptor of the m-th frame.

In order to make the video features more resistant to the conditions of observation (camera resolution, illumination, etc.) usually normalization in the Euclidean metric is applied [10]. Usually, the preliminary normalization of the features of each frame is performed. However, in this paper we also analyze the normalization of aggregated video features [18].

3. Proposed video data organizing system

We implemented a special software prototype using PyCharm from JetBrains (Python 3.6 language), OpenCV [19], Caffe and TensorFlow libraries. The data flow in our system is presented in Figure 1.

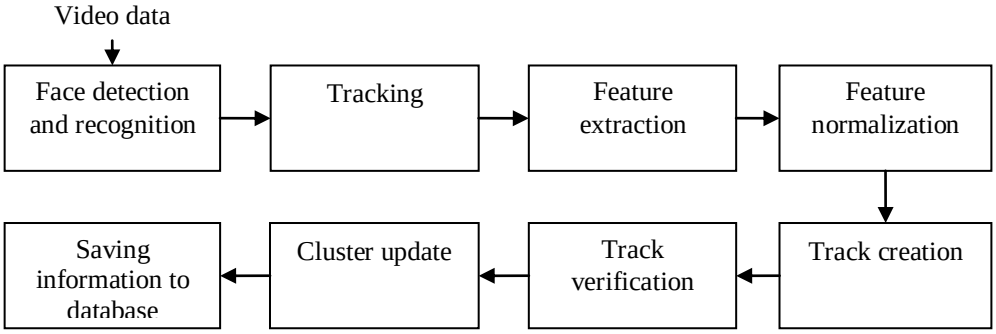


Figure 1. Dataflow of data organization in video surveillance systems.

At first, the faces are detected in each frame with the TensorFlow Models [20]. This repository contains different preliminary trained models of neural networks and provides interface for object detection (TensorFlow Object Detection API). In particular, faces are detected with the MobileNet SSD trained on the WiderFace dataset [21]. Further tracking of highlighted faces is processed but face detection is repeated periodically in order to: 1) precise tracking results; 2) obtain new faces and 3) mark disappeared faces. After that we extract features using particular CNN and normalize the features vector. Then consecutive frames of one person are united into one track (homogeneous segment). On the final step (Fig.1) subsequent clustering is processed: features vector of the last track is matched with the features of previously detected clusters. If the distance to the nearest cluster does not exceed a certain threshold, this track is added to the cluster and the information about the last is updated. The resulted set of selected clusters of homogeneous tracks is saved into NoSQL database Cassandra.

The architecture of the proposed software is shown in Fig. 2.

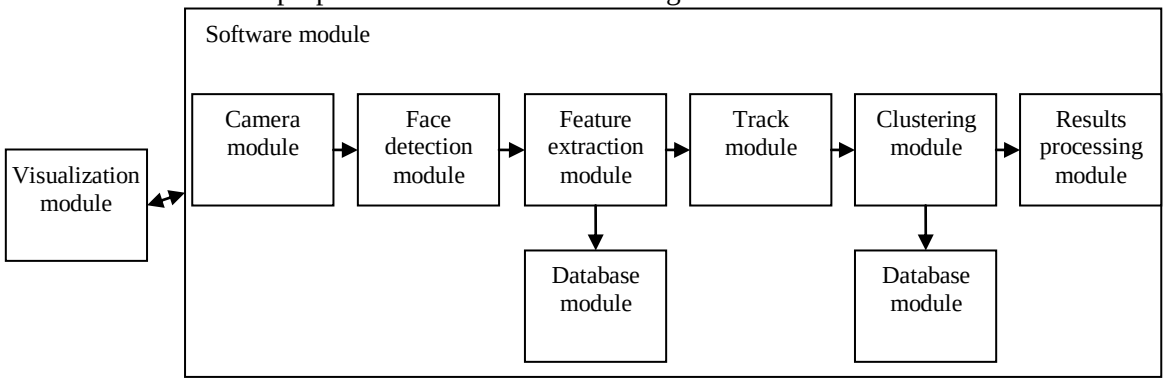


Figure 2. The architecture of the proposed software.

Using the cross-platform Qt framework for software development a graphical user interface was created (Fig. 3).

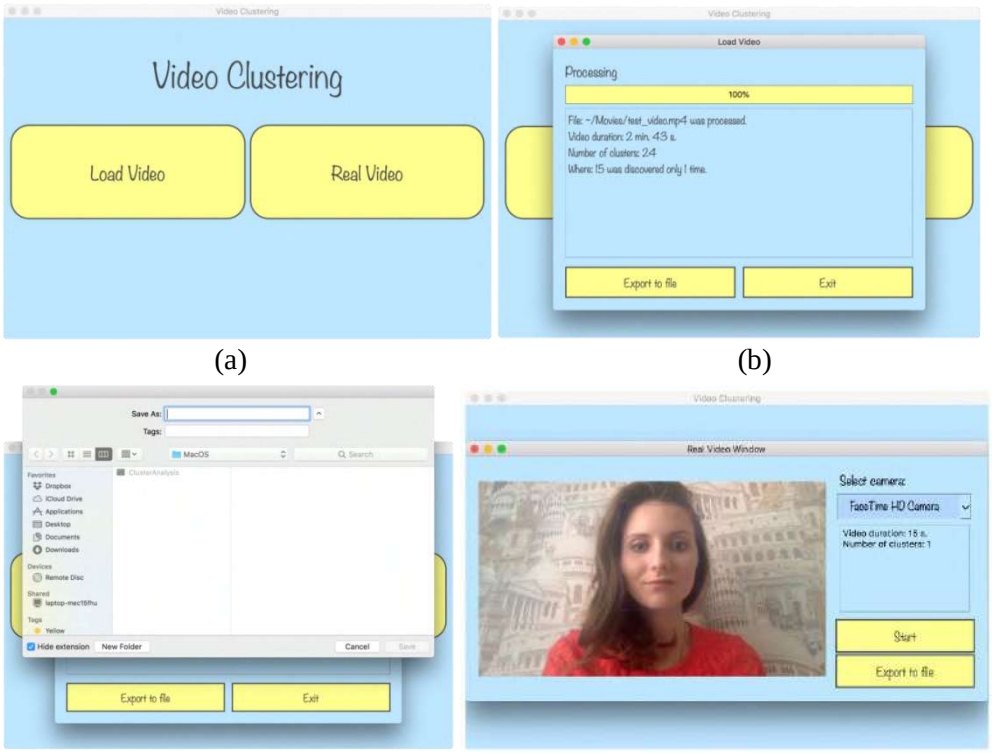


Figure 3. Results of working system (*a* – the main window, *b* – loading video processing and results, *c* – export data to file, *d* – real video processing and results).

4. Experimental results

In this section we describe the experimental results of our pipeline (Fig. 1). To extract features we used the Caffe framework [9] and three publicly available CNNs suitable for face recognition, namely, the VggNet [22], Lightened CNN (version C) [10], VggFace2 [23]. The VggNet extracts $D = 4096$ features vector in the output of "fc7" layer from 224x224 RGB images. The Lightened CNN extracts $D = 256$ features vector ("eltwise fc2" layer) is computed from 128x128 grayscale image. The VggFace2 is the ResNet50 model, which extracts $D = 2048$ features vector from "pool5_7x7_s1". Their advantages are high velocity of image processing and high accuracy of detection.

In this paper we analysed two types of distance between frames[24, 25]: traditional L2 (Euclidiean) metric and the Student criterion (t-test):

$$t = \frac{\rho(X(m_1), X(m_2))}{\sqrt{\frac{D(m_1)}{\Delta t(m_1)} + \frac{D(m_2)}{\Delta t(m_2)}}} \quad (8)$$

Here we used the YTF dataset [26], which contains 3,425 videos of 1,595 different people. The shortest track duration is 48 frames, the longest track contains 6,070 frames, and the average length of a video clip is 181.3. We have calculated following indexes: AUC (Area under curve) and FRR (False Reject Rate) for fixed FAR (False Accept Rate) = 1%. Weights in attention modules (5)-(7) were trained on 1895 videos of 500 subjects from the IJB-A dataset.

The AUC (area under curve) and the false reject rate (FRR) for 1% false accept rate (FAR) for Lightened CNN, VggNet and VggFace2 features are shown in Table 1, Table 2 and Table 3, respectively. Average face detection time for one frame is 60 ms at MacBook Pro laptop.

According to these tables, one can clearly see that the proper normalization plays a significant role. The most efficient algorithm is computation of average vector of normalized features. AUC of average vector search is more by 10-12% than AUC of medoid (2) comparison and by 14-16% than AUC of aggregation module (5)-(7) training.

Table 1. Video-based face verification results, Lightened CNN.

Measure	Distance between tracks	AUC (%)	FRR@FAR=1%
Distance (1)	L_2	90.7±0.6	77.0±8.4
L_2-norm->Distance (1)	L_2	98.2±0.4	14.1±3.6
Medoids (2)	L_2	89.7±0.6	80.6±6.4
	t-test (8)	84.7±0.7	72.9±7.8
L_2-norm (2) medoids	L_2	97.2±0.6	19.1±4.3
	t-test (8)	88.8±0.6	54.1±5.9
Attributes averaging(3)	L_2	91.3±1.3	71.8±10.0
	t-test (8)	91.8±1.4	72.3±11.5
L_2-norm attributes averaging (3)	L_2	97.7±0.5	21.4±6.4
	t-test (8)	96.8±0.5	37.2±7.6
L_2-norm of average attributes vector (3)	L_2	98.3±0.7	12.4±3.1
	t-test (8)	97.6±0.5	12.5±3.1
L_2-norm of median (4)	L_2	96.7±0.6	22.3±7.2
	t-test (8)	94.4±0.5	37.0±7.5
Attention module (5)-(7)	L_2	87.6±0.9	65.3±6.5
	t-test (8)	89.3±1.2	64.8±7.0
L_2-norm of attention module (5)-(7)	L_2	90.2±0.7	66.7±5.4
	t-test (8)	90.4±0.8	54.5±9.3

Table 2. Video-based face verification results, VggNet.

Measure	Distance between tracks	AUC (%)	FRR@FAR=1%
Distance (1)	L_2	83.3±0.8	85.8±9.0
L_2-norm->Distance (1)	L_2	97.9±0.6	23.2±6.3
Medoids (2)	L_2	85.7±1.8	86.0±8.4
	t-test (8)	80.8±1.2	83.9±7.7
L_2-norm (2) medoids	L_2	93.5±1.1	25.4±6.2
	t-test (8)	85.2±0.7	69.9±7.9
Attributes averaging(3)	L_2	89.1±0.9	79.7±7.8
	t-test (8)	87.4±1.2	81.2±5.8
L_2-norm attributes averaging (3)	L_2	97.2±0.6	54.4±6.3
	t-test (8)	96.3±0.7	76.9±6.8
L_2-norm of average attributes vector (3)	L_2	98.1±1.0	19.4±5.9
	t-test (8)	97.7±0.6	25.3±7.8
L_2-norm of median (4)	L_2	96.2±0.7	32.4±6.5
	t-test (8)	94.8±0.7	41.1±7.3
Attention module (5)-(7)	L_2	87.8±0.4	54.3±5.8
	t-test (8)	89.7±0.8	52.7±6.0
L_2-norm of attention module (5)-(7)	L_2	90.1±0.7	46.9±5.9
	t-test (8)	90.0±0.6	44.5±6.4

Table 3. Video-based face verification results, VggFace2.

Measure	Distance between tracks	AUC (%)	FRR@FAR=1%
Distance (1)	L_2	98.5±0.5	94.4±3.3
L_2-norm->Distance (1)	L_2	99.0±0.6	94.4±3.2
Medoids (2)	L_2	96.0±0.4	94.9±4.5
	t-test (8)	87.7±0.6	94.6±3.5
L_2-norm (2) medoids	L_2	98.4±0.7	94.9±3.6
	t-test (8)	86.2±0.7	94.6±4.1
Attributes averaging(3)	L_2	98.8±0.6	93.0±4.4
	t-test (8)	98.0±0.5	93.1±3.8
L_2-norm attributes averaging (3)	L_2	99.1±0.6	92.1±4.0
	t-test (8)	95.8±0.8	92.3±5.5
L_2-norm of average attributes vector (3)	L_2	98.5±0.5	92.9±6.3

L_2-norm of median (4)	t-test (8)	94.0±0.6	93.1±4.8
	L_2	94.9±0.7	94.6±5.6
Attention module (5)-(7)	t-test (8)	98.5±0.4	94.7±3.0
	L_2	88.4±0.6	95.7±3.4
L_2-norm of attention module (5)-(7)	t-test (8)	89.9±0.8	92.6±6.2
	L_2	91.7±0.7	93.4±4.5
	t-test (8)	92.2±0.7	94.4±4.5

Also we implemented sequential clustering where threshold for resulting clusters is set by fixing the FAR value. In addition, we examined the clustering algorithm [27] from the DominantSet library [28] and the Rank-Order hierarchical clustering [29]. The results are presented in Table 4.

Table 4. Clustering results.

	Lightened CNN		**VggNet**		**VggFace2**	
	Total quantity of clusters	**Quantity of incorrect clusters**	**Total quantity of clusters**	**Quantity of incorrect clusters**	**Total quantity of clusters**	**Quantity of incorrect clusters**
Sequential clustering, FAR=1%	2492	35	2634	44	2102	23
Sequential clustering, FAR=10%	2147	162	2195	171	1956	121
DominantSet	1836	180	1902	225	1801	177
Rank-Order	2105	36	2213	42	2007	31

Total quantity of clusters is bigger than quantity of different people from YTF dataset (1,595) because different videos with one person could be mentioned in different clusters. Moreover, average processing time of YTF tracks for hierarchical algorithm is 8 minutes, while algorithm DominantSet utilized for video tracks grouping consumed more than 20 minutes. The most effective results were demonstrated by Rank-Order clustering algorithm. As usual, the most accurate results are obtained for the VggFace2 facial features.

5. Conclusion

In this paper we solved the problem of video subsequences clustering for video surveillance systems. In particular, we focused on calculating the degree of proximity of video tracks using the aggregation of features vectors extracted by deep CNNs. Experimental study demonstrated that the features vectors averaging of all frames and subsequent normalization lead to the highest accuracy of video face verification. In the future work we plan to analyze other clustering algorithms deeper in order to achieve low calculation complexity and high accuracy of data processing.

6. References

[1] Chellappa R, Du M, Turaga P and Zhou S K 2011 Face Tracking and Recognition in Video *Handbook of Face Recognition* 323-351
[2] Shan C 2016 Face Recognition and Retrieval in Video *Video Search and Mining, Studies in Computational Intelligence* **287** 235-260
[3] Savchenko A V 2016 Search Techniques in Intelligent Classification Systems *Springer International Publishing*
[4] Savchenko A V and Belova N S 2018 Unconstrained Face Identification Using Maximum Likelihood of Distances Between Deep Off-the-shelf Features *Systems with Applications* **108** 170-182

[5] Chen J C, Ranjan R, Kumar A, Chen C H, Patel V M and Chellappa R 2015 An end-to-end system for unconstrained face verification with deep convolutional neural networks *IEEE International Conference on Computer Vision Workshops* 118-126
[6] Li H, Hua G, Shen X, Lin Z and Brandt J 2014 Eigen-PEP for video face recognition *Asian Conference on Computer Vision, LNCS* **9005** 17-33
[7] Savchenko A V 2017 Deep neural networks and maximum likelihood search for approximate nearest neighbor in video-based image recognition *Optical Memory and Neural Networks (Information Optics)* **26(2)** 129-136
[8] Sokolova A D, Kharchevnikova A S and Savchenko A V 2017 Organizing Multimedia Data in Video Surveillance Systems Based on Face Verification with Convolutional Neural Networks *Lecture Notes in Computer Science* **10716** 213-220
[9] Jia Y et al 2015 Caffe: Convolutional architecture for fast feature embedding *Proceedings of the 22nd ACM international conference on Multimedia ACM* 675-678
[10] Wu X, He R and Sun Z 2015 A Lightened CNN for Deep Face Representation *Preprint arXiv*: *1511.02683*
[11] Kaufman L and Rousseeuw P J 2009 *Finding groups in data: an introduction to cluster analysis* (John Wiley & Sons)
[12] Savchenko A V 2017 Clustering and maximum likelihood search for efficient statistical classification with medium-sized database *Optimization Letters* **11(2)** 329-341
[13] Savchenko A V 2018 Trigonometric series in orthogonal expansions for density estimates of deep image features *Computer Optics* **42(1)** 149-158 DOI: 10.18287/2412-6179-2018-42-1-149-158
[14] Savchenko A V 2017 Maximum-likelihood dissimilarities in image recognition with deep neural networks *Computer Optics* **41(3)** 422-430 DOI: 10.18287/2412-6179-2017-41-3-422-430
[15] Nikitin M Yu, Konushin V S and Konushin A S 2017 Neural network model for video-based face recognition with frames quality assessment *Computer Optics* **41(5)** 732-742 DOI: 10.18287/2412-6179-2017-41-5-732-742
[16] Goodfellow I, Bengio Y and Courville A I 2016 *Deep learning* (MIT press)
[17] Yang J 2017 Neural aggregation network for video face recognition *arXiv Preprint*
[18] Savchenko A V and Belova N S 2015 Statistical testing of segment homogeneity in classification of piecewise- regular objects *International Journal of Applied Mathematics and Computer Science* **25(4)** 915-925
[19] *OpenCV* (Access mode: http://opencv.org/)
[20] *TensorFlow API* (Access mode: http://github.com/tensorflow/models/tree/master/research/object_detection/)
[21] *Wider Face: A face detection benchmark* (Access mode: http://mmlab.ie.cuhk.edu.hk/projects/WIDERFace/)
[22] Parkhi O M, Vedaldi A and Zisserman A 2015 Deep face recognition *Proceedings of the British Machine Vision* 6-17
[23] Cao Q, Shen L, Xie W, Parkhi O M and Zisserman A 2017 VGGFace2: A dataset for recognising faces across pose and age *Preprint arXiv: 1710.08092*
[24] Nemirovskiy V B, Stoyanov A K and Goremykina D S 2016 Face recognition based on the proximity measure clustering *Computer Optics* **40(5)** 740-745 DOI: 10.18287/2412-6179-2016-40-5-740-745
[25] Nemirovskiy V B and Stoyanov A K 2017 Clustering face images *Computer Optics* **41(1)** 59-66 DOI: 10.18287/2412-6179-2017-41-1-59-66
[26] Wolf L, Hassner T and Maoz I 2011 Face recognition in unconstrained videos with matched background similarity *IEEE International Conference on Computer Vision and Pattern Recognition* 529-534
[27] Pelilo M and Pavan M 2007 Dominant sets and pairwise clustering *IEEE Transactions on Pattern Analysis and Machine intelligence* **29(1)** 167-172
[28] *Dominant Set Library* (Access mode: https://github.com/xwasco/DominantSetLibrary)

[29] Zhu C, Wen F and Sun J 2011 A rank-order distance based clustering algorithm for face tagging *IEEE International Conference on Computer Vision and Pattern Recognition* 481-488

Acknowledgements
The work was conducted at Laboratory of Algorithms and Technologies for Network Analysis, National Research University Higher School of Economics and supported by Russian Federation President grant MD-306.2017.9.

Technology for automatically characterizing the primary colors of the image and counting the number of objects

M I Khotilin[1], R A Paringer[1], I A Rytsarev[1] and N S Kravtsova[1]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086

Abstract. This work is devoted to the analysis of images and finding objects located on them, with the subsequent calculation of their number. The method of preprocessing based on, so-called, allocation of priority colors of the image is considered. The paper reflects the results of a study of the developed algorithm for object selection and clustering using priority image color analysis.

1. Introduction

Images have always occupied an important place in people's lives: whether they are rock carvings in primitive society, paintings of great artists and especially photographs With the advent of social networks, photographs began to move from the personal category to the public domain and became not only a new way of self-presentation, but also gave rise to other ways of thinking and seeing [1]. People literally translate their "I": they express their opinion, download the so-called "selfies", arrange discussions, share impressions and emotions in comments to photos or by color processing of photos with the help of various editors or filters.

The color characteristic is one of the most significant characteristics, which describes an image. With the help of it you can extract a variety of information that can be useful in many areas of human activity, such as psychology and sociology. In 2015, the Laboratory for the Study of Big Data, under the leadership of Lev Manovich, collected and analyzed 2.3 million photographs of major cities such as New York, Tokyo, Bangkok, San Francisco and eleven other global cities from the social network Instagram. With the help of the laboratory's analysis, we were able to see the "biological rhythms" by which the metropolis lives, and also how the life of a resident of one city differs from that of another, how time frames differ, etc. [2].

Also, the extracted data can be useful for example in medicine, in the selection and recognition of objects and in many other areas.

Recognition and selection of objects on images are one of the most significant tasks of computer vision and image processing. It finds its solution, for example, in finding various kinds of tumors in medicine, tracking objects in security systems, restoring images, and many others. To select objects of the image, various methods are used, both classical and modern, which use neural networks and large data processing tools, or BigData.

In this paper, we describe in detail the method of preprocessing an image by highlighting priority colors with their subsequent clustering. This preprocessing allows you to simplify the process of selecting image objects.

2. Highlighting the priority colors of the image

What is an image? Each image can be viewed as a scene, which is a set of objects of different colors and sizes. Whatever we explore: a frame from a computer's webcam, or a high-resolution photo taken with a digital camera - any digital images are a collection of dots, which called “pixels”. The method of selecting priority colors is based on working with pixels.

The essence of this method is as follows: since any digital image is a set of pixels characterized by their color brightness values, the entire image is treated as a three-layer table, each layer of which is a table with brightnesses of red, green and blue. Next, there is a color distribution analysis of each layer, which results in a corrected table (matrix) of the color. The next stage of the resulting new three-layer matrix is the resulting image. At this stage, the preprocessing ends.

Further, after preprocessing, the process of selecting image objects takes place. It consists in the passage of the entire image by a window of a fixed size, which makes it possible to determine the closedness of the color boundary. By determining the closedness of the color boundary, it is possible to determine various image objects.

At the moment, the development process is underway, with the purpose of further refinement of the method of selecting objects for working with large amounts of data.

We demonstrate the operation of the preprocessing method described above with an example. Consider the image shown in Figure 1.

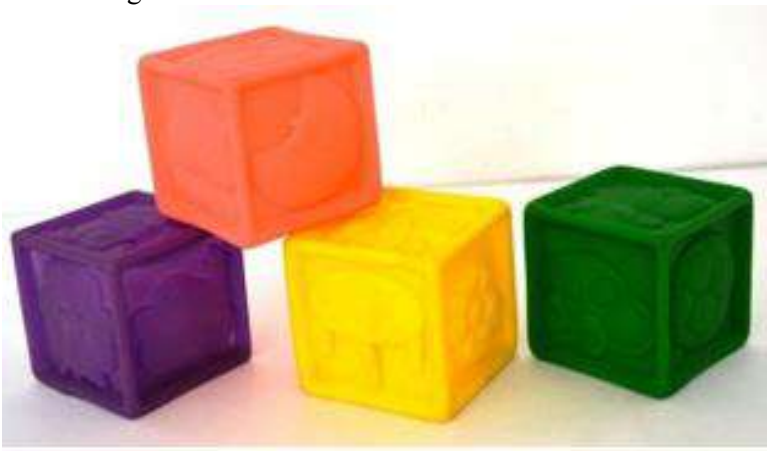


Figure 1. The original image.

In this image, each object has a certain color, and the relief of the object is highlighted by changing the intensity of the color. This image allows us to demonstrate all the steps of the algorithm.

Suppose that each object is uniquely determined by its colors - the main feature by which it can be designated. For the further analysis, a software package MATLAB R2014a was used to solve the problems of technical calculations.

Each pixel of the image in the RGB palette is represented as a blend of three colors: red, green and blue, the first letters of which bear the name of the palette. The distributions of the brightness of pixels on the main colors of the palette are shown in Figure 2.

However, the such distributions are not suitable for implementing the process of determining the priority colors of the image, due to the large number of extremums of these functions.

To solve this problem, it is necessary to bring the distributions to the peak type indicated in Figure 3.

The next step is the application of the method of mean values to color distributions. During the application of this method, the entire color distribution is divided into intervals, and the brightness values are summed. Next, changing the brightness in the original color table of the image, we get the resulting image, shown in Figure 4.

The main colors that occupy the largest area of the object of the resulting image are shown in Table 1.

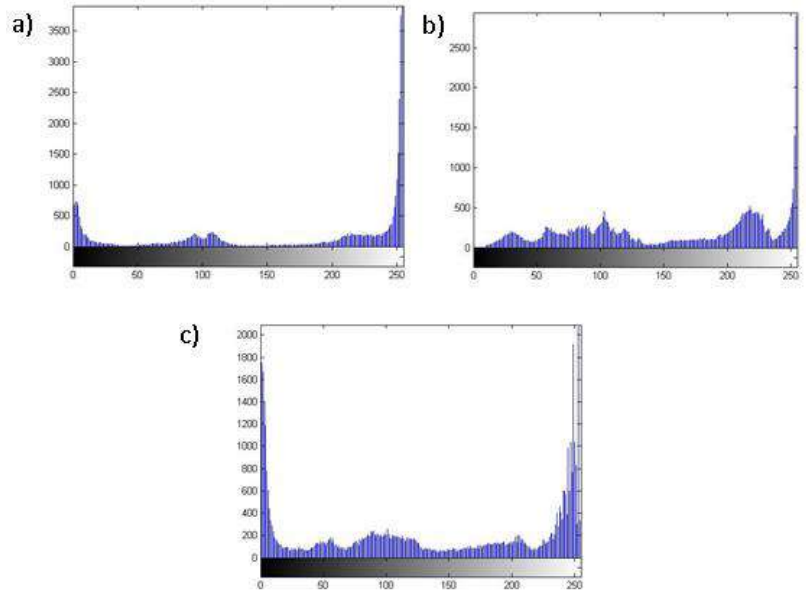


Figure 2. The distribution of the brightness of the pixels of the original image: a) red; b) green; c) blue.

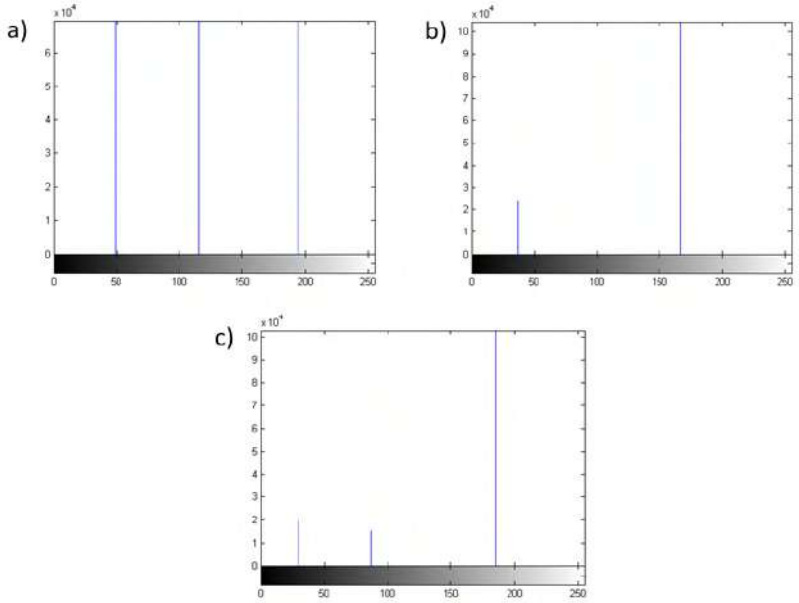


Figure 3. Resulting histograms of transformations.

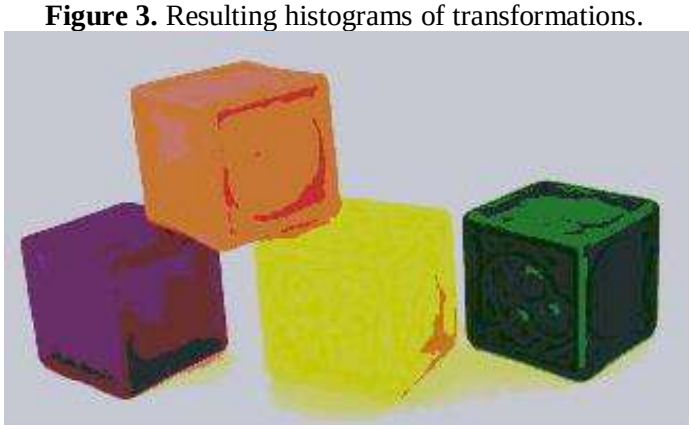


Figure 4. Resulting image.

Table 1. The main colors of the image.

	R: 200 G: 117 B:51
	R: 104 G:46 B:104
	R:199 G:202 B:35
	R:44 G:112 B:53
	R:198 G:200 B:212

After completing this set of operations, it was decided to use a more saturated image with a large number of constituent colors. The new original image is shown in Figure 5.

Figure 5. New original image.

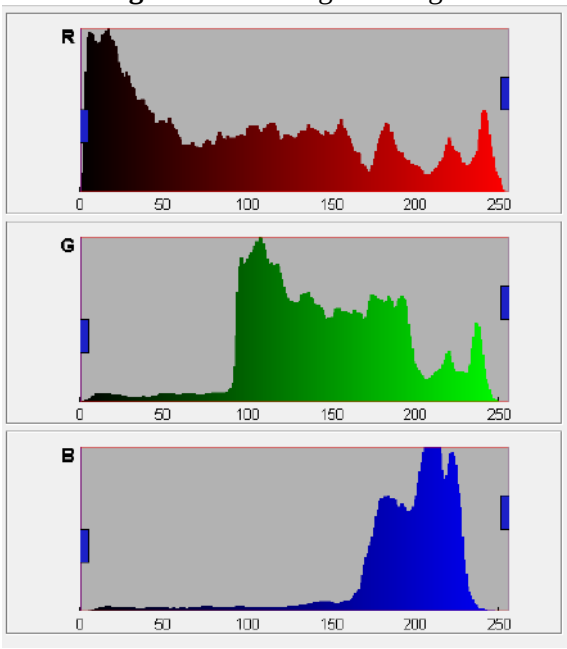


Figure 6. Color distribution of the original image.

The brightness distribution of the original image is shown in Figure 6. The result of passing the image through all stages of the algorithm is the resulting image, shown in Figure 7. The main colors that occupy the largest area of the object of the resulting image are shown in Table 2.

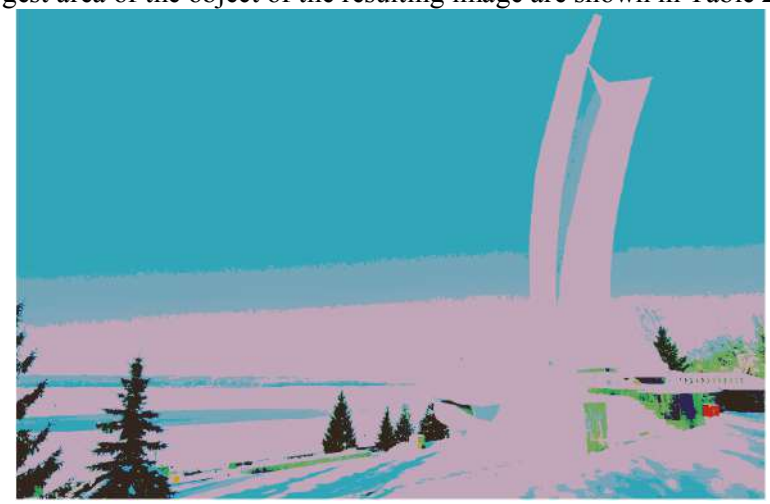


Figure 7. New resulting image.

Table 2. The main colors of the image.

	R: 49 G: 166 B:186
	R: 115 G:166 B:221
	R:194 G:166 B:217
	R:50 G:37 B:29
	R:112 G:168 B:159

3. Conclusion

Recognizing and highlighting key priority colors on images is one of the most significant tasks of computer vision and image processing. It finds its solution in various spheres of human activity. Often, existing methods of image processing and selection of objects on them, work well with relatively small amounts of raw data. Processing of large arrays of source images takes considerable time, which is absolutely unacceptable in a number of problems.

Currently, the authors of this article are working on the adaptation of the above methods to work with large amounts of data or BigData, and sre developing a software tool that allows appropriate processing of image arrays and recognition of objects on them. Also, the authors are completing the algorithm for recognizing image objects for using it in further work with the above technologies.

4. References

[1] Manovich L Z 2015 *Instagram is a window in the mind and imagination of a person* (MediaProfi)
[2] Khaykin S 2008 *Neural networks: full course* (Williams Publishing House) p 1131
[3] Sergeev V V 1998 Application of the methodology of pattern recognition in digital image processing tasks *Autometry* **2** 63-76

[4] Verhagen K, Doyne R and Groon F 1985 *Pattern Recognition: Status and Prospects* (Moscow: R&S) p 104
[5] *The use of neural networks in image recognition* (Access mode: https://geektimes.ru/post/74326) (15.10.2017)
[6] Samal D I 2002 Algorithms for human identification from a photo portrait based on geometric transformations *Thesis for the degree of Candidate of Science. Institute of Technical Sciences. cybern. NAS of Belarus* (Minsk) p 170
[7] *Graphic filters based on the torsion matrix* (Access mode: https://habrahabr.ru/post/43895) (01.11.2017)
[8] Kravtsova N S, Paringer R A and Kupriyanov A V 2017 Parallel implementation of the informative areas generation method in the spatial spectrum domain *Computer Optics* **41(4)** 585-587 DOI: 10.18287/2412-6179-2017-41-4-585-587
[9] Khotilin M and Blagov A 2016 Visualisation and cluster analysis of social networks *CEUR Workshop Proceedings* **1638** 843-850

Acknowledgment
This work was partially supported by the Ministry of education and science of the Russian Federation in the framework of the implementation of the Program of increasing the competitiveness of Samara University among the world's leading scientific and educational centers for 2013-2020 years; by the Russian Foundation for Basic Research grants (# 15-29-03823, # 16-41-630761, # 17-01-00972, # 18-37-00418), in the framework of the state task #0026-2018-0102 "Optoinformation technologies for obtaining and processing hyperspectral data".

Using spatial-temporal maps for visualization of the karst development dynamics

R V Sharapov[1]

[1]Vladimir State University, Orlovskaya street 23, Murom, Russia, 602264

Abstract. There is a problem of visualization dynamics of karst processes during territory monitoring. For this can be used geographic information system. All information on existing forms of karst processes can connected to the map based on the coordinates. But there is a need to include into GIS the information about the dynamics of the karst processes. There is a big problem to solve this task in traditional GIS. Development of karst processes has features that affect their visualization in GIS: the shape and the state of karst objects may change; existing karst forms can be replaced with new objects; the frequency of state changes and appearance of new objects is small; cartographic basis (map) may change as a result of the impact of natural phenomena and human activities. To solve the problem can be used spatial-temporal GIS. In paper the basic features of the karst processes development that affect their storage and display on maps are given. The questions of organization of karst processes dynamics visualization on the basis of spatial-temporal maps are presented. Practical issues of linking the objects in map, display of them geometry, determine temporal boundaries of objects existence are describes.

1. Introduction

In 2009, there were plans to build a Nizhny Novgorod nuclear power plant near the village of Monakovo. But there is one serious problem: Monakovo lies in the zone of active karst-suffusion processes.Karst processes have a significant impact on the safety of civil and industrial facilities.Actual tasks are objective assessment of karst hazard of a particular territory and reducing to a minimum the probability of catastrophes on newly constructed buildings caused by the development of karst processes.

For a long time, the staff of Murum institute of Vladimir State University are monitoring the development of karst processes in the area of the village Monakovo and the karst lake Svyato [1].In this case there was a problem of the linking karst forms data to cartographic basis and its visualization. With the advent of geographic information systems (GIS) there was new direction - the computer cartography. It has been actively used in different spheres of activity. Unlike paper maps, the use of computer technology allows users to select the desired information and produce its visualization in a user friendly manner.This may change the level of detail - enables or disables display of various layers of maps (topography, water bodies, roads, buildings, etc.), carried out generalization when zooming. In addition, GIS allow interactive display of information with support for interacting with objects on the map (for example, information about the selected object). All this inspired us to implement linking karst processes observing system to GIS.

At the initial stage of the project there was almost no difficulties - all information on existing forms of karst processes was linked to the area map based on the coordinates.But later it became necessary to

make information on the dynamics of development of karst processes. For example, in early 2014 formed a new sinkhole in village Chud that has absorbed already existing sinkhole (Fig. 1). There is a problem: what to do with information on the sinkhole - to link it to new sinkhole, remove or remain unchanged? A similar problem arose after filling with water several karst forms formation and in their place a small pond.

a)  b) 

c) 

Figure 1. Sinkhole in village Chud: a – The place of sinkhole in 2013, b – The place of sinkhole in 2014, c – State of sinkhole in the summer of 2014.

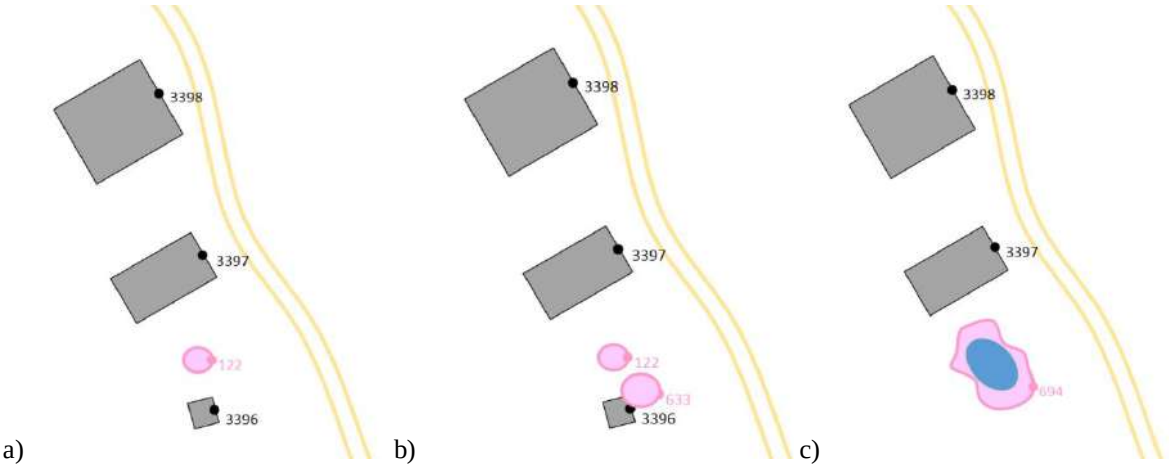
Of course, to display the current status sufficient to store information about currently available forms of karst. In this case, the mapping of the previous states is not required. On the other hand, to obtain a full picture of the development of karst processes necessary to have data about their dynamics. This gives rise to new task: to visualize not only the current state of karst processes, but also the previous state (getting maps for a given timestamp, for example, one year ago, 10 years ago, etc.). In other words, there is a need to compile retrospective (time) maps.

Purpose of work - consider the practical aspects of data linking of karst processes observing system to the geographic information system for the visualization of their development.

2. Main features of karst processes observations

In the observation system for karst processes for each of karst forms there are one or more the descriptions (observations). Each observation can be given as a full or partial (only the important characteristics or changes) description. For a number of observations can be traced to the development of karst processes and the changes (evolution) karst forms.

Let’s consider the example of the evolution of karst forms (see. fig. 2). At the initial time there is a sinkhole №122 (Fig. 2a), which is located between a residential house and a bath. During the development of karst processes formation a new karst form №633 (sinkhole), affecting existing object №3396 (bath), and causing its partial destruction (Fig. 2b). During the further lowering of the surface object №122 and №633 combined into a single complex objects №694, which is partially filled with water (Fig. 2c). As can be seen, in the latter case there was a change of the cartographic basis - disappeared (was disassembled) object number 3396 (Bath).



a) b) c)

Figure 2. An example of the evolution of karst forms on the map: a – One karst form, b – The appearance of the second karst forms with the partial destruction of the object, c – Union of karst forms in a partially filled with water.

Let's consider the features of karst processes, affecting their visualization in GIS:

- form and the state of observation objects (karst forms) can be changed;
- the position of karst forms in space remains constant (except for changing their geometry);
- existing karst forms can be replaced with new objects (such as merging several sinkholes into a single object);
- the frequency of changes in the state and the emergence of new objects is small;
- cartographic base may vary as a result of the impact of natural phenomena and human activities.

3. Approaches to building spatial-temporal maps

There are two basic approaches to create a spatial-temporal maps [2, 3, 4]:

- Creating a set of spatial-temporal maps showing the status of research in specific periods of time (e.g., once a month, once a year, etc.)
- Automatically generating maps based on observation data (including animation dynamic changes).

The most famous example of the implementation of maps of the first kind are the historical maps (satelliteimagery) in Google Earth (fig. 3). Creation a set of spatial-temporal maps is quite laborious. It requires the involvement of specialists to create the next map. Due to this, the process of creating maps is moderated and controlled to eliminate human error. Moreover, in the preparation of each new map it is possible to use the updated cartographic base (map of area).

The second variant does not require specialists to generate a maps - it happens automatically. As a result - errors may occur. However, the variant of automatic generation of maps is more attractive.

There are quite a number of works devoted to the creation and use of spatial-temporal maps. Such maps are used to monitor traffic, road load, display the current status and weather forecasts, the spread of pollution, etc. There are solutions for displaying dynamics of various processes, including real-time. For this purpose special Spatial-Temporal GIS and Temporal GIS.

In [2] discussed questions retrospective (temporal) presentation information in a GIS. A model of time in cartography and variants for presentation of spatial-temporal data using models of space-time cube, sequential images, the base of states with changes and the space-time combination.

In [5] given the basic model for the implementation of temporal GIS: models based on the position of objects, on the objects or their features, on the time, event model, process-oriented models, causal models.

Each object can be described as a triplet (o_i, s_i, t_i), where o_i – object features, s_i– the spatial position of the object, t_i– timestamp when o_i exists in s_i (see. fig. 4) [5].

Models based on the object's position, considering moving along the axis of "space". In this case, the model is a set of parallel planes "Feature-Time" $\{(o, s_j, t)\}$, where $1 \leq j \leq n$ [2].

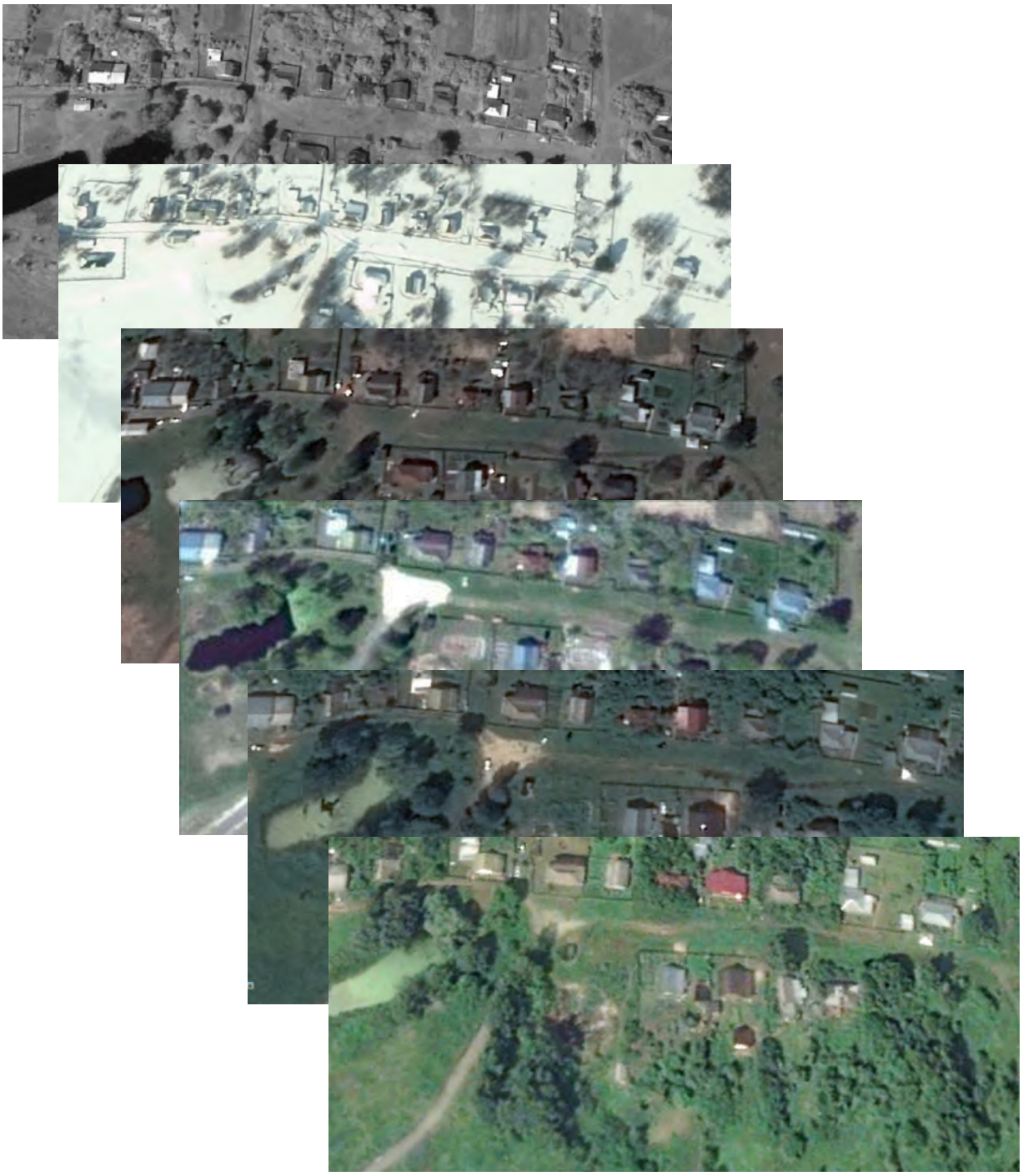


Figure 3. Set of satellite images of territory taken at different times.

Models based on the objects or features considering moving along the axis of "Feature", and is a set of parallel planes, "Space-Time" $\{(o_i, s, t)\}$, where $1 \leq i \leq m$ [6, 7, 8, 9, 10].

The models based on time, make temporary snapshots of reality (of planes "Space-Feature") $\{(o, s, t_k)\}$, where $1 \leq k \leq l$ [5, 11].

In the event model, state transition represents an event. Events can be represented as lines connecting the two states in the data space [12, 13, 14]. In the a process-oriented models are considered occurring processes, such as changing a single object, the functional relationship between objects or modify the spatial structure associated with some aspects [15].

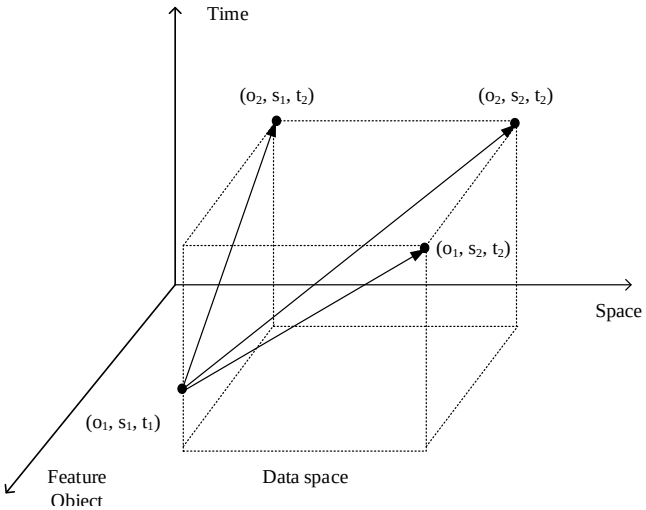


Figure 4. 3D data space.

Questions of modelling and management of spatial and temporal data considered in papers [16, 17]. Thus, there are a number of solutions in the field of presentation space-time (retrospective) information in a GIS. The use of a decision is largely due to the set in each case the task.

4. Presentation of data for visualization

4.1. Data model.
In accordance with the task, it is needed only display the time information, i.e. creation of retrospective maps. For this reason, the basic information for visualization to be objects (karst) and their status (geometry, water availability, etc.).

4.2. Spatial reference object.
As a basis for spatial reference karst forms the basis used to map GPS coordinates of objects. Because objects do not move in space, their coordinates on the axis "space" remain constant.

4.3. Object geometry.
Manifestations of karst processes (sinkhole) are closed, often circular, shape. For this reason, the map can be presented either in the form of an ellipse or a closed polyline. In the first case it is sufficient to store the coordinates of the ellipse center and the radius the major and minor semi-axes. Additionally, it may indicate the angle of rotation of the axes of the ellipse. In the second case, a sequence of pairs of coordinates storage defining the contours of the object. In this last pair of coordinates must be equal to the first.

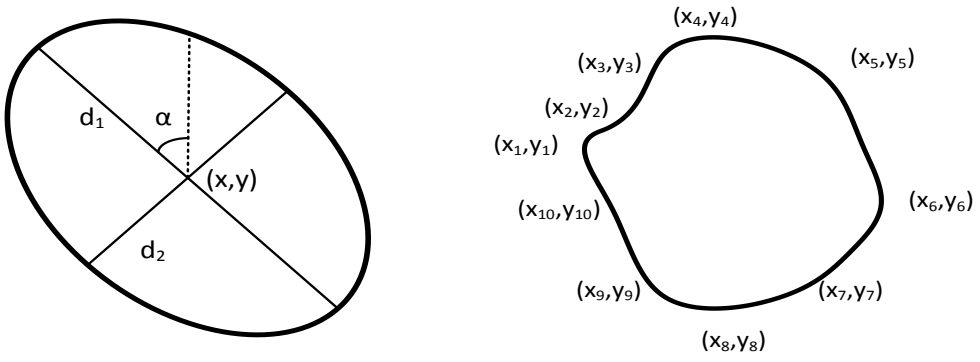


Figure 5. Object geometry.

4.4. Evaluation time presence of the object.
Since karst forms can occur at different times, there is the problem of determining time slots, where a form will be displayed on the map (the axis of "Time"). To determine the initial interval may use information about the time of occurrence the karst forms or data first describe it. The time of occurrence can not always be determined: sinkholes are formed slowly, and the old sinkholes can not be properly described. In this case, the first time is used to describe an object.

The upper boundary of the time display object is set only if the status of an object indicates that it has ceased to exist (moved to another facility, or has been removed). In the this case, the date of the last of the description (in which the status was changed and the object) and read by the upper boundary of the time display.

4.5. Feature of the object
To determine the features of karst forms (axis “Feature”) used observations. Specific features of the objects selected from a set of data represented in each observation. Karst object can change its state. For example, sinkholes can be filled with water, turn into swamp or dry.

Since observations may be incomplete and often describe only certain changes to detect feature changes necessary to analyze not only current, but previous observations of objects. Status change has s_i if $s_i \neq s_j$, where $s_j \neq 0$ (a set value) and $i > j$. Since not all states can be described in the observations, adopted a set of default conditions (such as lack of water in karst, the lack of vegetation, etc.).

4.6. Object status.
The development of karst-suffusion processes, objects can be combined with each other. With active business karst may disappear (e.g., falling sinkholes at the earthworks). For this reason, it is necessary the introduction of the status of the object. The default status for any object – “object exists”. When combined with other objects are labelled as “object merged”, the transformation of the object (for example, the conversion of the sinkhole to water bodies), set the status “object is transformed”. With the disappearance of the object status is set to “object does not exist”.

4.7. Temporal Database
It is necessary to develop a database that allows to effectively manage all necessary information for storage of spatially temporary data. A lot of works are devoted to the creation of such databases [18-25]. For this, both specialized structures and relational databases can be used [26-27].

Implementing data storage with relational databases seems to us more promising and convenient. Currently, there is a large number of both commercial (Microsoft SQL Server, Oracle Database, Informix) and freely distributed DBMS (MySQL, SQLite, FireBird). The solutions implemented in them allow users to work with large amounts of data, perform fast search and processing of information, perform transactions and control access to data.

In the general case, each object in the database must have the attributes of the observation time. This corresponds to the axis of "Time". Such attributes must have the objects themselves, their properties and their geometry. In most DBMSs, these attributes can be a pair (FromDate, ToDate) of data type values.

The FromDate value indicates the date from which the property becomes relevant. It contains the date of observation. The property ToDate contains the date when the next observations were received and the property ceased to be relevant (replaced by a new value). This property for the currently relevant data contains the NULL value. The value of ToDate takes on the value of FromDate of new observation, when new observations appear. It is necessary to correctly update ToDate values to exclude anomalies in the database. For this purpose triggers can be usedin DBMS.

5. Practical implementation

On the basis of presented approaches has been implemented visualization module of karst processes observations in a GIS. Open Source Geographic Information System QGIS was used for practical implementation. By default, the visualization of the current status of the map with karst, confirmed in the most recent observations.The user can select the desired time stamp and the module calculates the

existence of objects. All objects, while the presence of (being) which overlaps with the selected time stamp, is automatically drawn in separate dynamic layers (karst holes, sinkholes, water bodies, etc.).
The results of the practical implementation of the visualization system on the example of studies of particular area in Nizhny Novgorod region [28].

During the implementation of the project, we faced several serious problems.Firstly, this is a great deal of time consuming filling the database. This is due to the need to make not only observations and geometry of objects, but also accounting for the transformation and disappearance of objects. And these operations are most often performed in manual mode.

Secondly, the question arose: timestamps should be set for each graphic primitive and object property or for the whole data packet of one observation. The first option is very laborious, complicates the database and increases the redundancy. The second option complicates the search for changes, because as a result of new observations, not all data change.In the end we had to choose the first option.

Thirdly, we are faced with incomplete data. This is due to the periodicity of observations, not complete coverage of the territory with each new observation, fixing not all characteristics of the observed objects.Because of this we had to look through the entire list of subsequent observations for many objects, in order to avoid erroneous disappearances and subsequent occurrences of objects on the map.

Fourthly, in the study area, earthwork were carried out in connection with the preparation of the site for construction Nizhny Novgorod nuclear power plant and road construction.Because of this, the relief was changed, part of the karst forms was covered with soil.Due to incomplete data last observation was not clear - whether some karst forms were destroyed (transformed) or their study simply was not carried out. To eliminate this ambiguity, we were forced to conduct an additional expedition for field research.Nevertheless, the issue of eliminating these ambiguities in the future remains open.When conducting new observations in the future, it is necessary either to collect a full range of data for all objects, or the status of unexplored objects will be inaccurate. The second option is not a very good solution, given the importance and danger of the planned NPP construction.

In this way, implementation of a system for the Nizhny Novgorod region has allowed us to see areas where karst processes are actively developing at present time. In addition, we were able to see the areas in which there are many karst forms, but they formed a long time ago and is currently active karst processes is not observed.

6. Conclusions

The proposed solutions connect the data observation system for karst processes to geographic information system and solve the problem of visualizing their development. Due to the fact that the display of karst on the map is carried out on a given timestamp automatically according to the observations, it is possible to observe the dynamics of development of karst processes in user-defined time frame.

Obtained with the help of GIS visualization of karst processes in time allows anew look at the problem of studies of karst territories. With this solution, the user can see the area as it was in the past. Without the use of a spatial-temporal GIS to user available maps showing only the current status of the territory.Observing the dynamics of karst process development is an important step for assessing the karst hazard of the territory.

It should be noted that the construction of spatial-temporal maps is a much more labor-intensive process than conventional maps. For this reason it makes sense to create such maps for areas with a dynamic terrain change and territories with important objects.

7. References

[1] Kuzichkin O R and Sharapov R V 2013 Monitoring of karst-suffusion formation in area of nuclear power plant *Proc. 7th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems* **2** 810-813

[2] Stuart N 1992 Time in Geographic Information-Systems *International Journal of Geographical Information Systems* **6** 547

[3] Al-Taha K and Barrera R 1990 Temporal data and GIS: An overview *Proc.GIS/LIS* (Bethesda, MD: American Congress on Surveying and Mapping) 244-254
[4] Egenhofer M and Golledge R 1998 *Spatial and Temporal Reasoning in Geographic Information Systems* (Oxford University Press: New York, NY) p 276
[5] El-Geresy B A, Abdelmoty A I and Jones C B 2002 Spatiotemporal geographic information systems: a causal perspective *Advances in Databases and Information Systems* 191-203
[6] Becker L, Voigtmann A and Hinrichs K H 1996 Temporal support for geo-data in object-oriented databases *Proc. 7th International Conference on Database and Expert Systems Applications DEXA* 79-93
[7] Voigtmann A 1997 *An Object-Oriented Database Kernel for Spatio-Temporal Geo-Applications* (Münster: WestfälischeWilhelms-Universität) p 260
[8] Ramachandran S, McLeod F and Dowers S 1996 Modeling temporal changes in a GIS using an object-oriented approach *Proc. 7th International Symposium on Spatial Data Handling* **2** 518-537
[9] Frank A 1998 Different types of "times" in GIS S*patial and Temporal Reasoning in Geographic Information Systems* 40-62
[10] Segev A and Shoshani A 1993 A temporal data model based on time sequences *Temporal Databases: Theory, Design, and Implementation* 248-270
[11] Hornsby K and Egenhofer M 1997 Qualitative Representation of Change *Spatial Information Theory -A Theoretical Basis for GIS, International Conference* 15-33
[12] Peuquet D J and Duan N 1995 An Event-Based Spatiotemporal Data Model (ESTDM) for temporal Analysis of Geographical Data *International Journal of Geographic Information Systems* **9** 7-24
[13] Claramunt C and Theriault M 1995 Managing time in GIS: An event-oriented approach *Recent Advances on Temporal Databases* 23-42
[14] Dorling D and Openshaw S 1992 Using computer animation to visualise space–time patterns *Environment and Planning B: Planning and Design* **19** 639-650
[15] Cheng T and Molenaar M 1998 A process-oriented spatio-temporal data model to support physical environment modelling *Proc. 8th International Symposium on Spatial Data Handling* 418-429
[16] Wang X Y, Zhou X F and Lu S L 2000 Spatiotemporal data modelling and management: a survey *Proc. 36th International Conference on Technology of Object-Oriented Languages and Systems TOOLS-Asia* 202-211
[17] Worboys M F 1992 A model for spatio-temporal information *Proc.Fifth International Symposium on Spatial Data Handling* **1** 602-611
[18] Freksa C 1992 Temporal reasoning based on semi-intervals *Artificial Intelligence* **54** 199-227
[19] Montanari A and Pernici B 1993 Temporal reasoning *Temporal Databases: Theory, Design, and Implementation* (The Benjamin/Cummings Publishing Company, Inc.) 534-562
[20] Tansel A U, Clifford J, Gadia S, Jajodia S, Segev A and Snodgrass R 1993 *Temporal Databases: Theory, Design, and Implementation* (The Benjamin /Cummings Publishing Company, Inc.) 633
[21] Wuu G and Dayal U 1992 A uniform model for temporal object-oriented databases *Proc. of the International Conference on Data Engineering* 584-593
[22] Snodgrass R 1992 Temporal databases *Theories and methods of spatio-temporal reasoning in geographic space* 22-64
[23] Edsall R, Kraak M-J, MacEachren A, Peuquet D 1997 Assessing the effectiveness of temporal legends in environmental visualization *Proc. GIS/LIS* (Cincinnati, USA) 677-685
[24] Denisova A Y and Sergeev V V 2015 Impulse response identification for remote sensing images using gis data *Computer Optics* **39(4)** 557-563 DOI: 10.18287/0134-2452-2015-39-4-557-563
[25] Belim S V and Larionov S B 2016 An algorithm of image segmentation based on community detection in graphs *Computer Optics* **40(6)** 904-910 DOI: 10.18287/2412-6179-2016-40-6-904-910

[26] Gadia S K and Yeung C S 1988 A generalized model for a relational temporal database *Proc. of ACM SIGMOD International Conference on Management of Data* 251-259
[27] Vorobiova N S, Sergeyev V V and Chernov A V 2016 Information technology of early crop identification by using satellite images *Computer Optics* **40(6)** 929-938 DOI: 10.18287/2412-6179-2016-40-6-929-938
[28] Sharapov R 2015 Visualization of the karst development dynamics by spatial-temporal maps *Proc. 15th International Multidisciplinary Scientific GeoConference SGEM* **2(1)** 913-920

Spatio-Temporal Slices for Frame Cut Detection in Video

N A Sorokina[1] and V A Fedoseev[1,2]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086
[2]Image Processing Systems Institute - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001

Abstract. The paper proposes an approach for unauthorized inter-frame video change detection using spatio-temporal slices. This approach can significantly reduce the amount of data processed and replace video processing with image processing that could be performed much faster. To test the efficiency of this approach, we consider a simple algorithm in analyzing adjacent rows of a slice and then classifying the rows based on its result. Experimental studies have revealed that this algorithm shows moderate results in terms of quality, but it has a great potential for improvement, which confirms the prospects of spatio-temporal slices for the given problem.

1. Introduction

1.1. Problem statement

Today digital video plays an increasingly important role in the society. In 2015, according to the Sandvine report [1], the percentage of video and audio in North American traffic exceeded 70%. In addition, according to the Ericsson report [2], by 2019 the percentage of video in mobile traffic should exceed 50% (and already now it is above 40%). The reasons for that is not only entertainment industry development, but also the growing market for video surveillance systems (up to 20% per annum, according to the Markets and Markets analysts' report [3]), and their widespread introduction to both large business structures and small companies. As a consequence, the data received by surveillance systems are increasingly used in investigative activities or as evidence in legal proceedings. For this reason, such data must be reliably protected from unauthorized alteration by intruders.

One of the most common methods of unauthorized video alteration is inter-frame modification, which includes the removal of video fragments or their replacement with copies of other ones. Such changes can remove crime evidence or data about movements of persons or vehicles that are important in a particular context. In video signals obtained from a stationary camera, such changes can be practically invisible. In the case of a moving camera, such changes can also be hard to detect when an intruder cuts or replaces short-term fragments.

In this paper, we consider the problem of detecting inter-frame artificial changes in video signals taken from a moving or stationary camera. The detection method should work well with video signals stored in various formats, and also combine high detection accuracy with high speed.

1.2. Review of related studies

In practice, the detection of artificial changes in video can be carried out using digital forensics methods developed since the second half of the 2000s [4-6]. The main achievements in this direction are associated with H. Farid, A.C. Popescu, S. Prasad, J. Fridrich, A. Piva, M. Barni. So,

the latest two in the review paper [7] classified methods of digital video forensics onto the following groups:

1) camera-based methods that analyze various video artifacts to determine the optical system of the camera;

2) coding-based methods that identify artifacts resulting from encoding video using certain codecs;

3) geometry, or physics-based methods that detect violations in the physical or geometric parameters of the objects observed;

4) pixel-based methods based on detecting changes at the pixel level of the video.

Examples of algorithms aimed to detect inter-frame modifications can be found, in particular, in [8-14]. Most of them are coding-based methods [8-11] and geometry / physics-based [12-14].

The coding-based methods [8-11] are based on the properties of certain video formats (usually different versions of MPEG) and assume the separation of frames into different types (P-frames, I-frames, etc.). Therefore, these methods do not satisfy the claimed universality requirement regarding the data format. In addition, a significant number of such methods (in particular, [8-10]) may be used to determine the fact of video change, but do not allow to determine the exact location of the changes (in the time domain).

As for the geometry / physics-based methods, many of them (in particular, [12-13]) are based on the use of an optical flow to track changes frame by frame. This technology works well for the case of a stationary camera, but for a moving camera, it needs to take into account camera movements which can be unknown. Moreover, the methods based on optical flow do not provide a high computational performance. Another method from this group [14] utilizes another approach. As a vivid example, this paper considers a simple scene with moving balls. To detect the removal of a video fragment, the method first tracks ball positions, and then detects physically unjustified deviations of the found trajectories, which is possible evidence of unauthorized changes. The main drawback of this method is the need to solve a complex problem of tracking several moving objects. However, the basic idea of this method is rather attractive: to detect unauthorized alterations, we can try to analyze video data over a long time interval.

In this paper, we test a method based on the same idea of detecting deviations in time. However, as the analyzed data, we propose to use the so-called spatio-temporal slices of video images that are slices of the video data cube along the time axis and one of the spatial axes (for example, the horizontal one). If you build several such horizontal slices of video at a certain vertical interval, the resulting set of images will give quite enough information about object movements, although the data will have much less volume comparing with the original video. Moreover, for the processing of these data, we can use computationally effective image processing methods, in particular, parallel-recursive FIR filters [15, 16].

In papers [17, 18] a similar approach is used for road object detection in the problem of autonomous navigation., The algorithms [17, 18] also process not all the pixels of each frame, but only horizontal lines spaced from each other at equal intervals. This allows the authors to ensure the solution of the problem with satisfactory accuracy in real time.

The paper is organized as follows. Section 2 illustrates the traces of natural events in spatio-temporal slices and outlines the principles for detecting inter-frame video changes with their help. Section 3 describes a simple cut detection method based on these principles. Finally, Section 4 describes the experimental studies.

2. Object movements at spatio-temporal slices

Figure 1 shows an example of a spatio-temporal slice of a video obtained by a stationary camera. In this figure, the following events occurred are marked with numbers: (1) – movements by the hands of a standing person, (2) – movements of a person who emerged from one door and entered another, (3) – appearance and stopping of a car, (4) - the appearance of a person from the left border of the frame and their movement. As the figure shows, these events lead to the appearance of smooth curves on the cuts characterizing object movements. You can also note the local background shifts due to camera deviations or fluctuations of observable objects (trees, advertising signs, etc.). Figure 2 illustrates four typical types of object movements observed in spatio-temporal slices.

In case of unauthorized video alteration in the time domain, video slices can clearly visualize them by horizontal offsets, as shown in Figure 3. These offsets can appear over the entire frame width. Furthermore, if we consider a set of slices made at different vertical positions, such offsets will be observed in the same lines of the slices corresponding to the same time shift. Thus, we assume that the considered video slices contain enough information to detect inter-frame changes.

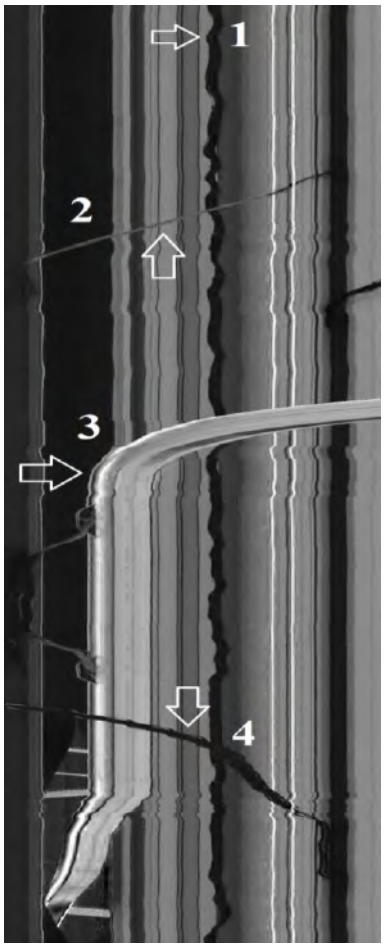


Figure 1. A spatio-temporal slice of a video containing several events.

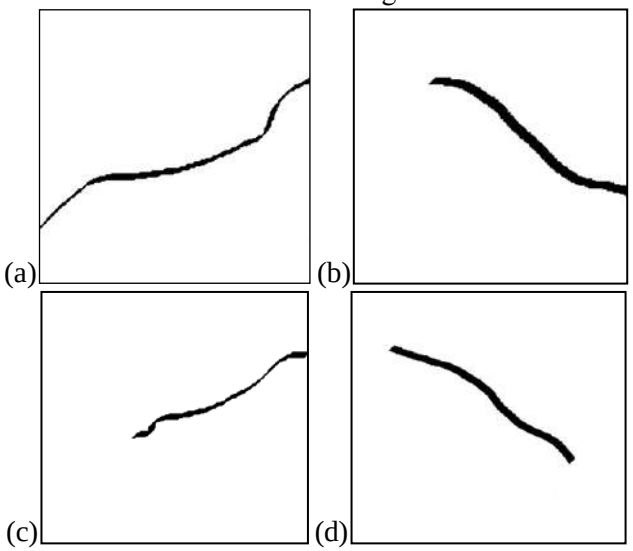


Figure 2. Four types of object movement in spatio-temporal slices: (a) an object enters the camera view and goes beyond it, (b) an object appears in the camera view and goes beyond it, (c) an object enters the camera view and disappears, (d) an object appears and disappears within the camera view.

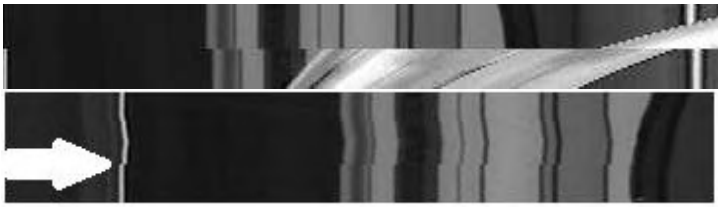


Figure 3. Examples of video fragment deletion shown in spatio-temporal slices.

3. Cut detection method

As noted in Section 2, the evidence for unnatural inter-frame changes is sharp horizontal shifts in the slice image (see Figure 3). To detect them, we can estimate the displacement of each next row relative to the previous one. Formally speaking, we are aimed to find the shift Δ_k of the (k+1)-th row relative to the k-th row, which would provide the smallest error between neighboring samples:

$$\Delta_k = \arg\min_{\delta} \varepsilon_k(\delta), \tag{1}$$

where δ is an integer argument characterizing the line shift, and $\varepsilon_k(\delta)$ is the error function estimated from the equation:

$$\varepsilon_k(\delta) = \sum_{j=\max(0,-\delta)}^{\min(W-1-\delta,W-1)} \left(I(k,j) - I(k+1,j+\delta)\right)^2. \tag{2}$$

In (2), $I(k,j)$ is the grayscale value of the slice image at k-th row and j-th column, W is the frame width. We can numerically solve problem (1)-(2) using the correlation approach.

Next, we can use the obtained Δ_k and $\varepsilon_k(\Delta_k)$ (for simplicity, we will sign the latest one as ε_k, i.e. without the argument) to detect artificial changes based on the following assumptions. Low values of Δ_k and ε_k indicate unaltered videos, while sharp leaps in either Δ_k or ε_k may give evidence of artificial changes (see Fig. 4-5).

To detect artificial changes using Δ_k and ε_k, we used the algorithm based on supervised learning. It classified slice image rows into two classes: “Cut” and “Non-cut”. To train the algorithm, we used the following features of the rows:

$$p_1 = |\Delta_k|, \tag{3}$$

$$p_2 = \varepsilon_k, \tag{4}$$

$$p_3 = \min\left(|\Delta_k - \Delta_{k-1}|, |\Delta_{k+1} - \Delta_k|\right), \tag{5}$$

$$p_4 = \min\left(|\varepsilon_k - \varepsilon_{k-1}|, |\varepsilon_{k+1} - \varepsilon_k|\right), \tag{6}$$

$$p_5 = |\Delta_k - med(\Delta_k)|, \tag{7}$$

$$p_6 = |\varepsilon_k - med(\varepsilon_k)|. \tag{8}$$

We should note especially that we calculated the features (3)-(8) in the local maxima only. Function $med(x)$ in equations (7)-(8) means the median of the 4-point neighborhood of pixel x excluding x itself.

To speed up the algorithm, we classified the only rows corresponding to the local maxima of p_3 and p_4.

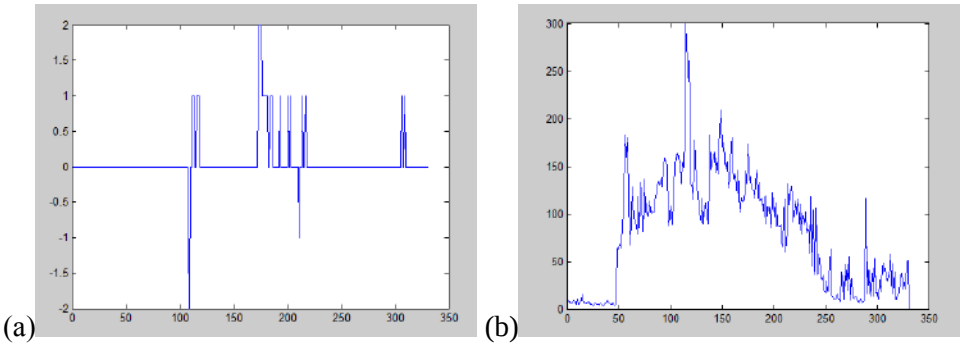


Figure 4. Dependence of Δ_k (a) and ε_k (b) on k for an unaltered video.

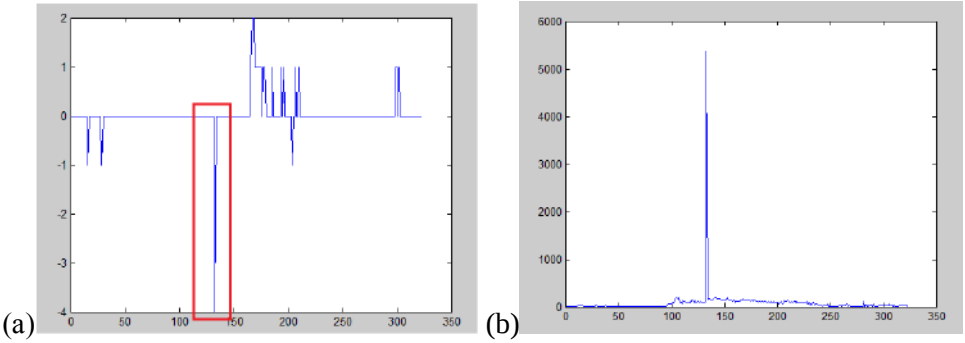


Figure 5. Dependence of Δ_k (a) and ε_k (b) on k for a video with a cut at frame 130.

4. Experimental investigations

To test the proposed method, we used two types of video: DVR recordings (made from a moving car), and model records (made mainly by a stationary camera and containing typical pedestrian movements).

To conduct the experiments, we made spatio-temporal slices of the source videos, and then divide them onto 100-line fragments. A half of these fragments were obtained from 100 consecutive frames, while the other half was composite and contained a frame cut in the 50th line. The length of the gap in

the frames was a parameter of the experiment. Then for each image, we calculated feature sets. 70% of the data obtained was used as a training set, whereas the remaining 30% was a test sample. When testing the method, it did not use the information about the line of the gap. The experimental studies were carried out in two stages. The first one was aimed to select the most appropriate feature set and classifier model. We considered two models: linear SVM and non-linear SVM with radial basis function. At the second stage, we investigated the algorithm performance for various gap lengths and for different types of video. In addition, at the second stage, we analyzed the efficiency of combining data from different lines of several slices, which was made by summarizing lines.

The first stage of the experiments was carried out on DVR recordings with the gap length of 30 and without slice summation. The results of this stage in terms of classification accuracy (equal to the fraction of the correct classifications) are given in Figure 6. The diagram in Figure 6 shows that the best accuracy values are resulted from the use of the (p_3, p_4) feature set and linear SVM. Therefore, these options were further used in the second stage (see results in Table 1). The obtained results show that slice summation improves the classification quality sufficiently. We may also notice that the algorithm works better on DVR videos which contain a rapidly changing background. In general, the final results allow us to conclude that the proposed method is able to solve the considered problem, even in the simplified version of the algorithm described in Section 2.

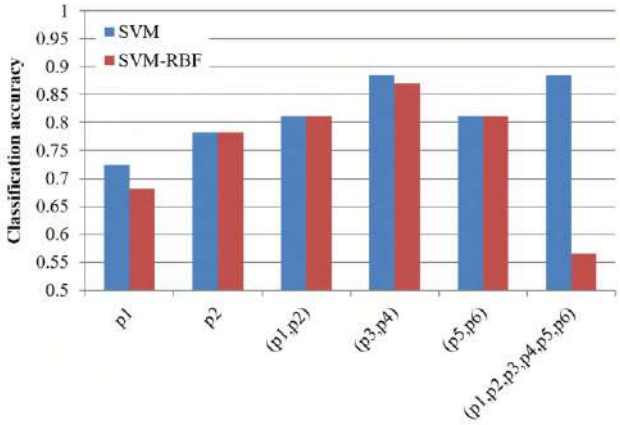


Figure 6. Results of the second stage of the experiment: classification model and feature set selection.

Table 1. Results of the second stage of the experiment: accuracy estimation for different video types.

Video Type	Cut Length	Slice Summation	Cut Detection Accuracy
DVR	10	+	0.8406
		–	0.8400
	30	+	0.9063
		–	0.8696
	60	+	0.9375
		–	0.8732
Model	10	+	0.8732
		–	0.7654
	30	+	0.7971
		–	0.7200
	60	+	0.8485
		–	0.7308

5. Conclusion

In this paper, we have tested the approach based on spatio-temporal video slices to solve the problem of detecting unauthorized inter-frame changes in video. This method is theoretically capable to solve this problem with a high speed due to the processing of only a part of the video signal, as well as using

fast image processing techniques. To test the efficiency of this approach, we proposed a simple algorithm for detecting inter-frame changes and performed some numerical experiments. Our studies showed that the algorithm provides an accuracy of not less than 0.8 and works better for video captured with a moving camera. The results of the studies allow us to conclude that the method of spatio-temporal image slices looks promising, but the algorithm should be significantly improved in terms of increasing the accuracy at short gap lengths and when using video from a stationary camera.

6. References

[1] *Anon CNW Sandvine: Over 70% of North American traffic is now streaming video and audio* (Access mode: http://www.newswire.ca/news-releases/sandvine-over-70-of-north-american-traffic-is-now-streaming-video-and-audio-560769981.html)

[2] *Anon Internet Video Streaming to Dominate Mobile Data Traffic by 2019 - ISPreview UK* (Access mode: http://www.ispreview.co.uk/index.php/2014/06/internet-video-streaming-dominate-mobile-data-traffic-2019.html)

[3] *Anon Video Surveillance Market by Applications & Management Services* 2015 MarketsandMarkets (Access mode: http://www.marketsandmarkets.com/Market-Reports/surveillance-277.html)

[4] Poisel R and Tjoa S 2011 Forensics Investigations of Multimedia Data: A Review of the State-of-the-Art *Sixth International Conference on IT Security Incident Management and IT Forensics* 48-61

[5] Rocha A, Scheirer W, Boult T and Goldenstein S 2011 Vision of the Unseen: Current Trends and Challenges in Digital Image and Video Forensics *ACM Comput. Surv.* **43** 26:1-26:42

[6] Gashnikov M V, Glumov N I, Kuznetsov A V, Mitekin V A, Myasnikov V V and Sergeev V V 2016 Hyperspectral remote sensing data compression and protection *Computer Optics* **40(5)** 689-712 DOI: 10.18287/2412-6179-2016-40-5-689-712

[7] Bestagini P, Fontani K M, Milani S, Barni M, Piva A, Tagliasacchi M and Tubaro K S 2012 An overview on video forensics *Proceedings of the 20th European Signal Processing Conference (EUSIPCO)* 1229-1233

[8] Shanableh T 2013 Detection of frame deletion for digital video forensics *Digital Investigation* **10** 350-360

[9] Stamm M C, Lin W S and Liu K J R 2012 Temporal Forensics and Anti-Forensics for Motion Compensated Video *IEEE Transactions on Information Forensics and Security* **7** 1315-1329

[10] Gironi A, Fontani M, Bianchi T, Piva A and Barni M 2014 A video forensic technique for detecting frame deletion and insertion *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 6226-6230

[11] Wu Y, Jiang X, Sun T and Wang W 2014 Exposing video inter-frame forgery based on velocity field consistency *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 2674-2678

[12] Chao J, Jiang X and Sun T 2013 A Novel Video Inter-frame Forgery Model Detection Scheme Based on Optical Flow Consistency *The International Workshop on Digital Forensics and Watermarking* 267-281

[13] Wang W and Farid H 2006 Exposing Digital Forgeries in Video by Detecting Double MPEG Compression *Proceedings of the 8th Workshop on Multimedia and Security* (New York, NY, USA: ACM) 37-47

[14] Zhang J, Su Y and Zhang M 2009 Exposing Digital Video Forgery by Ghost Shadow Artifact *Proceedings of the First ACM Workshop on Multimedia in Forensics* 49-54

[15] Glumov N I, Myasnikov V V and Sergeyev V V 1996 Parallel-recursive local image processing and polynomial bases *Proceedings of Third International Conference on Electronics, Circuits, and Systems* **2** 696-699

[16] Myasnikov V V 2007 Fast algorithm for recursive computation of the convolution of an image with a two-dimensional inseparable polynomial FIR filter *Pattern Recognit. Image Anal.* **17** 421-427
[17] Kiy K I and Dickmanns E D 2004 A color vision system for real-time analysis of road scenes *IEEE Intelligent Vehicles Symposium* 54-59
[18] Kiy K I 2015 A New Real-Time Method of Contextual Image Description and Its Application in Robot Navigation and Intelligent Control *Computer Vision in Control Systems* **2** 109-133

Acknowledgments
This work was supported by the Russian Foundation for Basic Research (grants 16-29-09494, 16-41-630676), by the Ministry of Education and Science (grant MK-1907.2017.9), and by the Federal Agency for Scientific Organizations (Agreement 007-GZ/43363/26).

Use autoregressions with multiple roots of the characteristic equations to image representation and filtering

N A Andriyanov[1] and K K Vasiliev[1]

[1]Ulyanovsk State Technical University, Severny Venets str. 32, Ulyanovsk, Russia, 432027

Abstract. The article describes probabilistic properties of autoregressive models with multiple roots of characteristic equations as well as the results of such properties investigation. Particular attention is paid to the study of covariance functions. However, we investigate only such models of different orders that provides equal intervals of correlation in row and column. So the multiplicity of such models by row and column is the same. Asymptotic graphs of covariance functions cross-sections are constructed. It is shown that the cross-sections tend to ellipses when model orders are improving. We also describe in detail the problem of filtering images generated by autoregressions with multiple roots. Dependences of variances of filtering errors on the correlation parameter are obtained. The efficiency of filtering at various signal noise ratios and various orders of models is investigated. The effectiveness and expediency of applying autoregressions with multiple roots are shown in comparison with other autoregressions of high orders.

1. Introduction

A lot of real images is characterized by a smooth change in brightness. It means that real images have rather strong correlation links in some neighborhood. The known autoregressive (AR) image models [1-3] either do not adequately describe the nature of such images even at high values of the correlation coefficient in case of first order model.

However, the use of higher order models requires complex operations to calculate the set of correlation parameters. Meanwhile, there are models of AR random field (RF), that generated by the characteristic equations in image processing and representation literature [4-6,12]. Such models are called AR with multiple roots. In this case, the application of algorithms based on such models, when processing real signals and in various applied problems, can contribute to improving the efficiency of solving such problems. In recent years, particular interest is caused by the processing of satellite images [13-15].

A lot of tasks connecting with preliminary processing are known, but the important task is to suppress noise or filter images [7,8]. The advantage of models with multiple roots is the possibility of applying recurrent filtration procedures [9]. The article deals with the problem of image modeling on the basis of such models as well as problems of analysis of probability properties and optimal filtering of random fields generated by such models.

2. Autoregressions with multiple roots of characteristic equations

The widely known class of image models is the AR model of RF [1-3]. One of the main reasons for the spread of AR models is the efficient mathematical apparatus developed for modeling random

sequences. The AR class of RF models can be obtained on the basis of linear stochastic difference equations of the following form:

$$x_{\bar{i}} = \sum_{\bar{j} \in D} \alpha_{\bar{j}} x_{\bar{i}-\bar{j}} + \beta \xi_{\bar{i}} \quad \bar{i} \in \Omega, \tag{1}$$

where $X = \{x_{\bar{i}}, \bar{i} \in \Omega\}$ is modeled RF defined on the N-size grid $\Omega = \{\bar{i} = (i_1, i_2, ..., i_N): \quad \{i_k = \overline{1..M_k}\}, \; k = \overline{1..N}\}$; $\{\alpha_{\bar{j}}, \beta, \bar{j} \in D\}$ are coefficients of the model; $\{\xi_{\bar{i}}, \bar{i} \in \Omega\}$ is a random variable with Gaussian distribution; $D \subset \Omega$ is a causal region of local states.

The choice of a normally distributed RF with independent components is convenient and simple enough to describe some brightness properties of the image by the model. In this case, the RF X will also obey the Gaussian distribution. As an example, let us consider the formation of a two-dimensional RF using of the first order AR model known also as Habibie model:

$$x_{i,j} = \rho_x x_{i-1,j} + \rho_y x_{i,j-1} - \rho_x \rho_y x_{i-1,j-1} + \xi_{i,j}, i = \overline{1..M_1}; j = \overline{1..M_2}, \tag{2}$$

where ρ_x and ρ_y are coefficients of correlation of neighboring elements in row and column, respectively; $\{\xi_{i,j}\}$ is a two-dimensional field of independent Gaussian random variables with zero mean $M\{\xi_{i,j}\} = 0$ and variance $\sigma_\xi^2 = M\{\xi_{i,j}^2\} = (1-\rho_x^2)(1-\rho_y^2)\sigma_x^2$; $\sigma_x^2 = M\{x_{i,j}^2\}$; $M_1 \times M_2$ *is* the size of the image that we imitate.

Due to the specifics of this way to generate the image, the process of estimating the parameters of the AR model can be performed without special difficulties. Nevertheless, such models are not good enough for describing real images, since they provide only small regions of local states. And this is an important factor in modeling multi-zone images. Therefore, if it is necessary to use the AR equation, we must increase the order of the model for the expansion of local states. So this fact entails an inevitable growth of computational costs. Sometimes the image can be so difficult that it is required to perform imitation in a class other than AR. Otherwise, to achieve adequate replacement of the image is unlikely to succeed.

In addition, the weakness of such models is that they are not suitable for an adequate description of isotropic RFs, for example, with a covariation function (CF) of the form $R(\bar{k}) = R(\bar{k} = \sqrt{k_1^2 + k_2^2 + ... + k_N^2})$. To overcome this drawback, one can use the characteristic equations [10]. In work [6] models based on the generalization of one-dimensional ARs to the multidimensional case were proposed. For example, if we take an AR with multiple roots of characteristic equations of the second order

$$x_i = 2\rho x_{i-1} - \rho^2 x_{i-2} + \xi_i, \tag{3}$$

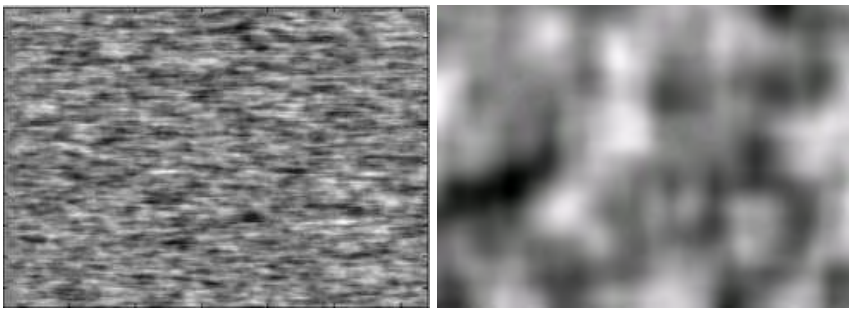
then it is possible to obtain on its basis a model of a two-dimensional second-order RF

$$\begin{aligned} x_{ij} = {} & 2\rho_x x_{i-1,j} + 2\rho_y x_{i,j-1} - 4\rho_x \rho_y x_{i-1,j-1} - \\ & - \rho_x^2 x_{i-2,j} - \rho_y^2 x_{i,j-2} + 2\rho_x^2 \rho_y x_{i-2,j-1} + \\ & + 2\rho_y^2 \rho_x x_{i-1,j-2} - \rho_x^2 \rho_y^2 x_{i-2,j-2} + b\xi_{ij}, \end{aligned} \tag{4}$$

where b is normalizing coefficient, which makes it possible to obtain a stationary RF with a given variance.

It can be noted that the model (4) is an eight-point model, i.e in it to form the next element of the RF $\{x\}$ we use 8 preceding elements from the neighborhood. Similarly, for the multiplicity model (3.3), we can obtain a 15-point model, for the multiplicity model (4.4), we obtain a 24-point model.

Figure 1(a) shows an image formed from the values of RF (4) with model parameters $\rho_x = 0.9$ and $\rho_y = 0.7$. Figure 1(b) shows the implementation of an RF based on AR with multiple roots of characteristic equations having model parameters $\rho_x = 0.95$ and $\rho_y = 0.95$ with multiplicity (3,3). The size of images is *640×480* pixels.



(a) multiplicity (2;2) (b) multiplicity (3;3)

Figure 1. Implementation of models having different multiplicities.

Thus, using a model with multiple roots of characteristic equations when imitating images, it is possible to simply obtain such implementations of RF that will be close to real images. In this case, an important property of the generated RF will be their quasi-isotropy property. A general formula for models of different multiplicities can be written in the form:

$$x_{i,j} = \beta\xi_{i,j} - \sum_{i_1=0}^{N_1}\sum_{j_1=0}^{N_2}\alpha_{i_1,j_1}x_{i-i_1,j-j_1}, \tag{5}$$

where N_1 and N_2 characterize the multiplicity of the model; coefficients $\alpha_{i_1,j_1}(\alpha_{0,0}=0)$ are products of the corresponding coefficients of one-dimensional AR along the axes x and y:

$$\alpha_{i_1,j_1} = \alpha x_{i_1}\alpha y_{j_1}. \tag{6}$$

The coefficients of one-dimensional AP (6) can be obtained from the expression

$$\begin{aligned}\alpha x_{i_1}(\rho_x, N_1) &= (-1)^{i_1+1}C_{N_1}^{i_1}\rho_x^{i_1}\\ \alpha y_{j_1}(\rho_y, N_2) &= (-1)^{j_1+1}C_{N_2}^{j_1}\rho_x^{j_1}\end{aligned} \tag{7}$$

where $C_n^m = \dfrac{n!}{m!(n-m)!}$ is number of combinations of n by m; ρ_x, ρ_y are parameters of the model.

Finally, the two-dimensional model coefficient β is the normalized product of the corresponding coefficients of one-dimensional ARs along the axes x and y:

$$\beta = \frac{\sigma_x}{\sigma_\xi}\beta_x\beta_y. \tag{8}$$

These coefficients can be found from the following formulas

$$\beta_x = \sqrt{\frac{\left(1-\rho_x^2\right)^{2N_1-1}}{\sum_{i=0}^{N_1-1}\left(C_{N_1-1}^i\rho_x^i\right)^2}}, \beta_y = \sqrt{\frac{\left(1-\rho_y^2\right)^{2N_2-1}}{\sum_{i=0}^{N_2-1}\left(C_{N_2-1}^i\rho_y^i\right)^2}}. \tag{9}$$

So we can use expressions (5) - (9) to show, for example, that the AR model of the first order (2) can be represented by a model with multiple roots of multiplicity (1,1), and the eight-point model (4) can be represented by a model with multiple roots with multiplicities (2,2).

Thus, it is possible to construct models of arbitrary multiplicity, obtaining images with slow varying brightness properties. It should be noted that speed of properties change will depend on parameters and orders of models.

3. Covariance functions of autoregressive random fields with multiple roots

In order to construct AR models of images with multiple roots of characteristic equations having given statistical characteristics, we perform an analysis of the data of the characteristics of the considered models. An important task arising in the course of statistical analysis of the mathematical model of an RF is to determine the type of its CF.

An important property of models with multiple roots is the factorizability of CF. For example, the RF generated by the model (1) is anisotropic, and its covariance CF, by virtue of anisotropy, is a

generalization of the CF of a one-dimensional first order AR to a two-dimensional case. It can be shown [4] that it is described by the following expression:

$$B(k_1,k_2)=\sigma_x^2\rho_x^{|k_1|}\rho_y^{|k_2|}, \quad (10)$$

where σ_x^2 is the variance of the RF X; ρ_x and ρ_y are parameters of the model; k_1 and k_2 are the distances between the elements of the RF X along the axes *x* and *y*.

The use of the model (4) provides an increase in the links in the AR model due to the expansion of the range of significant preceding states and the type of CF changes. For model (4), the CF significantly differs from the CF of the first order AR model and it takes the following form [4]:

$$B(k_1,k_2)=\sigma_x^2(1+\frac{1-\rho_x^2}{1+\rho_x^2}|k_1|)(1+\frac{1-\rho_y^2}{1+\rho_y^2}|k_2|)\rho_x^{|k1|}\rho_y^{|k2|}, \quad (11)$$

where σ_x^2 is variance of the RF X; ρ_x, ρ_y are correlation parameters of the model; k_1 and k_2 are the distances between the elements of the RF X along the axes *x* and *y*.

In order to obtain the CF of models with arbitrary orders, one can use expressions for one-dimensional CF of AR with multiple roots of the characteristic equations

$$B_x(k)=\sigma_x^2\sum_{l=0}^{m-1}g(m,l,k)\frac{\rho^{2(m-l-1)}}{(1-\rho^2)^{2k-l-1}}, \quad (12)$$

where $g(m,l,k)=\frac{(m+k-1)!(2m-l-2)!}{l!(m-1)!(m-l-1)!(m+k-l-1)!}$. The variance of RF ξ_i, $i=1,2,...,n$ can be found by using condition $B_x(0)=\sigma_x^2$

$$\sigma_\xi^2=\sigma_x^2(1-\rho^2)^{2m-1}\Big/\sum_{l=0}^{m-1}(C_{m-1}^l\rho^l)^2. \quad (13)$$

Correspondingly, if the AR multiplicity is (m_1,m_2) then the expression for the CF can be written as the product of the RF variance and expressions of the form (12) for one-dimensional CF

$$B_{xy}(k_1,k_2)=\sigma_x^2\sum_{l=0}^{m_1-1}g(m_1,l,k_1)\frac{\rho_x^{2(m_1-l-1)}}{(1-\rho_x^{\,2})^{2k_1-l-1}}\sum_{l=0}^{m_2-1}g(m_2,l,k_2)\frac{\rho_y^{2(m_2-l-1)}}{(1-\rho_y^{\,2})^{2k_2-l-1}}. \quad (14)$$

The obtained relations completely determine the CF and the parameters of the RFs generated by the AR with multiple roots of the characteristic equations of multiplicity (m_1,m_2).

For analysis, an interesting case is when the correlation coefficients of the model provide the same correlation intervals on row and column for models of different orders, i.e.

$$B_{m=1}(\rho_{x1},\rho_{y2},k_0,k_0)=B_{m=2}(\rho_{x2},\rho_{y2},k_0,k_0)=...=B_{m=n}(\rho_{xn},\rho_{yn},k_0,k_0)=...=\frac{\sigma_x^2}{e}. \quad (15)$$

For simplicity, we will assume that the multiplicity of the AR for each of the axes is the same, and the parameter ρ is also the same for both axes. Then we can reduce condition (15) to the following form

$$B_{m=n}(\rho,k_0)=\frac{\sigma_x}{\sqrt{e}}. \quad (16)$$

This simplification makes it possible to obtain a set of parameters that ensure the same correlation intervals, analogous to the one-dimensional case [5]. Table 1 presents the dependences between the correlation parameter ρ and correlation interval k_0.

Analysis of the data presented in Table 1 shows that in order to ensure equal correlation interval, it is necessary to decrease the value of the parameter ρ. Furthermore, it is possible to single out the dependence of expression $\gamma_m=2(1-\rho_m)k_0$. The parameter γ_m tends to 1 if $m=1$. Figure 2 shows dependence of parameter γ_m on correlation interval k_0.

Table 1. Results of numerical calculations of parameters for different correlation intervals.

k_0	1	10	20	50	100	500	1000	10000
$\rho_{m=1}$	0,606	0,9048	0,9512	0,9672	0,9802	0,99004	0,99502	0,9994998
$\rho_{m=2}$	0,338	0,7657	0,8737	0,9137	0,9472	0,9732	0,98653	0,998644
$\rho_{m=3}$	0,2345	0,685	0,8257	0,8797	0,926	0,96225	0,98093	0,998077
$\rho_{m=4}$	0,1795	0,6275	0,7895	0,854	0,9095	0,9536	0,97653	0,99763

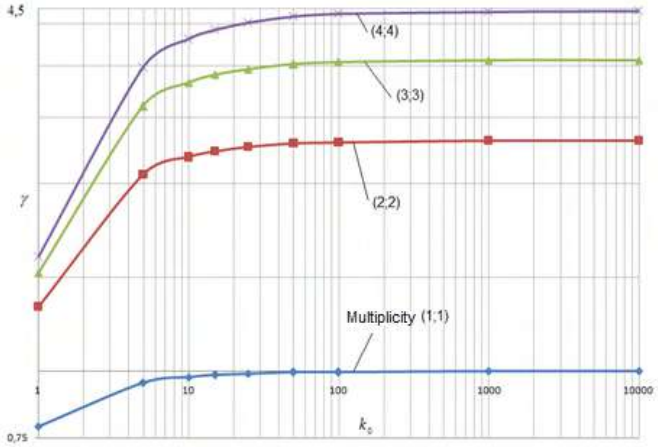
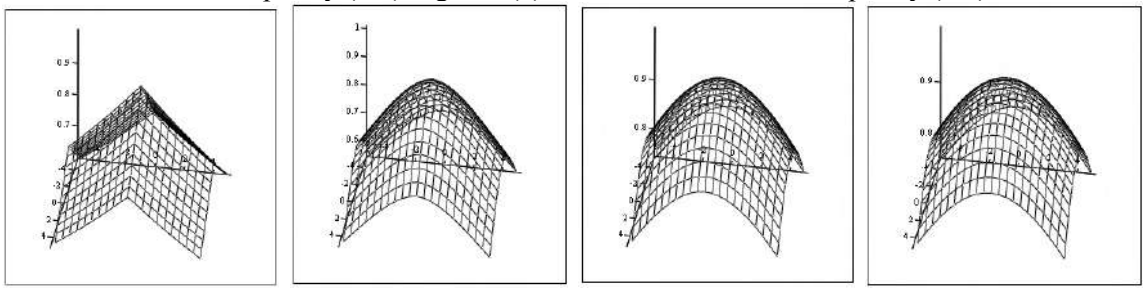


Figure 2. Dependence of the correlation characteristics for models of different orders.

Analysis of the curves in Figure 2 shows that the parameter γ_m tends to a constant value for large k_0. At the same time, the greater the multiplicity of the model, then the greater the parameter γ_m.

Figure 3 shows CF of different order models providing a correlation interval $k_0 = 10$. Figure 3(a) shows the model of multiplicity (1,1), figure 3(b) shows the model of multiplicity (2,2), Figure 3(c) shows the model of multiplicity (3,3), figure 3(d) shows the model of multiplicity (4,4).



(a) (b) (c) (d)

Figure 3. CF of the AR RF with multiple roots of the characteristic equations.

Figure 4 shows the CF cross-sections for the models of the 1st order (solid line) and the 2nd order (dashed line) with the correlation interval $k_0 = 15$. Figure 5 shows CF cross-sections for models of different orders with the same value of the parameter $\rho = 0.8615$.

As can be seen from Figure 4, the second order CF model has a "bell-shaped" vertex, and the correlation links between the RF elements generated by such models are stronger within the correlation interval.

Fig. 5 shows that the CF cross-sections tend to ellipsoids with increasing multiplicities AR. However, with a significant increase in the orders, the CF decreases much more slowly than when using the first and second order ARs.

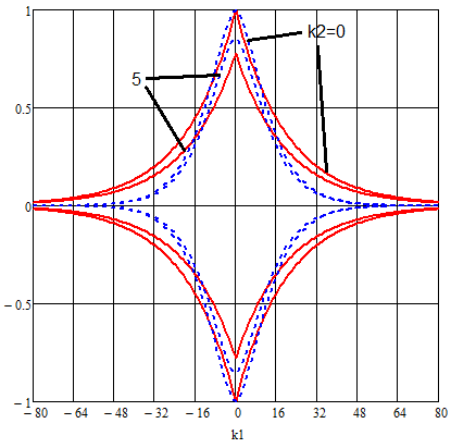


Figure 4. CF cross-sections with the same correlation interval.

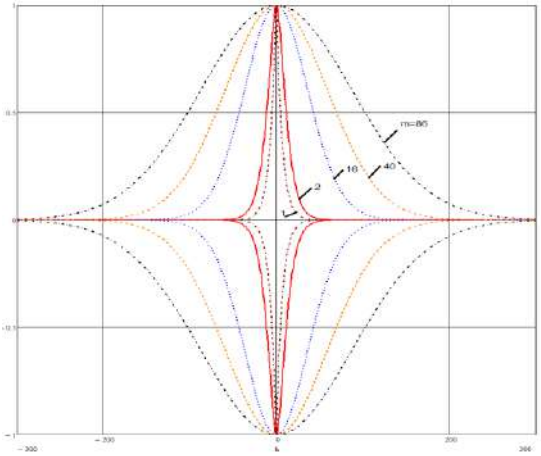


Figure 5. CF cross-sections with the same correlation parameter.

Thus, the correlation properties of AR with multiple roots of the characteristic equations ensuring equal correlation intervals are investigated. Dependences of the correlation parameter of such ARs on the correlation interval are obtained.

4. Covariance functions of autoregressive random fields with multiple roots

Let us consider the case when against a background of white noise, it is necessary to perform a filtration of an RF simulated by the AR with multiple roots of the characteristic equations

$$z_{ij} = x_{ij} + n_{ij}, \tag{17}$$

where white noise has zero mathematical expectation and variance σ_n^2.

We will use the following simple algorithm for image filtering. At the first stage, we perform line by line filtering of all the elements in each row. Then the same procedure applies to each column. Finally, we get the total estimation of each element as the average between the estimates for the row and the column. So, to filter a string, you can use the following algorithm [5]. We introduce the extended state vector:

$$\overline{x}_i = (x_i x_{i-1} ... x_{i-m+1})^T. \tag{18}$$

Then the observation model will be written as:

$$z_i = C\overline{x}_i + n_i,\ i = 1,2,..., \tag{19}$$

where $C = (1\,0...0)$.

The equation of state of a string is also can be written in the vector-matrix form:

$$\overline{x}_i = \wp\overline{x}_{i-1} + \overline{\xi}_i, i = 1,2,..., \tag{20}$$

where $\wp = \begin{pmatrix} \rho_{11} & \rho_{12} & \dots & \rho_{1m} \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 1 \end{pmatrix}$; $\rho_{1j} = (-1)^{j+1} C_m^j \rho^j$; $\overline{\xi}_i = (\xi_i 0...0)^T; V_\xi = M\{\overline{\xi}_i \overline{\xi}_i^T\}$.

After the transformations, we use the standard Kalman linear filter equations to find the optimal estimates of the information RF [11]:

$$\hat{\hat{x}}_i = \hat{\hat{x}}_{эi} + P_i C^T \frac{1}{\sigma_n^2}(z_i - C\hat{\hat{x}}_{эi}), P_i = P_{эi}(E + \frac{1}{\sigma_n^2} C^T C P_{эi})^{-1}, \quad (21)$$

where $P_{эi} = \wp P_{i-1} \wp^T + V_\xi; \hat{\hat{x}}_{эi} = \wp \hat{\hat{x}}_{i-1}$.

At the every i-th estimation step we build the optimal forecast $\hat{x}_{эi} = \sum_{j-1}^{m} \rho_{1j} \hat{x}_{i-j}$ based on previous estimates $\hat{x}_{i-j}$, $j = 1,2,...,m$ and we find the best estimate (in the sense of the minimum variance of the error) $P_{11i} = M\left\{\left(x_i - \hat{x}_i\right)^2\right\}$

$$\hat{x}_i = \hat{x}_{эi} + P_{11i} \frac{1}{\sigma_n^2}(z_i - \hat{x}_{эi}), \quad (21)$$

where $P_{11i} = P_{э11i} \Big/ \left(1 + P_{э11i} \Big/ \sigma_n^2\right)$. The remaining components of the vector $\hat{\hat{x}}_i$ are calculated on the basis of interpolation of previous estimates taking into account the next observation z_i and the mutual covariance of estimation errors.

Figure 6 shows the variances of filtering errors obtained for models of different multiplicity and for different ratios of signal to noise $q = \sigma_x^2 / \sigma_n^2$. It is worth noting that the variance estimate of the filter error was averaged over five processed images of size *150×150*.

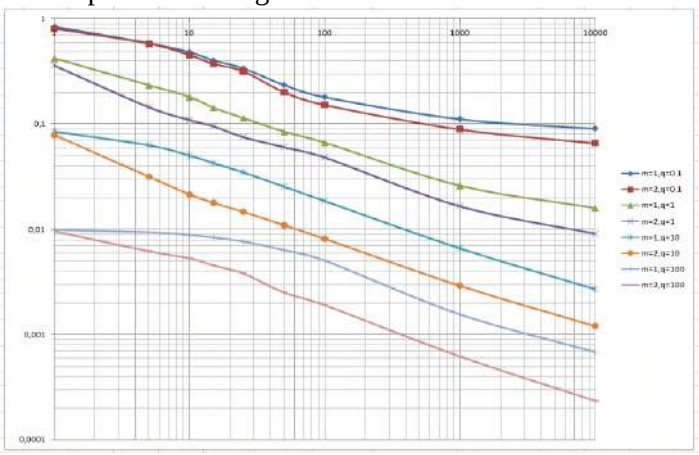


Figure 6. The filtration efficiency of two-dimensional ARs with multiple roots of characteristic equations.

Figure 7 shows the results of the proposed algorithm for the models with multiplicities (1;1) and (2;2). The filtering parameters are following: $q = 1$, $k_0 = 15$, image size is *300×300*. From the top down we show the original images, noisy images, images obtained after filtering.

An analysis of the results shows that the use of AR of higher orders, firstly, makes it possible to obtain smoother fields, and secondly, it provides a greater filtration efficiency (on average 32% for 1-st and 2-nd order models with q=1).

5. Conclusion

Thus, RF models based on AR with multiple roots of characteristic equations are presented. The equality of correlation interval is the interesting property of investigated models. For such models,

dependencies between correlation properties and multiplicities are found. An effective quasi-optimal filtering algorithm for such RFs is proposed. It is shown that the use of high-order models makes it possible to obtain gains in the case of filtering in the case of identical correlation intervals.

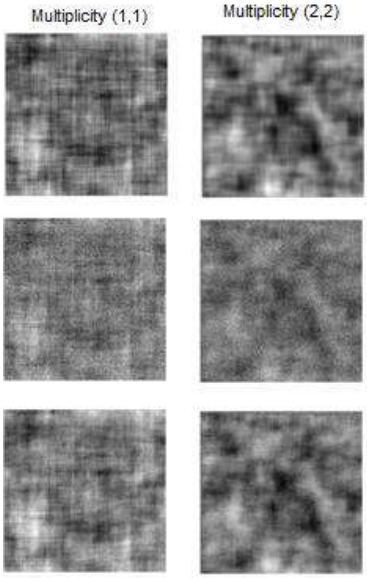


Figure 7. The results of filtering simulated images.

6. References

[1] Sage E and Mels J 1976 *The theory of estimation and its application in communication and management* (Moscow: Communication) p 496
[2] Bulinsky A V and Shashkin A P 2008 Limit theorems for associated random fields and related systems (Moscow: FIZMATLIT) p 480
[3] Glazov G N and Kostevich A G 2002 Modeling discrete Gaussian random fields *Intelligent Systems in Control, Design and Education* **2** 19-27
[4] Vasiliev K K and Krasheninnikov V R 2015 *Statistical analysis of images* (Ulyanovsk: UlSTU) p 214
[5] Vasiliev K K and Andriyanov N A 2017 Analysis of autoregressions with multiple roots of characteristic equations *Radiotekhnika* **6** 13-17
[6] Vasiliev K K, Dragan Ya P and Kazakov V A 1995 *Applied theory of random processes and fields* (Ulyanovsk: UlSTU) p 256
[7] Gonzalez R, Woods R 2005 *Digital image processing* (Moscow: Technosphere) p 1072
[8] Vasiliev K K, Dementiev V E and Andriyanov N A 2017 Filtration and restoration of satellite images using doubly stochastic random fields *Proceedings of the 1st International Workshop on Radio Electronics and Information Technologies* 10-20
[9] Andriyanov N A, Vasiliev K K, Dementiev V E 2017 Anomalies detection on spatially inhomogeneous polyzonal images *CEUR Workshop Proceedings* **1901** 10-15
[10] Vasiliev K K 2014 Autoregressions with multiple roots of characteristic equations *Radiotekhnika* **11** 74-78
[11] Vasiliev K K 2016 *Optimal signal processing in discrete time* (Moscow: Radio Engineering) p 288
[12] Andriyanov N A and Gavrilina Yu N 2018 Image Models and Segmentation Algorithms Based on Discrete Doubly Stochastic Autoregressions with Multiple Roots of Characteristic Equations *3rd International Workshop on Radio Electronics and Information Technologies CEUR Workshop Proceedings* **2076** 19-29
[13] Denisova A Yu and Sergeev V V 2015 Impulse response identification for remote sensing images using GIS data *Computer Optics* **39(4)** 557-563 DOI: 10.18287/0134-2452-2015-39-4-557-563

[14] Fursov V A, Goshin Ye V and Kotkov A P 2016 The hybrid CPU/GPU implementations of the computational procedure for digital terrain models generation from satellite images *Computer Optics* **40(5)** 721-728 DOI: 10.18287/2412-6179-2016-40-5-721-728
[15] Afanasyev A A and Zamyatin A V 2017 Hybrid methods for automatic landscape change detection in noisy data environment *Computer Optics* **41(3)** 431-440 DOI: 10.18287/2412-6179-2017-41-3-431-440

Acknowledgments
The study was supported by RFBR, project № 17-01-00179.

Unsupervised segmentation of ceramic proppant particles in 3D microCT images

E P Serkova[1,2], I V Safonov[1], I V Yakimchuk[1] and V Yu Evstefeeva[1,2]

[1]Schlumberger Moscow Research, Pudovkina 13, Moscow, Russia, 119285
[2]Moscow State University, Leninskie Gory 1, Moscow, Russia, 119991

Abstract. Oil and gas industry uses different types of ceramic proppants in millions of kilograms per year. X-ray microtomography (microCT) imaging can be applied for investigation of quality of the material. For analysis, it is necessary to segment spherical contacting particles of proppant. We apply a marker-controlled watershed for segmentation. The method of markers detection has several parameters which have crucial influence on segmentation outcome. To optimize segmentation quality, we propose unsupervised (non-reference) measure based on a compactness of 3D connected regions, where compactness is calculated via central geometric moments of second order. In addition, we demonstrate advantages of our technique for compactness estimation over the method based on ratio of surface area and volume of a region.

1. Introduction

Oil and gas industry uses various types of proppants (from phrase “propping agent”) in hydraulic fracturing technology. Annually in the world hundreds of thousands of tons of ceramic proppant are produced, the particles of which are granules of spherical shape with a size of about 1 mm. Mechanical strength and conductivity are the most important attributes of proppant pack for optimal fracturing job design. Crush test is one of conventional procedures for measurement of proppant characteristics. Proppant grains crush test should be performed at axial stress up to 100 MPa. Some particles are crushed during the test. The fraction of the crushed particles depends on stress and quality of the proppant. For some proppants the crush-rate is only a few percent. X-ray micro-tomography (microCT) makes capable measurement of morphometric characteristics of each particle as well as their fragments in initial state and after stress.

Figure 1 shows example of real reconstructed 3D microCT image of proppant pack before stress. The grayscale image has size 4000x4000x2000 voxels and 8 bit-depth. The image was obtained with SkyScan 1172 microCT system (Bruker MicroCT, Belgium). One can see proppant particles in the image. Also fragments of crushed particles can be found in images scanned after loading. The usual image of proppant pack contains several hundreds of particles. Intensities of particle regions are much lighter than dark background. However, there is many contacting regions of granules and their fragments. Separation of the contacting regions is a challenging task.

Marker-controlled watershed for distance map is a traditional method for segmentation of touching regions. An algorithm for detection of markers has several parameters, which influence to segmentation quality considerably. Since the fraction of broken particles can be equal to several percent, even single segmentation errors leads to bias of an assessment of quality of the material. Selecting the segmentation parameters manually by the operator is a long and unobvious process.

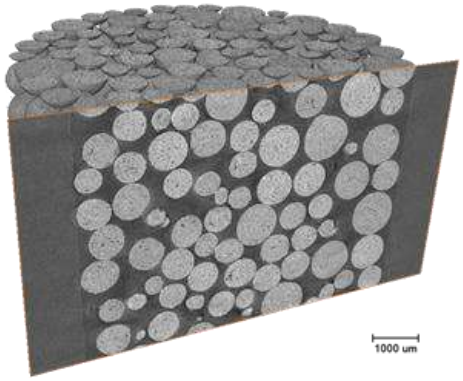


Figure 1. Example of reconstructed 3D microCT image of proppant particles.

In addition, it is quite difficult to detect segmentation errors in a 3D image visually. Our preliminary experiments have shown the optimal segmentation parameters vary from image to image; it is impossible to set parameters once based on a previously processed sample. Therefore, it is important to develop an unsupervised (non-reference) quality metrics that allow selection of the parameters automatically with aim to minimize the number of errors of segmentation of proppant particles.

The main contribution of the paper is proposed unsupervised metrics of segmentation of 3D spherical particles by means of maximization of average compactness of segmented regions, where compactness is calculated via second order central geometrical moments.

2. Segmentation of particles

Figure 2 demonstrates flow-chart of segmentation algorithm. Reconstructed grayscale 3D image is downsampled to image G having size 1000x1000x500 voxels and 8-bits depth. It is necessary to reduce requirements to memory space and decrease processing time. Lighter voxels of ceramic particles differ from dark voxels of background and holder. Histogram of intensities of the image is bi-modal. For images with such histograms thresholding by Otsu algorithm [1] is a good solution to distinguish voxels of particles from background. After thresholding we obtain binary image T, where voxels of solid are designated by 1 and voxels of voids are designated by 0. It is required to split regions formed by touching particles in T image.

Conventional way for separation of overlapping or contacting convex regions without holes is an application of watershed algorithm to the inverted outcome of the geodesic distance transform [2]. The general idea of watershed algorithm is the following: image is considered as a geological relief; a water source is placed in each regional minimum in the relief, to flood the entire relief from sources, and build barriers in the place where different water sources meet; the resulting set of barriers constitutes a watershed by flooding [3]. For volumetric images watershed algorithm operates identically to 2D one.

Before application of distance transform we need to fill holes, which are pores in particles. Several pores are open, they connect with background voids. That is why filling of holes in 3D keeps these pores unchanged. It is required to perform filling of holes for 2D slices. Theoretically, the filling of holes should be done for slices in all three mutually perpendicular directions. Such approach ensures that the open pores, which are penetrate a particle and are parallel to image axes, are filled. However, in practice open pores are tortuous, so, it is enough to fill the holes for 2D slices in only one direction. Binary image T_f is outcome of slice-by-slice filling of holes for T image. It is worth to note, filling of holes in 2D can lead to filling of space between touching particles. Fortunately, it happens seldom and it has no negative impact on the next processing steps.

On the next stage, geodesic distance transform builds the distance map D by calculation for each voxel of T_f image Euclidean distance to the nearest voxel equal to zero. Inverted D plays the role of relief for watershed algorithm. Local minima on (-D) are basins origins. Frequently there are several local minima inside a connected region because even survived particles have non-ideal convex shape,

a lot of fragments of crushed particles are concave. It leads to over-segmentation. To avoid over-segmentation due to the huge number of local minima the marker-controlled watershed is used, where markers play role of water sources [4].

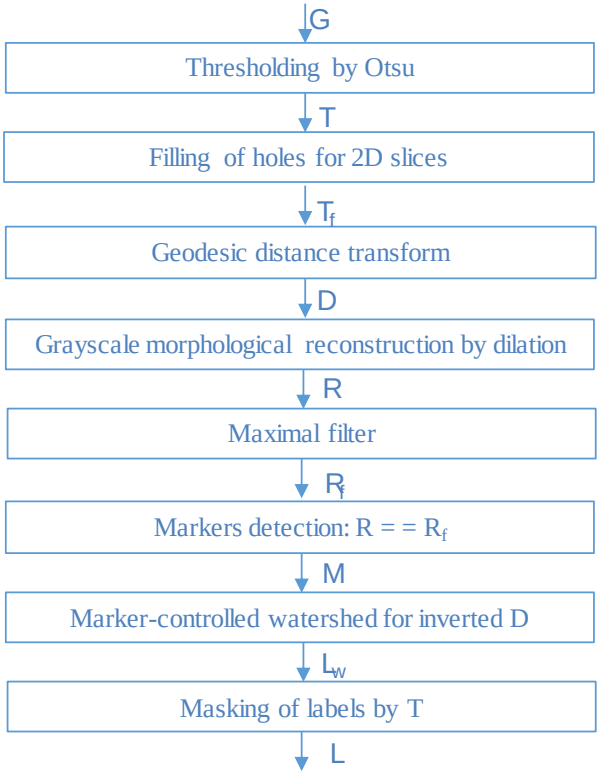


Figure 2. Flow-chart of our segmentation technique.

There is a plenty of approaches for markers generation. We tried several of them and found out the following steps providing the best segmentation outcome. The first step is grayscale morphological reconstruction by dilation [5]. Morphological reconstruction by dilation uses two images: "seed" image (D – *delta_h*), which specifies the values that spread, and "mask" image D, which gives the maximum allowed value at each voxel. The mask image limits the spread of high-intensity values. The resulting reconstructed image R looks exactly like seed image, but with the peaks cut off, where *delta_h* determines height of peaks. Difference (D – R) is referenced as 'h-dome' operator, ((D – R) > 0) is referenced as 'h-maxima' operator.

Figure 3 demonstrates example of markers generation for simple image (see figure 3a) containing two contacted regions. Figure 3b shows distance map. Figure 3c shows plot of profile for the distance map along axis K. One can see four local maxima in the plot, where three local maxima relate to the right region. We need to suppress redundant local maxima. Adjustment of *delta_h* mitigates issue of several local maxima for some particles, for example two local maxima of central peak are combined to single region. However, increasing of *delta_h* cannot solve the issue completely, e.g. the rightmost peak has local maximum that should be ignored. Application of maximal filter to image R and comparison of filtration outcome R_f with R employ to suppress the most of redundant local maxima. Binary image M containing markers is: M = (R == R_f) (see figure 3c). The maximal filter uses cubical structural element with *max_filter_size* size. These three parts of figure 3 illustrate meaning of parameters *delta_h* and *max_filter_size*.

The next stage is marker-controlled watershed for inverted D and markers from M, where each connected region has unique label. Image L_w is outcome of segmentation by watershed. Finally, to obtain labelled particles in T image it is necessary to multiply L_w by T.

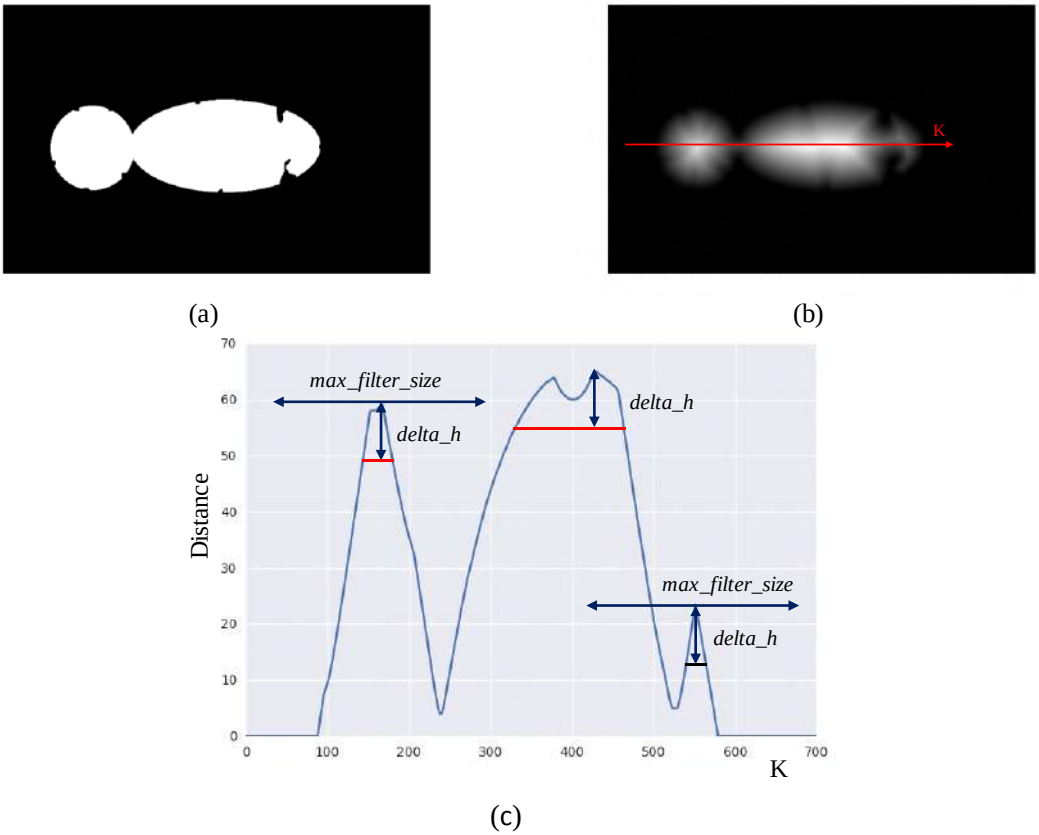


Figure 3. Illustration of parameters for generation of markers.

3. Unsupervised segmentation quality metrics

Segmentation outcome depends on parameters *delta_h* and *max_filter_size*. Decrease of both leads to over-segmentation, because regions of granules split in fragments. Increase of both leads to under-segmentation, because neighbouring particles and fragments merge. How to set the parameters properly? Subjective choice based on visual analysis of segmentation result is a troublesome due to 3D nature of data. It is hard to find optimal parameters visually. Paper [6] analyses supervised quality measures for segmentation such as Global Consistency Error and Rand Index. However, we have no ground truth to be able to apply supervised metrics. How to automatically choose the optimal parameter values, based only on the analysed image? Survey [7] describes a few dozen of unsupervised metrics for estimation of segmentation quality of 2D images. All of them are not universal, but depend on the task being solved. When segmenting regions have more or less the same shape, the shape factor can be used as metrics.

We know a-priori that studied particles have a rounded shape. We use the fact for formulation of criterion of unsupervised segmentation. A sphere is the most compact body in three-dimensional space. In the next section, we describe a *compactness* for characterizing the closeness of a region shape to a sphere. Maximal average compactness of segmented regions corresponds to the best segmentation:

$$Q = \frac{1}{N}\sum_{i}^{N} C_i,$$

where N is the number of segmented regions having volume greater then Vt; C_i is compactness of i-th region. The threshold Vt is introduced to exclude from consideration too small fragments. In addition, it is hard issue: what is compactness for regions consisting of only a few voxels?

Let's consider how Q changes in the case of improper segmentation. Over-segmentation is splitting of spherical regions. It leads to decreasing of average compactness Q. Under-segmentation is merging of spherical regions. As rule, it leads to decreasing of average compactness too. It is worth to

note, proposed segmentation criterion is correct in assumption of existing of at least several tens of survived spherical particles in image after loading. If almost all particles are crushed then the metrics is ineffective.

4. Compactness

Many 3D shape factors are a natural extension of corresponding measures for 2D images. Compactness in 2D is sometimes called the *circularity* or the *roundness*. Shape compactness in 2D is generally understood as the degree to which a given shape differs from a region bounded by a circle. There are plenty publications, which describe compactness as ratio of area of region to squared perimeter. Analogous definition of compactness for 3D regions is ratio of squared volume to area of surface of a region in cubic power [9]:

$$C_a = \frac{36\pi V^2}{S^3},$$

where V is volume of 3D region, S is area of surface.

Paper [8] performs detailed analysis of drawbacks of approaches, which use perimeter and surface area for calculation of compactness for 2D and 3D images correspondingly. First, a perimeter and a surface area are difficult to calculate invariantly to rotation and quite precisely due to the digital nature of images. Second, the surface of the regions is not ideal, there are various types of noises that arise during the registration and segmentation of regions. Third, the evaluation of compactness is seriously affected by holes in regions.

Bribiesca [11] describes compactness for 3D regions that differs from C_a, but it is based on volume and area of surface also. The paper [11] demonstrates Bribiesca's is less sensitive to small distortions of surface in comparison with C_a.

There are several approaches for calculation of surface area. For example, the paper [9] considers marching cubes algorithm [10] for estimation of outer surface area. A surface area can be estimated as the number of voxels of the outer shell, where the shell is the difference in the region and the result of its erosion with a 3x3x3 structural element in the form of a cube (26-connectivity), or a ball (18-connectivity), or a cross (6-connectivity). Surely, all pores should be filled in advance. Which method is preferable?

Zunic et al. [12] proves properties of 3D shapes compactness based on second order central geometric moments and demonstrates its advantages over other algorithms for compactness calculation:

$$C_m = \frac{3^{5/3}}{5(4\pi)^{2/3}} \frac{{\mu_{000}}^{5/3}}{\mu_{200} + \mu_{020} + \mu_{002}},$$

where central geometric moments are:

$$\mu_{pqr} = \sum_x \sum_y \sum_z I(x,y,z)(x - m_{100})^p (y - m_{010})^q (z - m_{001})^r,$$

where geometric moments are:

$$m_{pqr} = \sum_x \sum_y \sum_z I(x,y,z) x^p y^q z^r,$$

where $I(x,y,z)$ is indicator function of particle region, I equals one for voxels belonging the region, and I equals zero otherwise. Zero order geometric moments equal to volume of particle: $m_{000} = \mu_{000} = V$. First order geometric moments are centroids. Compactness lies in the range from 0 to 1. Compactness of ideal sphere equals 1.

Paper [12] is based on outcomes from Mamistvalov's theory. Mamistvalov published in English mathematical theory for recognition of n-dimensional solids via geometric moment invariants in 1998 [13]. In Russian, it was done in 1974. Unfortunately, those papers are not widely known. Ratio $\frac{\mu_{200}+\mu_{020}+\mu_{002}}{{\mu_{000}}^{5/3}}$ is so-called *first 3D moment invariant* based on second order central geometric moments. The invariant is constant or slightly changes for origin translation, uniform scaling and rotation. C_m is the inverted *first 3D moment invariant* multiplied by the constant, so, it is an invariant for translation, rotation and scaling as well. In general, geometric moments have been widely used in

statistics for description of the form of a probability density function and in classic rigid-body mechanics to measure the mass distribution of a body. Usage of moment invariants is a prospective direction in an image processing for creation of shape factors as well as for development of algorithms for shapes analysis and recognition. According to monography [14], last decade an interest to shape analysis and classification via moment invariants grows. Not only geometric invariants are applied, but orthogonal moments such as Gaussian-Hermite, Zernike, Chebyshev, Legendre and Fourier-Mellin.

Let's consider which method of computing the compactness is better suited for our problem. We calculate the compactness coefficients for spheres of different radii with different random noises such as protuberances and cavities on a surface. Figure 4 shows a slice of the sphere with typical noise on the surface. In the next experiment, we determine which method for compactness calculation allows us to distinguish sphere from regions formed by cutting off segments of different sizes from the sphere better.

Figure 4. Slice of sphere with cavities and outliers on a surface.

Plots in figure 5 show various compactnesses depending on radius of sphere: C_m; C_a, where surface area is calculated by marching cubes algorithm; C_a, where surface area is calculated by means of erosion, Bribiesca's compactness; theoretical ideal case, where compactness equals 1. C_a for both considered approaches are very different from 1 and have significant fluctuations. C_m is close to 1 except for spheres having radius less than 10. Bribiesca's compactness is almost 1 for all range of considered radiuses.

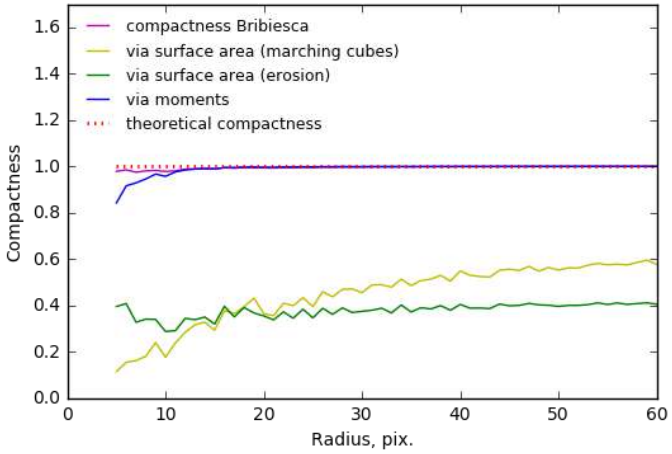


Figure 5. Compactness of spheres of different radiuses with noisy surface.

Plots in figure 6 show compactnesses for bodies from hemisphere (L=60) to sphere (L=0). Figure 7 illustrates meaning of L parameter. C_m and C_a allow to distinguish sphere and sphere with clipped segment. Bribiesca's compactness is too similar even for sphere and hemisphere.

We conclude, application of compactness C_m calculated via second order central geometrical moments has obvious advantages in comparison with well-known approaches based on surface area

and volume of 3D region. In addition, in contrast to C_a, C_m has robustness to presence of the modest number of open and close pores.

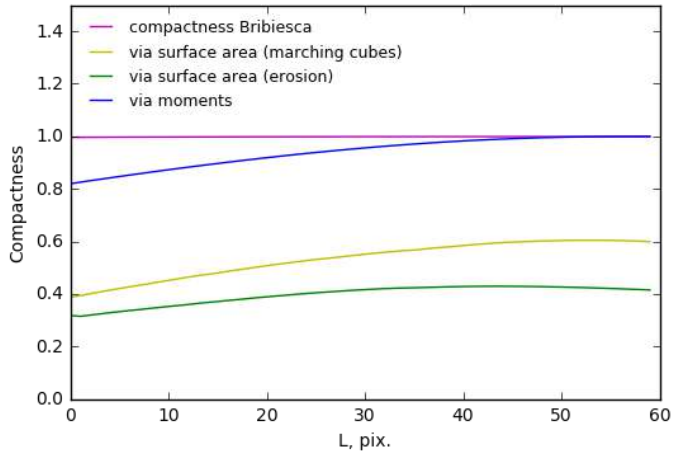


Figure 6. Compactness when the shape of the body changes from the hemisphere to the sphere.

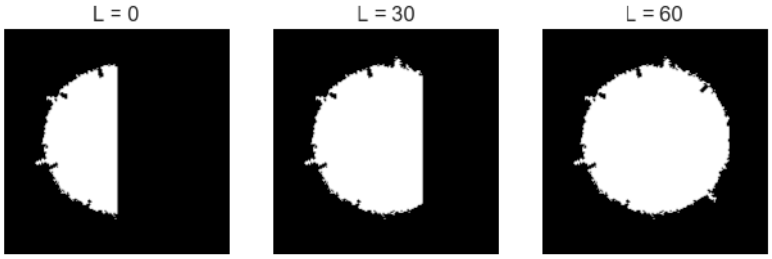


Figure 7. Illustration of meaning of L parameter for figure 6.

5. Results and discussion

For estimation of benefits of proposed unsupervised segmentation technique, we processed five 3D images with fixed parameters *delta_h* and *max_filter_size* as well as with parameters obtained by maximization of Q metrics. It is worth to note, those fixed parameters were optimal according to Q criterion for two images segmented previously. We apply gradient descent algorithm for looking for maximal Q. Figure 8 shows Q metrics depending on parameters *delta_h* and *max_filter_size*.

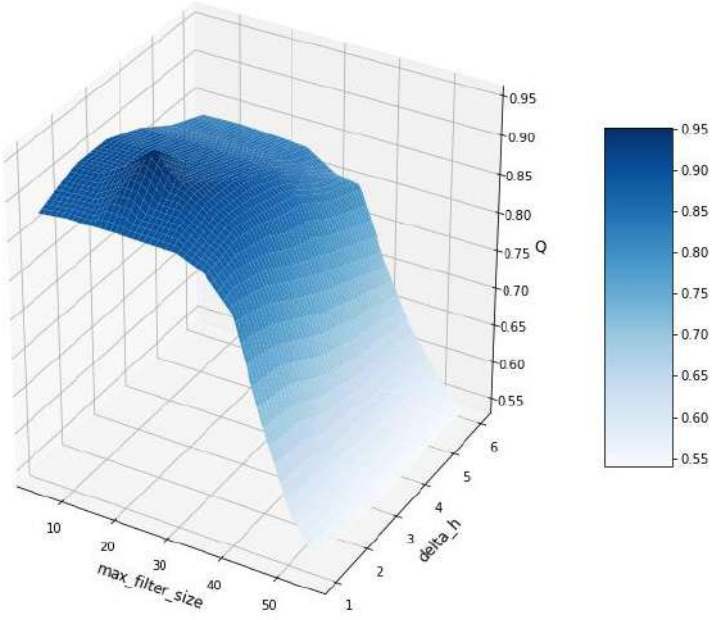


Figure 8. Segmentation quality metrics *Q* depending on parameters *delta_h* and *max_filter_size*.

Table 1. The number of erroneously segmented particles.

Image	Fixed parameters	Parameters according to maximal Q
1	37	0
2	37	0
3	54	0
4	44	2
5	38	1

Table 1 contains for each tested image the number of erroneously segmented particles by segmentation with fixed parameters and by proposed unsupervised segmentation. Total number of particles in each image is 615. Segmentation with fixed parameters leads to 30-50 errors. Figure 9 shows slice of segmented and labeled image by segmentation with fixed parameters. One can see, one ellipsoidal particle was segmented as two regions, ten pairs of neighboring particles were pairwise combined into one region. Unsupervised segmentation had no errors, or just one or two errors in the worst case. Figure 10 demonstrates the slice of 3D image, where unsupervised segmentation was applied. Segmentation outcome is fully correct.

So, proposed unsupervised criterion for the segmentation of 3D images of ceramic proppant allows providing high quality segmentation in automatic mode.

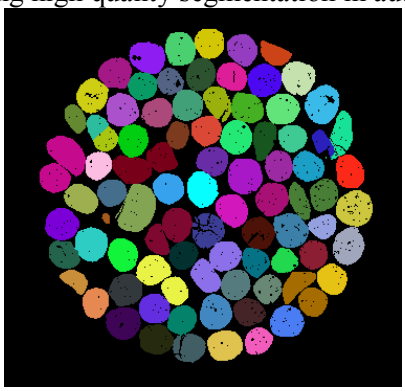


Figure 9. Slice of segmented 3D image with several separated and merged particles; fixed parameters were used for segmentation.

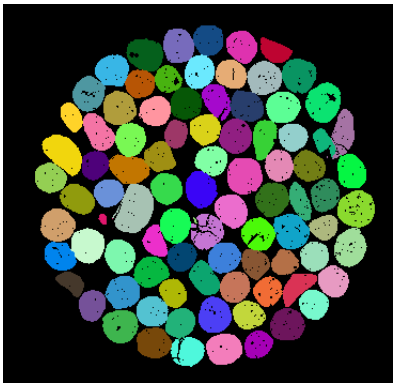


Figure 10. Slice of segmented 3D image with proper segmentation; parameters according to maximal Q were selected.

6. References

[1] Otsu N A 1979 Threshold selection method from gray-level histograms *IEEE Trans. Syst. Man Cybern.* **9(1)** 62-66
[2] Vincent L and Soille P 1991 Watersheds in digital spaces: an efficient algorithm based on immersion simulations *IEEE Trans. Pattern Anal. Mach. Intell.* **6** 583-598
[3] Gonzalez R C and Woods R E 2012 *Digital image processing*
[4] Safonov I V, Mavrin G N and Kryzhanovsky K A 2006 Segmentation of convex cells with partially undefined boundaries *Pattern Recogn. Im. Analysis* **16(1)** 46-49
[5] Vincent L 1993 Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms *IEEE Trans. Image Process* **2(2)** 176-201
[6] Myasnikov E V 2017 Hyperspectral image segmentation using dimensionality reduction and classical segmentation approaches *Computer Optics* **41(4)** 564-72 DOI: 10.18287/2412-6179-2017-41-4-564-572
[7] Zhang H, Fritts J E and Goldman S A 2008 Image segmentation evaluation: A survey of unsupervised methods *Comput. Vis. Image Underst.* **110(2)** 260-280

[8] Montero R S and Bribiesca E 2009 State of the art of compactness and circularity measure *Internat. Math. Forum* **4(27)** 1305-1335
[9] Zhao B and Wang J 2016 3D quantitative shape analysis on form, roundness, and compactness with μCT *Powder Technol.* **291** 262-275
[10] Lorensen W E and Cline H E 1987 Marching cubes: A high resolution 3D surface construction algorithm *ACM Siggraph Comp. Graphics* **21(4)** 163-169
[11] Bribiesca E 2008 An easy measure of compactness for 2D and 3D shapes *Pattern Recognit.* **41(2)** 543-554
[12] Žunić J, Hirota K and Martinez-Ortiz C 2012 Compactness measure for 3d shapes *ICIEV* 1180-1184
[13] Mamistvalov A G 1998 N-dimensional moment invariants and conceptual mathematical theory of recognition n-dimensional solids *IEEE Trans. Pattern. Anal. Mach. Intell*. **20(8)** 819-831
[14] Flusser J, Suk T and Zitová B 2016 2D and 3D *Image Analysis by Moments* (John Wiley & Sons)

An image understanding system based on the geometrized histograms method: finding the sky in road scenes

K I Kiy[1]

[1]Keldysh Institute of Applied Mathematics of RAS, Miusskaya square 4, Moscow, Russia, 145047

Abstract. In this paper, the technique provided by the geometrized histogram method for segmentation and description of color images is developed and improved in order to analyze the adjacency relation of left and right germs of contrast objects (left and right contrast curves) on the *STG*. This adjacency relation involves and generalizes the adjacency relation for regions in classical segmentation methods (the so-called RAG). Using this order relation, the adjacency relation for left and right germs of contrast objects is established. This order relation is also employed for finding relations between left and right germs with prescribed geometric and color-intensity characteristic that are not adjacent and lay apart at a distance. In addition, the concept of contours that are close to vertical on *STG* is introduced. Based on the adjacency relation proposed, a technique for constructing complex contrast objects with a prescribed geometric shape and color-intensity description is proposed. The developed technique is applied to analyzing road scenes in order to find the sky in video sequences. The results of finding the object in video sequences by a program complex, implementing these ideas, are presented and discussed.

1. Introduction

In spite of serious progress in image segmentation and analysis [1-4] and many new ideas arising in machine learning and deep learning of convolutional networks, there are still serious difficulties in implementing global image analysis in real time. These difficulties are mainly connected with many different objects occurred in the scene, occlusion, and difficult and diverse illumination conditions in the real world. This makes it difficult to analyze the joint behavior of several real objects and even to assemble parts of objects separated by occlusion using one or another method of image analysis (classical segmentation, sliding windows with machine learning or deep learning of convolutional networks). Moreover, as the mortal accident with the Tesla pilotless vehicle has shown, it is crucially important not only to classify any frame, but also to have clear understanding of the state of important objects in the image and to produce their conceptual description in order to recognize the case of their complete change due to possible occlusion occurred. It is also necessary to analyze the dynamics of the motion and changes in the shape of objects in video sequences. In this case, it can prevent us from recognizing the body of a blue van as a part of the sky. It is necessary for the image understanding system of a robot to be able to select the sky region, to determine its shape and the location of its boundary. It is also desirable to describe the semantic type of the sky region. The information about the sky region has to contain its color and intensity characteristics and a semantic interpretation of the regions over which it lies. The complete change in this information within a small number of frames

can inform the system about the dangerous occlusion occurred. This information has to be used in order to prevent a possible accident.

In this paper, we propose an image understanding system that can solve such problems in real time using only standard computational facilities. The approach to designing image understanding systems of such a type is based on the geometrized histograms method proposed by the author [5-7]. This method not only segments color images in real time, but also makes it possible to construct adjacency relations between detected objects and to introduce higher-order adjacency relations for objects that are rather distant in the image. This technique is applied to designing an image understanding system for finding and analyzing sky regions in images and video sequences of images of road scenes. The designed image understanding system finds the sky in video sequences of suburban and country roads very efficiently. The results can be found on the sites [9, 10].

2. A brief description of the geometrized histograms method

This method combines the advantages of statistical methods connected with studying conventional histograms of color or multichannel images [11] (real-time results) and conventional segmentation methods based on regions and contours [1, 2] (detailed shape-description). It was designed keeping in mind the application to constructing real-time image understanding systems. The origin of the method is dated back to the last 1980s, when the very early version of the method was applied to designing a vision system of a pilotless vehicle [12]. In addition, many papers are devoted to the problem of separating contours that belong to the boundaries of real regions. For example it is the main point in [13] and other papers can be found in the references of this paper. This problem can also be solved within the scope of the geometrized histograms method. Moreover each such contour can be furnished with the data characterizing the part of the region which it bounds (the so-called left and right contrast boundary curves). This data is very convenient for constructing real objects from different parts in the case of complex illumination, using both intensity-color characteristics and shape description of the considered parts.

A detailed description of the geometrized histograms method can be found in [5-7]. Let us explain briefly the concept of the geometrized histogram of a color image. To construct the geometrized histogram, the image is divided into strips St_i, $i = 1,\ldots n$, of the same width W with boundaries parallel to the horizontal or vertical axis of the image plane Os. Suppose that we deal with horizontal strips. The case of vertical strips is considered in a similar way. To describe approximately the image in a chosen narrow image strip, it is necessary to describe approximately the distribution of values of the vector function specifying it. The vector functions (R, G, B), (H, S, I), or $(G/(G+B), G/(G+R), I)$, introduced by the author, can be examples of this vector function. This approximate description will be called the geometrized histogram of the image in the strip. Let us explain first how to construct the geometrized histogram for a scalar function $f(x, y)$, giving a grayscale image. The geometrized histogram describes approximately the level sets L_z of $f(x, y)$, i.e., the set of points (x, y) of the strip $\boldsymbol{St}_n$, where $f(x, y) = z$. Since we deal with the discrete representation of the image, the projection of L_z onto Os is a union of intervals (segments) I_{kz} on this axis $Pr(L_z) = \cup_k I_{kz}$. For each segment I_{kz}, its cardinality is the number of the points of the level set L_z in the strip $\boldsymbol{St}_n$ that are projected onto this interval. It is clear that the set of cardinalities of the intervals I_{kz} for all possible z determines the classical histogram of $f(x, y)$ in the strip $\boldsymbol{St}_n$. The collection of intervals I_{kz} approximately describes L_z, since the set of level z belongs to the preimage of $\cup_k I_{kz}$, $L_z \subset Pr^{-1}(\cup_k \mathbf{I}_{kz})$, and the strip is narrow. The union of I_{kz} for all z determines the space of intervals on Os with the scalar function of cardinality on them. Note that intervals I_{kz} for different z may have a nonempty intersection on Os. This occurs when the intervals correspond to different objects in the strip and one object lies over another in it. The space of intervals $\cup_k^z I_{kz}$ is called the local geometrized histogram ($\mathbf{HG}_n$) of $f(x, y)$ in $\boldsymbol{St}_n$.

Let us show how to generalize this construction to the case of a vector function giving a color image. We deal with the function $(G/(G+B), G/(G+R), I)$ [5, 7], representing the color image. Let us introduce a characteristic function CF. If the hue of the point belongs to the yellow part of the color triangle, then CF coincides with $G/(G+B)$. When passing to the next range (green, blue, red), the value of $G/(G+B)$ is shifted by M, where M is the number of grades of the function $G/(G+B)$. The

geometrized histogram of *CF*, added for each interval I_{kz} by the classical histogram of the other color component *G/(G+R)*, is called the geometrized histogram of the color image in $\boldsymbol{St}_{\text{n}}$. Each interval I_{kz} of the geometrized histogram of *CF* is called the localization interval $\mathbf{Int}_{kz} = [beg_{kz}, end_{kz}]$ of the geometrized histogram of the color image in $\boldsymbol{St}_{\text{n}}$. Since each interval I_{kz} is furnished with the classical histogram of the other color component *G/(G+R)*, we can attach to $\mathbf{Int}_{kz}$ definite ranges of color characteristics and the mean values of these color features. Therefore, for each localization interval $\mathbf{Int}_{kz}$, it is possible to find the range and the mean value of its hue $\Delta_{\text{H}}^{kz} = [H_{min}^{kz}, H_{max}^{kz}]$ and H_{mean}^{kz}, the range and the mean value of its saturation $\Delta_{\text{S}}^{kz} = [S_{min}^{kz}, S_{max}^{kz}]$ and S_{mean}^{kz}, and the range and the mean value of its grayscale intensity $\Delta_{\text{I}}^{kz} = [I_{min}^{kz}, I_{max}^{kz}]$ and I_{mean}^{kz} [5, 7]. In addition, each interval of the geometrized histogram has the cardinality $Card^{kz}$.

Usually, there are too many intervals of the geometrized histogram I_{kz} to solve real problems. To reduce the number of them, a clustering procedure is introduced [5, 7], which joins intervals $\mathbf{Int}_{kz}$ that are close as intervals on *Os* and have close intensity-color characteristics. The joined intervals are called color bunches. Each strip St_{i} is described by the set of color bunches B_{i}. Each color bunch $b \in B_{\text{i}}$ is characterized by the following parameters:

1. the localization interval $int_{\text{b}} = [beg_b, end_b]$, belonging to *Os*;
2. $\Delta_{\text{H}}^{b} = [H_{min}^{b}, H_{max}^{b}]$ and H_{mean}^{b} – the range and the mean value of the hue of *b*;
3. $\Delta_{\text{S}}^{b} = [S_{min}^{b}, S_{max}^{b}]$ and S_{mean}^{b} – the range and mean value of saturation;
4. $\Delta_{\text{I}}^{b} = [I_{min}^{b}, I_{max}^{b}]$ and I_{mean}^{b} – the range and the mean value of the grayscale intensity;
5. the cardinality $Card^{b}$ (approximately, the number of points in the strip St_{i} whose coordinate *x* belongs to the localization interval $[beg_b, end_b]$ that have the color characteristics belonging to the ranges Δ_{H}^{b}, Δ_{S}^{b}, and Δ_{I}^{b} of the color bunch).

In this way, we can attach to each color image the graph of color bunches *STG* (STructural Graph). $B = \cup B_{\text{i}}$ is the set on nodes of *STG*. Color bunches b_1 and b_2 lying in the same strip are called adjacent if their localization intervals int_{b1} and int_{b2} are adjacent. Color bunches lying in the adjacent strips are called adjacent if their localization intervals have nonempty intersection. Edges of *STG* join all adjacent color bunches.

Informally, each bunch describes a certain part of a real object in the strip, its projection on *Os* and the description of numerical characteristics of this part of the object. The graph *STG* can be interpreted geometrically by superimposing localization intervals of bunches ($[beg_b, end_b]$), belonging to it, on the middle lines of the corresponding strips. Figure 1 demonstrates the representation of an image by the *STG* graph. Color bunches of each strip are superimposed near middle lines of the corresponding strips of the grayscale, image corresponding to the considered color image.

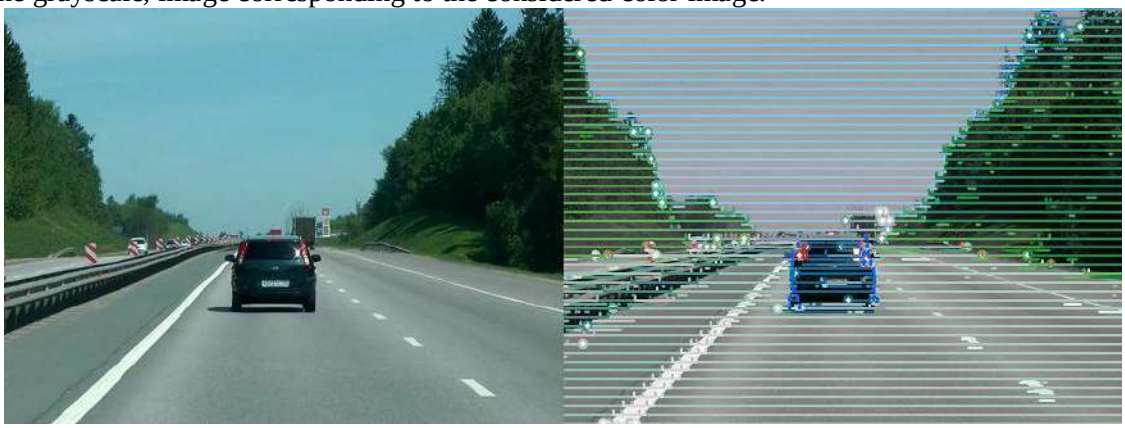


Figure 1. A road scene and the corresponding image of color bunches of the *STG* graph.

There are two types of color bunches. Color bunches of the first type are called dominating bunches. A dominating bunch is a bunch that at some points of its localization interval int_{b} has a maximum density $dens_{\text{b}} = Card^{b}/l(int_{\text{b}})$, where $l(int_{\text{b}})$ is the length of the interval int_{b}. It is clear that the localization intervals of dominating bunches generate a covering of the middle line of the corresponding strip. In this visualization, the localization intervals of dominating bunches are

superimposed on the entire middle lines of strips. In addition, we have some kinds of color bunches that are not dominating. These bunches may also be very important. For example as a rule, the signal zones of a distant vehicle (side-lights, brake lights) may have densities less than the densities of bunches, corresponding to the body of the vehicle. However, these bunches are very important in order to recognize the next actions of the driver of a vehicle going in front of our car. In the visualization, the color bunches of the second type are put slightly below the middle line. The procedure of construction of color bunches was prepared keeping in mind the possibility of detecting any connected colored set having a contrast with surrounding objects in the image. Numerous experiments with images have shown that color bunches represent any connected color object that have a contrast with in the real image with the size greater than three pixels. The description of a color image by color bunches compresses the information on images from millions of pixels to several hundreds of bunches. However, this image description contains all important features of the image, including a description of the geometry of objects belonging to it.

2.1. Continuous object on STG

In [6, 7] the concepts of left and right contrast curves (left and right germs of global contrast objects) in *STG* were introduced, and a bipartite graph of left and right contrast curves *LRG* was constructed. Let the image be divided into horizontal strips. A left (right) contrast curve is a chain of color bunches b_i with contrast right (left) neighbors located in adjacent strips St_i $i = k, k + 1, \ldots k + d$, such that the intensity-color characteristics of these bunches vary continuously from strip to strip, as well as the coordinate x of their left (right) ends [6, 7]. Such a chain is constructed upward, beginning from its lowest strip, finding the continuous extension of the previous bunch to the next strip [6, 7]. Figure 2 presents two examples of left (right) contrast curves (germs of contrast objects) in images.

Figure 2. Two parts of the sky represented by germs of global contrast objects in *STG*.

In this way, up to 256 different left and right contrast curves are found. By the construction, no more than one left and one right germ of global contrast objects can pass through any color bunch. On the set of all color bunches, functions $\mathrm{Germ}_{\mathrm{left}}(STG)$ and $\mathrm{Germ}_{\mathrm{right}}(STG)$ are determined. At each color bunch, these functions take as the value the number of the left (right) germ passing through this bunch or –1, if there is no such germ. Each color bunch of a left (right) contrast curve has a contrast contact with its right (left) neighbor. It is supposed that each left (right) contrast curve is a left (right) part of a certain hypothetical global object. Any left (right) contrast curve has its own linear geometric pattern determined by the discrete set of left (right) ends of the localization intervals of its color bunches. In the right image of Fig. 2, the presented left contrast curve (painted for visibility by dark intervals) is simultaneously a right contrast curve, since each of bunches belonging to it has both left and right contrast neighbors (parts of the forest or the boundary of the frame). Left and right ends of color bunches of the contrast curve of the right side of Fig. 2 specify standard boundaries of a sky region in a road scene. The right contrast curve in the left image of Fig. 2 has to be completed to generate the whole sky region. Of course, it is the most typical situation. In what follows, we present a reasoning system that performs the operation of extension sky regions. Together with a linear pattern, each germ of a global object G has an area pattern determined by the figure in the image plane generated by the localization intervals of its color bunches. For each germ G of a contrast object, we define its weight $W_G = \Sigma_j\, l(int_{bj})$ as the sum of lengths of localization intervals of its color bunches b_j. The substantial characteristic of an object in a perspective image is its behavior at infinity (in the motion to the upper

boundary of the image). To determine this behavior, the image is divided into zones by straight lines parallel to its lower boundary. Several strips of the image can be involved in each zone. Denote by W_G^i the part of the weight of G belonging to the zone with number i. The sequence $\{W_G^i\}$ determines the behavior of G at infinity. Both the linear and area patterns determine the full geometric pattern of G. In addition to the geometric pattern, each contrast curve has intensity-color characteristics determined by its color bunches. For the image of the right side of Fig. 2, together with the geometry of boundaries, we can produce the label "bright blue sky without clouds". To be able to assemble the sky region from available left and right germs of global contrast objects, we have to explain a new technique developed for this purpose in [6–8].

2.2 Adjacency relations graphs

In each strip, we are able to select among all color bunches a basic set of bunches dominating in some part of the strip (having the greatest density in it). It is obvious that localization intervals of dominating bunches give a covering of the middle line of the corresponding strip. For each dominating bunch, it is possible to find its closest left and right dominating neighbors. Using this construction, we can select a completely ordered basic subset of dominating color bunches that provide a covering of the middle line. It is possible to introduce a complete ordering in this basic subset and to number dominating color bunches of this subset from 0 to a certain k. Figure 3 demonstrates basic subsets for two strips of the image of Fig. 1. In addition, Figure 3 shows that all important parts of objects in these strips are taken into account in the descriptions of strips by color bunches.

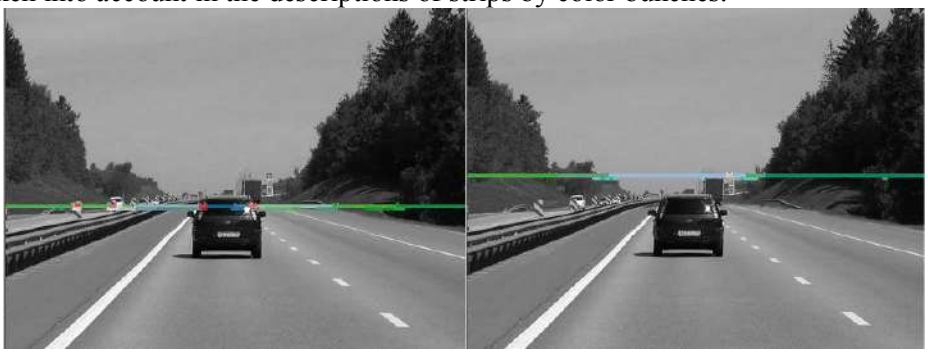


Figure 3. Two basic sets of color bunches in two different strips.

All linear ordered basic subsets of bunches, joined for all strips, generate on the image a "search lattice" SeachLat (*STG*) [8]. The constructed SeachLat (*STG*) (bunches are numbered with preservation of the adjacency relation) allows one to construct the adjacency graph *ADG*, which determines adjacency relation for left (right) germs of contrast objects in *STG*.

Each left (right) contrast curve (germ of global contrast object) is a continuous sequence of color bunches in a chain of adjacent strips (see Fig. 2). The values coordinates of left (right) ends of the localization intervals vary continuously, as well as their intensity-color characteristics [6]. By construction, only one left or right contrast curve can pass through any color bunch. If this curve exists, then it is uniquely determined by the functions $\mathrm{Germ}_{\mathrm{left}}(STG)$ and $\mathrm{Germ}_{\mathrm{right}}(STG)$. Suppose that we have a germ of a contrast global object G. Starting from the first bunch b_1 of this germ in its first strip and moving to the left and right of it, we find all adjacent germs of G in the considered strip. In this way, considering the germs passing through the direct neighbors of b_1, we can construct the direct adjacent germs of G in the strip. Moving from strip to strip, we are able to construct the part of the adjacency graph *ADG* connected with the left and right adjacent germs.

Consider the extension of adjacency relations in the downward and upward directions. In the construction of the set of color bunches in any strip, we generated a structure that informs us about all color bunches that pass through a definite point belonging the axis Os (the middle line of this strip). For each color bunch, we are able to find its first and last adjacent bunches in SeachLat (*STG*) in the upper and lower adjacent strips, based on this structure. Using this information, we are able to extend the adjacency relations downward and upward. Considering in each strip adjacent germs, passing through the next bunches of the search lattice, we are able to introduce a multiple adjacency graph

MADG or adjacency graphs of higher orders $\mathrm{Adg}_i(STG)$. *MADG* makes it possible to perform global image analysis, e.g., to analyze components of the same global object even in the case of occlusion [8]. For example, we are able to investigate two roadsides (left and right) simultaneously or two parts of any object separated by occlusion. The graph *MADG* makes it possible to assemble complex real objects which contain heterogeneous parts. In this graph, not only relations between objects that have common boundaries are established, but between objects separated by occlusion as well. New results connected with a detailed construction and application of *ADG* and *MADG* can be found in [8].

3. Construction of a reasoning system for finding the sky in road scenes

The problem of finding the sky is one of the problems solved in the course of developing the control system of the autonomous robot AvtoNiva, produced by a research group in Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences. For this purpose, it is not necessary to obtain a detailed description of the sky region at the pixel level. We need only approximate, qualitative and semantic description of the sky region that can be used in the control system for qualitative estimation of the road neighborhood and for detecting possible occlusion caused by unpredictable actions of other participants of the traffic. The statement of the problem under these conditions is described in the next subsection.

3.1 Problem statement and quality estimation

It is supposed that the image of a road scene is divided into a number of strips of the same width with the boundaries parallel to the horizontal axis of the image plane. For example for an image of resolution 640x480, we used 48 strips. We have to determine an array *Boun*(*n*), which specifies the pixel boundary of the sky for each column *n* of the image array. It is not supposed that the sky region is simply-connected. Due to occlusion it may contain several components. We have to find approximate color and intensity characteristics of each connected component and to describe its possible semantic type, e.g., “bright blue sky without clouds”. If the detected region of the sky in the form specified above takes into account about 90% of real pixel sky region (with minimum possible false positives) and the lower boundary of the sky region is found with the accuracy up to one strip, then the solution found is considered as quite successful.

Using this data, describing the character of the sky boundary, we can obtain certain useful information about the road behavior (a straight road, a forthcoming turn, descent, ascent, etc.) even in the case of heavy occlusion of the road caused by other vehicles. We are also able to recognize the dangerous occlusion caused by the car in front, taking into account among other features the complete change in the pattern of the sky region. It is also supposed that the problem has to be solved on a standard PC in real time. Since the accuracy of determining the sky boundary is up to one strip, it is proposed to find the set of color bunches B_S belonging to the sky region. Then to find the pixel boundary in each column *n*, the lowest bunch $b_l \in B_S$ passing through this column is found. We take the lower boundary of the strip within the localization interval of b_l as the pixel sky boundary *Boun*(*n*) of the sky region in this place. These assumptions make it possible to solve the problem in a real-time mode. It is clear that the boundary of the sky is given by a piecewise constant function of *n*.

3.2 Specific features of the problem and its solution

The careful study of a large number of images of scenes containing sky regions, taken under different illumination conditions, at different times of the day, and during different seasons, has shown that the sky region may be a very complex object. It may contain many different more or less homogeneous parts that are quite different in color and intensity. The sky region is especially diverse during sunset or sunrise. The appearance of clouds may change the pattern dramatically. The presence of sky-similar objects such as walls of buildings (especially without windows) makes the problem even more complex. Under these conditions, it may be impossible to solve the problem using only one frame. Sometimes even a human may fail to determine the boundary of the sky quite correctly using only one image. Only external knowledge and views from other positions may help. For instance in Fig. 4, a part of a circle of the white antenna left of the low white clouds over the roof of a Sberbank building gives an example of such a situation.

Therefore, to obtain an adequate solution, at least two stages of solving the problem are necessary. At the first stage, a single frame is analyzed and a preliminary solution is described. At the second stage, we compare and study a set of adjacent frames in order to provide the final solution.

Let us describe the first stage of solution. At the first step, we generate a preliminary conceptual and semantic description of all left and right contrast curves constructed by algorithms described in [6, 7]. To describe the geometry of any contrast curve, we use methods proposed in [8].

Let us briefly describe them. For this purpose, we divide the boundary points of any contrast of curves (a set of left (right) ends of localization intervals of the color bunches involved in this contrast curve) into branches on which the coordinate x (the horizontal coordinate in the image plane) of its nodes increases or decreases. This is aimed at finding the perspective in the image. As the additional constraint, we suppose that the absolute values of the differences $\mathrm{abs}(end_{b(k+1)} - end_{bk})$ (right curves) or $\mathrm{abs}(beg_{b(k+1)} - beg_{bk})$ (left curves) for the adjacent nodes of the curve on these branches are bounded by a constant connected with the width of the strip. Introducing these constraints, we eliminate the effect of sharp change of the shape of the boundary curve.

Then we test whether these branches belong to certain straight line segments or they are convex or concave. To test the linear hypothesis, we use histograms of inclines of the segments connecting adjacent nodes of the contrast curves. Details can be found in [8]. As was mentioned above, the distribution of lengths of localization intervals along the curve or the sums of lengths W_G^i within selected zones determines the behavior of the corresponding germ of a global object at infinity. The parameters of this distribution distinguish contrast curves that are long and narrow with decreasing lengths (like parts of the road) and long and wide with increasing and non-decreasing lengths (like forests, fields, parts of the sky, bodies of cars). Then we select both the left and right contrast curves with the maximum weight $W(b)$ that have color-intensity characteristics possible for parts of the sky and locate in the top part of the frame and have the corresponding behavior at infinity. Based on the left and right curves of the maximum weight found, using the search lattice on the image, we construct the whole region of the sky. Moving to the left or right on the search lattice, we add sky-similar germs and stop the extension of the sky region when the regions classified as forests, fields, roads, etc., occur. In a similar way, moving on the search lattice to the bottom of the image, we add sky-similar germs again and stop the extension when the regions mentioned above are met. It is especially difficult to eliminate regions generated by buildings, having intensity-color characteristics similar to those of sky regions. For this purpose, the reasoning system finds sky-similar regions with straight boundaries and tests whether these regions have subobjects inside with vertical boundaries (windows, doors).

To study suspicious regions, we need some new definitions. For this purpose, we introduce concepts of contours in *STG* having a rather big angle with the axis *Os* (the boundary lines of strips into which the image is divided). For horizontal (vertical) strips, we obtain contours in *STG* that are close to vertical (horizontal) ones. In turn, these contours give the corresponding vertical (horizontal) contours in the image if we consider pixel coordinates of the ends of the corresponding color bunches constituting the contours in *STG*. Let us give the definition of contours close to perpendicular to the axis *Os* (simply, contours in what follows) in *STG*. These contours are generated by left (right) ends of basic color bunches, belonging to the search lattice SeachLat. For each strip of the image, an array $Loc[i]$ of length DimX/k is generated, where DimX is the horizontal dimension of the image array, while k is a compression coefficient (e.g., $k = 4$). At each point i, $Loc[i] = d$, where d is the number of the basic bunch with localization interval passing through the point ki of the middle line of the corresponding strip. Consider a strip St_k. Let $b_k \in$ SeachLat be a basic color bunch with the localization interval $[beg_{bk}, end_{bk}]$. Remind that image strips are numbered bottom to top. Consider the next strip St_{k+1}. Using the array $Loc[i]$ of the next strip, we find the basic color bunch that passes through beg_{bk} (end_{bk}) in the next strip. Then moving along SeachLat, we find the basic color bunch b_{k+1} such that the distance between beg_{bk} (end_{bk}) and $beg_{b(k+1)}$ ($end_{bk(k+1)}$) is minimal. If this distance is less than a certain constant that bounds the angle of the shift, the contour is extended to b_{k+1}. Using left (right) ends of bunches, we obtain left (right) contours. For a contour of length n, the following characteristics are introduced: 1. the maximum deviation from the vertical max_i $\mathrm{abs}(end_{bk} - end_{bk+i})$; 2. the total deviation from the vertical direction $dtot = \mathrm{abs}(end_{bk} - end_{bk+n})/\mathrm{n}$.

Definition. A contour is close to the perpendicular to the axis *Os* if *dmax* and *dtot* are bounded by certain constants.

If we consider two graphs of color bunches STG_V and STG_H, constructed for vertical and horizontal strips, and find in both graphs contours close to perpendicular to the axes *Os*, we can generate in the image contours close to vertical and horizontal simultaneously. This technique is employed to select buildings in the image and to eliminate their regions from the sky region. Figure 4 demonstrates two complex examples of images of a city landscape.

In spite of several small mistakes, the level of the sky in both images is found quite correctly. The results of processing video sequences by the presented system can be found in [9, 10]. It is important to note that the results completely support the conclusion that the problem of finding the sky in images is not local and requires careful global analysis of the frame. At the end of the first stage, we find the boundary of the sky region in the form of an array specifying the number of the first pixel of the sky region in each column of the image *Boun*(*n*), where *n* is the number of the corresponding column. We also have the set *sk_germs* that contains of all germs (contrast curves) included in the sky region. Analyzing the parameters of the germs of the set, we produce the semantic description of the sky image.

Figure 4. Sky regions in images of a city landscape.

At the second stage of the solution, we compare arrays *Boun*(*n*) for the current and previous frames. We also compare the semantic descriptions of the sky regions found of adjacent frames. In the case of their correspondence, we adopt a new solution. Otherwise, we analyze the differences and decide whether a dangerous occlusion occurs, taking into account other features such as possible signal zones (side-lights, brake signals) and vertical and horizontal contours of a hypothetical vehicle in front.

The technique of contours close to vertical or horizontal ones can also be applied to the analysis of scenes in villages and towns to detect fences, cottages in order find roads in complex conditions of shadows and absence of road marking. This will be a subject of the next publications.

4. Software implementation, demonstration of the results and discussion

The image understanding system has been implemented by a program written in C++ and operating under Windows and Linux. This program processes video sequences in real time on standard computers with processors I3–I7 and records the results for each frame of the video sequence tested. For frames of resolution 640x480, the operation speed is about 20 fps. The program has been tested on dozens of video sequences taken from cars on different Russian roads under different seasons, times of the day, under different illumination conditions. Figure 5 presents several examples from records of the results for three video sequences. On country roads from the considered series the percentage of positive results varied from 98 to 100 %. Even on new video sequences of this type processed for the first time without modifying the program, a very high percentage of positive results were obtained. Some problems may appear when processing video sequences taken in villages and towns. This is connected with buildings with walls that cannot be distinguished from the adjacent parts of the sky using only one frame. The data base of rules and features are being modified, as well as the work with adjacent frames, in order to eliminate these problems. The results of processing several video sequences can be found in [9, 10]. In addition to the sky region, the video system of the control system

of AvtoNiva finds other regions interesting for controlling the vehicle such as the boundaries of the vegetation regions, road regions, and the other vehicles on the road. The solution to a part of these tasks was described in [8] and further results were presented in a brief publication [14]. The detailed publication on this topic is being prepared.

Figure 5. Examples of records of experiments with video sequences.

5. References

[1] Forsyth D A and Ponce J 2003 *Computer Vision, a Modern Approach* (London: Prentice Hall)
[2] Mishra A K and Aloimonos Y 2009 Active segmentation *Int. J. Humanoid Rob.* **6** 361-366
[3] Chen Ch, Papandreou G, Kokkinos I, Murphy K, and Yuille A L 2016 Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs *Preprint arXiv 1606.00915*
[4] Divvala S K 2012 *Context and subcategories for sliding window object recognition* PhD Thesis (Pittsburgh: Carnegie Mellon University)
[5] Kiy K I 2010 A new real-time method for description and generalized segmentation of color images *Pattern Recognit. Image Anal.* **20** 169-178
[6] Kiy K I 2015 Segmentation and detection of contrast objects and their application in robot navigation *Pattern Recognit. Image Anal.* **22** 338-346
[7] Kiy K I 2015 A new real-time method of contextual image description and its application in robot navigation and intelligent control *Computer Vision in Control Systems-2 Innovations in Practice Intelligent Systems Reference Library* **75** 109-133
[8] Kiy K I 2018 A new method of global image analysis and its application in understanding road scenes *Pattern Recognit. Image Anal.* **25**
[9] *Electronic Materials* (Access mode: http://video.mail.ru/kikip_46/_myvideo)
[10] *Electronic Materials* (Access mode: https://www.facebook.com/100004887018729/videos)
[11] Denisova A Y and Sergeev V V 2016 Algorithms for calculating multichannel image histograms using hierarchical data structures *Computer Optics* **40(4)** 535-542 DOI: 10.18287/2412-6179-2016-40-4-535-542
[12] Kiy K I, Klimantovich A V and Buivolov G A 1995 Vision-based system for road following in real time *Proc. 7th Int. Conf. on Advanced Robotics (San Feliu de Goixols, Catalonia, Spain)* **1** 517
[13] Belim S V and Kutlunin P E 2015 Boundary extraction in images using a clustering algorithm *Computer Optics* **39(1)** 119-124 DOI: 10.18287/0134-2452-2015-39-1-119-124
[14] Kiy K I 2018 Image understanding systems based on the geometrized histograms method *Proc. 7th Int. Conf. on Extreme Robotics and Conversion Tendencies (Saint-Petersburg)* 140

Acknowledgments

This work was supported by the Russian Foundation for Basic Research, projects no. 16-08-00880, 16-07-01264a, and 18-07-00127.

Accurate reconstruction of the 3D indoor environment map with a RGB-D camera based on multiple ICP

A N Ruchay[1,2], K A Dorofeev[2], A V Kober[1]

[1]Federal Research Centre of Biological Systems and Agro-technologies of the Russian Academy of Sciences, 9 Yanvarya street 29, Orenburg, Russia, 460000

[2]Department of Mathematics, Chelyabinsk State University, Bratiev Kashirinykh street 129, Chelyabinsk, Russia, 454001

Abstract. In this paper, we propose a new method for 3D map reconstruction using the Kinect sensor based on multiple ICP. The Kinect sensor provides RGB images as well as depth images. Since the depth and RGB color images are captured by one Kinect sensor with multiple views, each depth image should be related to the color image. After matching of the images (registration), point-to-point corresponding between two depth images is found, and they can be combined and represented in the 3D space. In order to obtain a dense 3D map of the 3D indoor environment, we design an algorithm to combine information from multiple views of the Kinect sensor. First, features extracted from color and depth images are used to localize them in a 3D scene. Next, Iterative Closest Point (ICP) algorithm is used to align all frames. As a result, a new frame is added to the dense 3D model. However, the spatial distribution and resolution of depth data affect to the performance of 3D scene reconstruction system based on ICP. In this paper we automatically divide the depth data into sub-clouds with similar resolution, to align them separately, and unify in the entire points cloud. This method is called the multiple ICP. The presented computer simulation results show an improvement in accuracy of 3D map reconstruction using real data.

1. Introduction

Surface reconstruction of real-world objects by using RGB-D sensors such as the Kinect is one of the most important topics in computer vision [1, 2, 3, 4]. These cameras can provide a high resolution RGB color image with depth data of the environment. In contrast to the stereo and monocular cameras, RGB-D sensors are able to obtain depth data directly, even in environments with poor visual textures. A full 3D mapping system that models indoor environment utilizing a joint optimization algorithm combining visual features and shape-based alignment was proposed in [5, 6, 7]. Accurate real-time mapping algorithm for complex indoor scenes in variable lighting conditions was suggected in [8]. It is a frame-to-global method and it maintains the single scene model with a global volumetric, truncated signed distance function representation.

Building dense 3D maps of environments is an important task for mobile robotics, with applications in navigation, manipulation, semantic mapping, face recognition [9, 10, 11, 12, 13, 14, 15, 16, 17], and in augmented reality applications, surveillance systems, medical applications [18, 19, 20, 21, 22].

Typically, a 3D mapping system based on RGB-D cameras starts with detection and analysis of distinctive points of the environment. These points are visual marks helping to create the map. However, due to the diversity of situations and conditions of the problem, such as noise introduced in the capture step, lighting changes, scaling, rotation, and other transformations, the most appropriate detector/descriptor for extraction of visual features has not yet been determined. So, it is important to identify the best visual marks of the environment. It requires a descriptor that provides the transformation invariance of points of interest, and stable detection of the same points along sequences with different variations in images.

Once the key points have been identified and matched with the points in the following frame, the next step is to align the three-dimensional data in consecutive frames, generally using the Iterative Closest Point (ICP) algorithm [23]. ICP based algorithms require an initial guess of the rigid transformation between two consecutive frames. This initial transformation, rotation and translation between two adjacent frames, can be calculated by the random sample consensus (RANSAC) method [23].

We are interested in improving the quality of the 3D map modeling [24, 25, 26, 27, 28, 29]. The following methods are proposed [30, 31, 32, 23, 3] to reduce the number of points to be processed, to select the best points, to improve the accuracy of location of the points, to improve the accuracy of alignment of frames and cloud points, to improve the map quality. In this paper we focus on a strategy to improve the accuracy in the alignment of 3D data by decomposing the data along depth axis and followed by the ICP method. This method is referred as multiple ICP.

A method for improving the accuracy of a global alignment with the ICP was proposed in paper [23]. Because depth data have different resolution, the depth data decomposed into several point regions (sub-clouds) with a similar resolution yields a better 3D scene restoration in terms of the accuracy. Also, one can observe that with a smaller number of points, a better result can be achieved than using all inliers generated by RANSAC. However the problem of automatical dividing of the depth data into sub-clouds with similar resolution has not solved yet. Therefore in this paper we proposed a new algorithm for automatical dividing the depth data into sub-clouds with a similar resolution.

The paper is organized as follows. Section 2 discusses related approaches of improving ICP algorithm. Section 3 presents the proposed algorithm of multiple ICP and the algorithm for automatic dividing of the depth data. Section 4 describes computer simulation results of accuracy of 3D map reconstruction using real data. Finally, Section 5 presents our conclusions.

2. Related work

An improved ICP algorithm that can automatically register multiple 3D data sets from unknown viewpoints is presented in [33]. The sensor projection that represents the mapping of the 3D data into its associated range image and a cross projection are used to determine the overlapping region of two range data sets. A combining ICP algorithm with the sensor projection makes an automatic registration of multiple 3D sets without pre-procedures that are prone to errors and any mechanical positioning device or manual assistance.

An ICP algorithm that uses matching points' center of gravity as reference points and point pair distance constraint to reject false point pairs for point cloud registration was proposed [34]. Authors use the nearest criterion to remove point pairs which contain same points by close point criterion in order to improve speed and accuracy of point cloud registration.

The paper [35] proposed an improved solution to overcome the problem, that the ICP algorithm is to require complicated computation steps. The proposed algorithm selects the best sample points in oder to reduce computation time while still retains accuracy.

An efficient and robust subset-ICP rigid registration method is proposed to reduce the computation complexity and improve the flexibility of ICP algorithm in the paper [36].

The paper [37] presented a more robust ICP algorithm. an extended ICP algorithm solve a problem of the least square registration by an inequality constraint of the rotation angle. A closed-form solution for the rotation is obtained according to the monotonicity of the model with respect to the rotation angle at each iterative step of the algorithm. The proposed approach extends the convergence domain of the ICP algorithm, and it can be used much more widely.

A variant of the ICP registration method by introducing a novel weight-bootstrap scheme was proposed in the paper [38]. The rigid transform between two point sets can be estimated, as long as proper correspondences are given. The accuracy of the estimated transform is mainly determined by the goodness of these matches. The proposed solution parameterizes the pair matching confidence and improves the optimization process to address the challenge, that the confidence of the correspondences is weakened by the observation error.

The paper [39] proposed a new approach for the registration of depth data with color data, which combines epipolar constraints with pointtoplane constraints to improve ICP algorithm and achieve accurate registration.

An improved ICP algorithm which based on k-d tree was presented for tree point cloud data registration [40]. The k-d tree is similar to the general tree structure and it can store, manage and search data efficiently. The proposed algorithm solves the problem, that massive point cloud data bring a great number of troubles to registration method. Also the improved ICP algorithm is presented in the paper, the proposed ICP algorithm is the combination of the original ICP algorithm and k-d tree [41].

3. The proposed algorithm of multiple ICP

In order to reconstruct 3D scene, first we solve the problem that depth data have different resolution [23]. As described above, once the key points have been identified and matched, the RANSAC algorithm is applied to remove outliers, and then the ICP method is applied to the data for aligning. Usually, supposing that transformation is orthogonal and objects are rigid, the ICP method is utilized over all inliers points provided by RANSAC. Note, that the use of a high number of points does not guarantee a better alignment. Moreover, it has been found that the relative motion between frames should be small enough for a good alignment. When the displacement between frames is relatively large, the algorithm takes a lot of time to converge or sometimes one diverges owing to big alignment errors. In this circumstance, the modeling map system yields a bad accuracy. A possible strategy for a lower computation time could be to reduce the number of points to be aligned. However, if the selected points are mainly located on the same deep plane, then the ICP yields wrong results.

On the other hand, it has been shown that the depth data provided by the Kinect sensor have different resolutions. With increasing distance from the sensor, the data resolution becomes worse [42].

With the aim of improving the accuracy in point clouds alignment, one can propose to divide data into regions with similar resolutions. In other words, in this manner we can get several sub-clouds and then align them separately. After, the entire points cloud is reconstructed by unifying of the aligned sub-clouds.

We show that by dividing the 3D data into sub-clouds, aligning them separately, and finally unifying them back, gives us a better 3D scene reconstruction than that if applying a single ICP for all inliers of the scene.

For each experiment presented on this work, the following basic steps carry out:

(i) Detection and matching of global key points in RGB data with SURF algorithm;

(ii) Remove outliers with RANSAC;

(iii) Align 3D data with ICP using the associate 3D points of the inliers;

(iv) Evaluate the quality of alignment;

(v) Construct 3D map.

In order to compare the accuracy of the alignment, ICP uses the root mean square error (RMSE) between two point clouds after each iteration,

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}|s_i - t_i|^2}, \tag{1}$$

where s_i and t_i are the corresponding pair points in two point clouds, and n the total number of valid pair points. The correspondence is determined by using the k-d tree search algorithm.

The proposed algorithm for automatic dividing the depth data into sub-clouds with similar resolution consists of several simple procedures written in Matlab.

(i) `AutoLength` divides the point cloud to n point sub-clouds evenly according to length of intervals. Generally, the point cloud contains points with depth values from 0 to 10 meters. So, a cloud can be divided into n equal sub-clouds with intervals $[0, 10/n), [10/n, 2 \cdot 10/n), \ldots, [9 \cdot 9/n, 10]$.

(ii) `AutoLengthIncrease` divides the point cloud into n point sub-clouds according to quadratic increase of $x, x^2, x^4, x^8, \ldots, x^{2^{n-1}}$ length of intervals of sub-clouds, respectively. We calculate x such a manner that $x + x^2 + x^4 + ... + x^{2^{n-1}} = 10$ (full length of the interval of interest from 0 to 10 meters).

(iii) `AutoLengthDecrease` divides the point cloud into n point sub-clouds according to quadratic decrease of $x^{2^{n-1}}, x^{2^{n-2}}, \ldots, x^4, x^2, x$ length of intervals of sub-clouds, respectively. We calculate x such a manner that $x + x^2 + x^4 + ... + x^{2^{n-1}} = 10$ (full length of the interval of interest from 0 to 10 meters).

(iv) Previous procedures do not take into account the number of points in each automatical divided sub-clouds. Sometimes it leads to serious problems because ICP algorithm does not owing to a sub-cloud contains a small number of points. A new procedure `AutoCountAllPoints` divides the point cloud into n sub-clouds with the same number of points. Let a point cloud contain points with depth values from 0 to 10 meters with the total number P of points. The point cloud can be automatically divided into n intervals with the number P/n of points based on a heuristic method which consistently increases the current interval of 0.01 meter. At each iteration we count the current number of points containing into the interval untill the number of point in sub-clouds reaches P/n.

(v) To divide the point cloud into n sub-clouds with the same number of inliers points from SURFF algorithm. Let a point cloud contain I inliers points (the total number of points is much larger). `AutoCountInlierPoints` divides the point cloud automatically into n intervals with the number I/n of points based on a heuristic method which consistently increases the current interval of 0.01 meter. At each iteration we count the current number of points containing into the interval untill the number of point in sub-clouds reaches I/n.

Note that by applying the ICP algorithm to sub-clouds a set of transformation matrices are obtained. The transformation matrices could be used in different ways for reconstruction of the 3D dense map of the environment: each matrix `tforms_i` from ICP algorithm to each i sub-cloud, or `tform_1` for all sub-clouds.

4. Computer simulation

In this section, computer simulation results of accuracy of 3D map reconstruction using real data are presented and discussed.

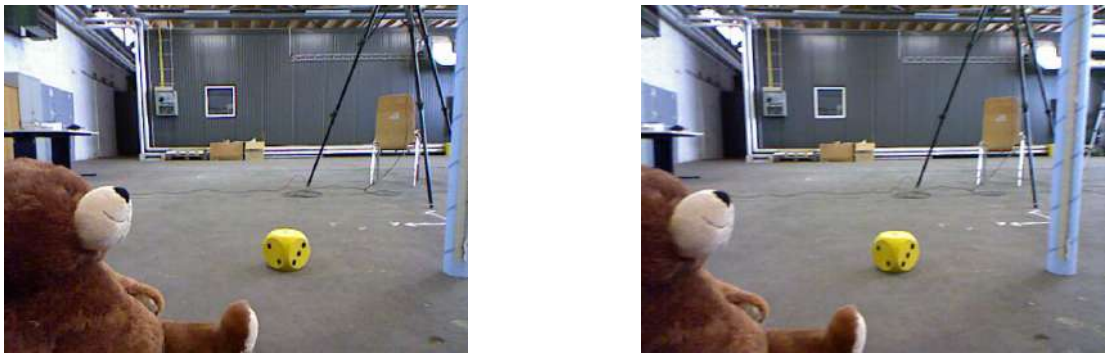


Figure 1. Pair of frames.

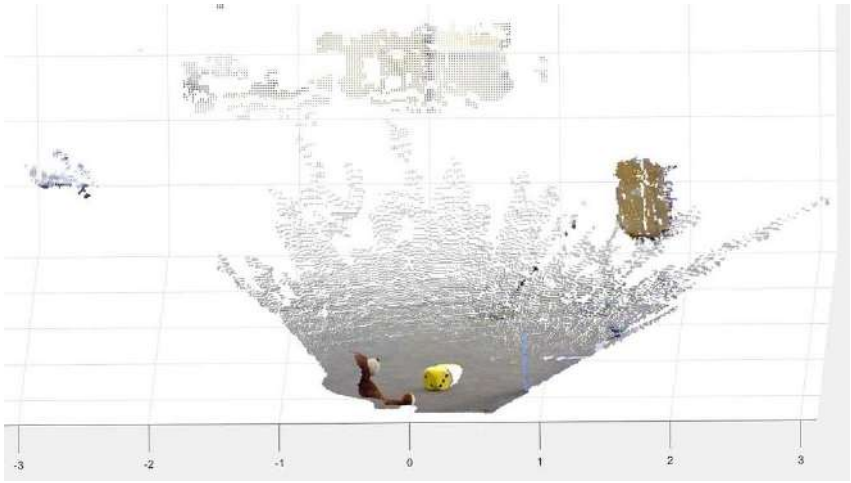


Figure 2. Point cloud from the first frame.

The tests were carried out with frames of "freiburg/pioneer slam", "freiburg2/pioneer slam", "freiburg3/pioneer slam" datasets of TUM RGB-D benchmark [4]. We calculate RMSE errors using real data for the proposed method of 3D map reconstruction based on multiple ICP and automatical dividing the depth data into sub-clouds. Fig. 1 shows the pair of frames to be aligned. If the point cloud is aligned, the data are transformed to the coordinate system of the ground truth for further evaluation using the RMSE. Fig. 2 shows the corresponding point cloud of the first frame.

Next we divide points clouds into sub-clouds by our proposed procedures. Fig. 3 shows results of automatical dividing into two point sub-clouds according to quadratic increase of the length and the distribution of all inlier valid SURFF points in the first frame after the RANSAC algorithm. The distribution of 283 and 153 inlier points for the sub-clouds is shown in Figs. 4. The procedure of automatical dividing into two point sub-clouds according to quadratic increase of the length yields the minimal RMSE errors comparing with other procedures of automatical dividing. We define a global alignment when the ICP uses all inliers for alignment of two point clouds. For the test frames in Fig. 1, we use 992 inlier points, which are distributed as shown in Fig. 3.

The transformation matrix computed with the help of the ICP algorithm for two frame is

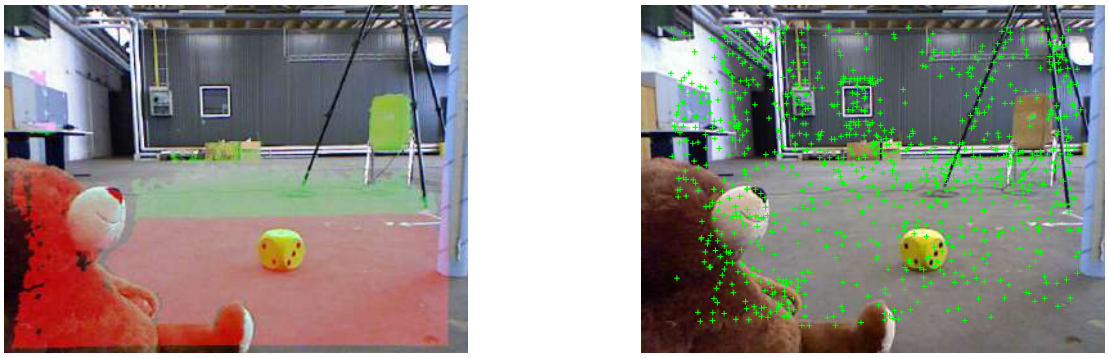


Figure 3. Left picture shows sub-clouds, and right picture shows all inlier points.

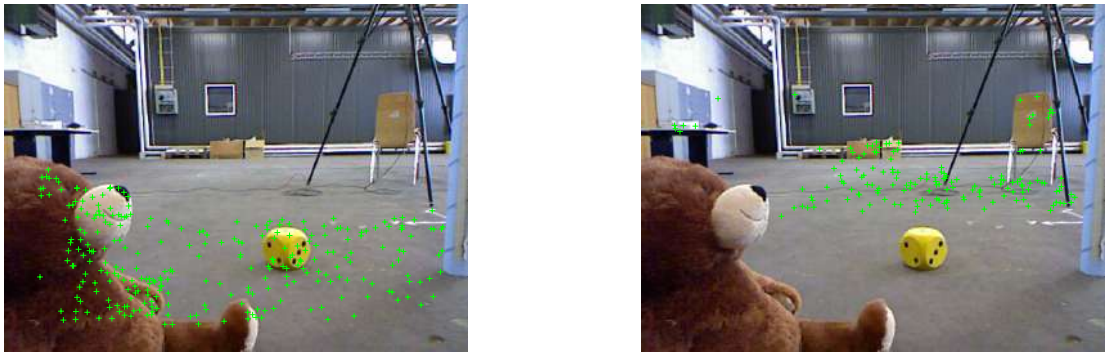


Figure 4. Inlier points for sub-clouds.

given as

$$\begin{pmatrix} 0.999381 & -0.012965 & -0.032708 & 0.000000 \\ 0.012638 & 0.999868 & -0.010166 & 0.000000 \\ 0.032836 & 0.009746 & 0.999413 & 0.000000 \\ -0.049280 & -0.024187 & 0.010850 & 1.00000 \end{pmatrix}$$

The multiple ICP algorithm returns the following transformation matrices for sub-clouds. Transformation matrix for the first sub-cloud is given as

$$\begin{pmatrix} 0.999920 & 0.000987 & -0.012579 & 0.000000 \\ -0.000981 & 0.999999 & 0.000433 & 0.000000 \\ 0.012579 & -0.000420 & 0.999921 & 0.000000 \\ -0.002609 & -0.000339 & 0.002048 & 1.000000 \end{pmatrix}$$

Transformation matrix for the second sub-cloud is given as

$$\begin{pmatrix} 0.999942 & -0.000517 & -0.010775 & 0.000000 \\ 0.000472 & 0.999991 & -0.004132 & 0.000000 \\ 0.010777 & 0.004126 & 0.999933 & 0.000000 \\ 0.002941 & -0.017613 & 0.002517 & 1.000000 \end{pmatrix}$$

The computed transformation is applied to the entire sub-clouds and compared with the ground truth. Finally, the reconstructed 3D model of the object is shown in Fig. 5.

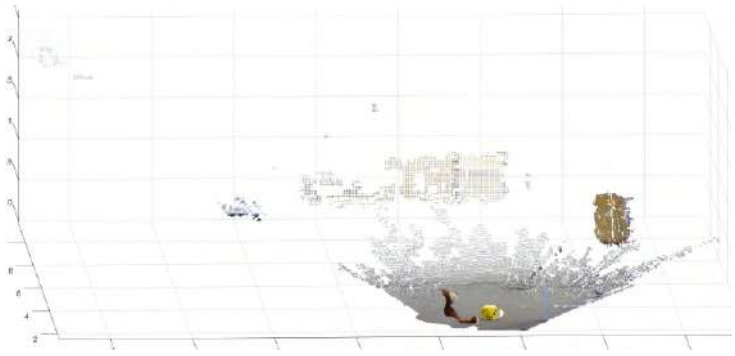


Figure 5. View of 3D scene reconstruction.

Table 1 shows RMSE errors for the proposed method of 3D map reconstruction based on multiple ICP and the various procedures of automatical dividing the depth data into sub-clouds with the real BEAR1 dataset [4].

Results of the proposed methods of automatical dividing into $n = 4$ and more sub-cloud were unsatisfactory with huge RMSE errors. The procedure `AutoLengthIncrease` with dividing the point cloud into $n = 2$ point sub-clouds according to quadratic increase of the length yields the best result in terms of RMSE. Note that the procedure can not be used for all datasets because of empty sub-cloud of inlier points. For other datasets the procedure `AutoCountInlierPoints` dividing the point cloud automatically into n intervals with the number I/n of points yields the best results. Finally, we can conclude that the accuracy of 3D reconstruction of a scene depends on the following reasons:

(i) Detection and matching of global points by feature detection algorithms. We evaluated SURFF, FAST, BRISK, Harris, MinEigen algorithms in terms of RMSE errors and the best results were obtained using SURFF algorithm, other algorithms return too few matched points;

(ii) Distribution of points in frame. If most points are close to the sensor position then no sense to divide point cloud into sub-clouds, a simple ICP algorithm provides good results.

5. Conclusion

We proposed a method of 3D map reconstruction based on multiple ICP and introduced various procedures for automatical dividing the depth data into sub-clouds. We evaluated RMSE errors using real data for the proposed method and procedures. The procedure `AutoLengthIncrease` which divides the point cloud into $n = 2$ point sub-clouds according to quadratic increase of length of intervals of sub-clouds yields the best result in terms of RMSE. In future, different datasets, especially with built accurate models will be studied and tested. Also we want to verify the performance of different descriptors for 3D scene restoration. Finally, the developed method will be incorporated into a complete SLAM system.

6. References

[1] Wang K, Zhang G and Bao H 2014 *IEEE Transactions on Image Processing* **23** 4893-4906
[2] Lee K R and Nguyen T 2016 *Mach. Vision Appl.* **27** 377-385
[3] Li H, Liu H, Cao N, Peng Y, Xie S, Luo J and Sun Y 2017 *International Journal of Advanced Robotic Systems* **14** 1729881417695560
[4] Cai Z, Han J, Liu L and Shao L 2017 *Multimedia Tools and Applications* **76** 4313-4355
[5] Henry P, Krainin M, Herbst E, Ren X and Fox D 2014 *RGB-D Mapping: Using Depth Cameras for Dense 3D Modeling of Indoor Environments* (Berlin, Heidelberg: Springer Berlin Heidelberg) 477-491

Table 1. RMSE of 3D map reconstruction for the tested algorithms using the BEAR1.

Algorithm	RMSE
SimpleICP	0.0397
n = 2 intervals	
`AutoLen tforms_i`	0.0172
`AutoLen tform_1`	0.0190
`AutoLenIncrease tform_i`	0.0082
`AutoLenIncrease tform_1`	**0.0074**
`AutoLenDecrease tform_i`	0.0184
`AutoLenDecrease tform_1`	0.0167
`AutoCountAllPoints tform_i`	0.0162
`AutoCountAllPoints tform_1`	0.0128
`AutoCountInlierPoints tform_i`	0.0111
`AutoCountInlierPoints tform_1`	0.0108
n = 3 intervals	
`AutoLen tform_i`	0.0177
`AutoLen tform_1`	0.0117
`AutoLenIncrease tform_i`	0.0122
`AutoLenIncrease tform_1`	0.0129
`AutoLenDecrease tform_i`	-
`AutoLenDecrease tform_1`	-
`AutoCountAllPoints tform_i`	0.0210
`AutoCountAllPoints tform_1`	0.0392
`AutoCountInlierPoints tform_i`	0.0101
`AutoCountInlierPoints tform_1`	0.0136
n = 4 intervals	
`AutoLen tform_i`	0.0144
`AutoLen tform_1`	0.0081
`AutoLenIncrease tform_i`	0.0130
`AutoLenIncrease tform_1`	0.0188
`AutoLenDecrease tform_i`	-
`AutoLenDecrease tform_1`	-
`AutoCountAllPoints tform_i`	0.0154
`AutoCountAllPoints tform_1`	0.0287
`AutoCountInlierPoints tform_i`	0.0180
`AutoCountInlierPoints tform_1`	0.0299

[6] Wang H, Wang J and Wang L 2016 *IEEE Conference on Computer Vision and Pattern Recognition* 3271-3279
[7] Li Y, Wang Y and Wang D 2017 *Computers & Electrical Engineering*
[8] Newcombe R A, Izadi S, Hilliges O, Molyneaux D, Kim D, Davison A J, Kohi P, Shotton J, Hodges S and Fitzgibbon A 2011 *10th IEEE International Symposium on Mixed and Augmented Reality* 127-136
[9] Echeagaray-Patron B A, Miramontes-Jaramillo D and Kober V 2015 *International Conference on Computational Science and Computational Intelligence (CSCI)* 843-844
[10] Echeagaray-Patron B A and Kober V 2015 *3D face recognition based on matching of facial surfaces* **9598** 95980V-8
[11] Sochenkov I and Vokhmintsev A 2015 Visual Duplicates Image Search for a Non-cooperative Person Recognition at a Distance **129** 440-445
[12] Vokhmintsev A, Makovetskii A, Kober V, Sochenkov I and Kuznetsov V 2015 A fusion algorithm for building three-dimensional maps **9599** 959929-7

[13] Tihonkih D, Makovetskii A and Kuznetsov V 2016 *A modified iterative closest point algorithm for shape registration* 9971 99712D-8
[14] Sochenkov I, Sochenkova A, Vokhmintsev A, Makovetskii A and Melnikov A 2016 *Effective indexing for face recognition* 9971 997124-9
[15] Picos K, Diaz-Ramirez V, Kober V, Montemayor A and Pantrigo J 2016 *Optical Engineering* **55** 55-55-11
[16] Echeagaray-Patron V K 2016 *Face recognition based on matching of local features on 3D dynamic range sequences* 9971 9971-6
[17] Echeagaray-Patrón B A, Kober V I, Karnaukhov V N and Kuznetsov V V 2017 *Journal of Communications Technology and Electronics* **62** 648-652
[18] Zou R, Ge X and Wang G 2016 *IEEE Chinese Guidance, Navigation and Control Conference (CGNCC)* 375-378
[19] Jun C, Kang J, Yeon S, Choi H, Chung T Y and Doh N L 2017 *Intelligent Automation & Soft Computing* **23** 207-218
[20] Guo K, Xu F, Yu T, Liu X, Dai Q and Liu Y 2017 *ACM Trans. Graph.* **36** 32:1-32:13
[21] Cheng S C, Su J Y, Chen J M and Hsieh J W 2017 *Model-Based 3D Scene Reconstruction Using a Moving RGB-D Camera* (Cham: Springer International Publishing) 214-225
[22] Holz D and Behnke S 2015 *European Conference on Mobile Robots (ECMR)* 1-8
[23] Gonzalez-Fraga J A, Kober V, Diaz-Ramirez V H, Gutierrez E and Alvarez-Xochihua O 2017 *Proc. SPIE* 10396 10396-7
[24] Aguilar-Gonzalez P M and Kober V 2011 *Optical Engineering* **50** 50-59
[25] Aguilar-Gonzalez P M and Kober V 2012 *Optics Communications* **285** 574-583
[26] Ruchay A and Kober V 2016 *Clustered impulse noise removal from color images with spatially connected rank filtering* **9971** 99712Y-99712Y-10
[27] Ruchay A 2017 *Removal of impulse noise clusters from color images with local order statistics* **10396** 10396-10
[28] Ruchay A 2017 *Impulsive noise removal from color video with morphological filtering* **10396** 10396-9
[29] Ruchay A and Kober V 2018 *Analysis of Images, Social Networks and Texts* (Cham: Springer International Publishing) 280–291
[30] Takimoto R Y, de Sales Guerra Tsuzuki M, Vogelaar R, de Castro Martins T, Sato A K, Iwao Y, Gotoh T and Kagei S 2016 *Mechatronics* **35** 11–22
[31] Kim D, Choi J, Leksut J T and Medioni G 2016 *IEEE International Conference on Image Processing* 3011–3015
[32] Song X, Zheng J, Zhong F and Qin X 2017 *Multimedia Tools and Applications* 1–07
[33] Kim S H, Hwang Y H, Hong H K and Choi M H 2004 *MICAI: Advances in Artificial Intelligence* 642-651
[34] Xin W and Pu J 2010 *International Conference on Computational and Information Sciences* 565-568
[35] Tran M D, Kang H J and Ro Y S 2010 *Improved ICP control algorithm in robot surgery application* 66-69
[36] Chen J and Belaton B 2014 *Machine Learning and Cybernetics* 55-263
[37] Zhang C, Du S, Xue J and Qi X 2014 *Foundations and Practical Applications of Cognitive Systems and Information Processing* 523-530
[38] Guo F, He Y and Guan L 2015 *An improved ICP registration algorithm with a weight-bootstrap scheme*
[39] Ye Q, Yao Y, Gui P and Lin Y 2016 *12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery* 2109-2114
[40] Li S, Wang J, Liang Z and Su L 2016 *Tree point clouds registration using an improved ICP algorithm based on kd-tree* 4545-4548
[41] Yang H, Jiang J, Zhao G and Zhao J 2015 *World Journal of Engineering and Technology* **03** 302-308
[42] Khoshelham K and Elberink S O 2012 *Sensors* **12** 1437-1454

Acknowledgments
This work was supported by the Russian Science Foundation, grant no. 17-76-20045.

Development of fuzzy fractal representation of the image

D G Privezentsev[1], A L Zhiznyakov[1] and E V Pugin[1]

[1]Murom Institute, Vladimir State University, Orlovskaya str. 23, Russia, 602264

Abstract. The article considered the way of the fuzzy image representation with using fractal models of images. According to it, the image presented in the form of a tree range blocks which correspond to image blocks. Everyone range block represents the structure describing similarity of this block to another block of the image. As a result, it turns out treelike fractal representation of the image describing the distribution of property of self-similarity in the image. The basic possibilities of construction of algorithms of digital processing of the images which based on given representation are listed

1. Introduction

Most methods of digital image processing (including defectoscopy images) refer to one of two groups: spatial or frequency. Spatial methods make it possible to identify an object in an image uniquely, and frequency methods allow one to adequately analyse the spectral components of a two-dimensional signal (useful component and noise). In this case, spatial methods, as a rule, are subject to interference and require the source signal to be cleaned of noise, possibly with the loss of useful information. For frequency methods, the main problem is the localisation of objects of interest in the coordinate system of the image.

The use of space-frequency approaches, for example, based on wavelet transforms, allows you to detect objects in images against background noise more accurately. This approach is since the wavelet spectrum contains complete information about the original signal, including the spatial coordinates of the frequency components, and at the same time separates the useful and noise components.

However, the overwhelming majority of digital image processing algorithms based on the algorithms for extracting useful information, which often does not allow the technical vision systems to solve non-trivial problems. In this regard, it is proposed to combine the mathematical theory of fuzzy sets and fuzzy logic with the proven fractal methods of digital image processing.

To develop a system of new features, need to create a new model of a digital image. It is proposed to modify the fractal model developed by the project manager by using a fuzzy distance in it as a measure of similarity of the image areas. This modification allows expanding the hierarchy of representations of the original image, thereby increasing the amount of useful information about the original image.

With increasing complexity of the tasks solved, it is expedient to use the specially oriented methods that to the construction of models, methods and algorithms that take into account the incompleteness and inaccuracy of the initial data. It is in such situations that the theory of fuzzy sets and fuzzy logic proves to be the most constructive. Already achieved theoretical results allow

to consider fuzzy logic as a promising tool for the analysis of heterogeneous information, in the form of complex qualitative linguistic descriptions and quantitative data. In modern works, we considered methods and algorithms for using fuzzy logic and fuzzy set theory to solve problems of digital image processing.

Analysis of the used fuzzy image processing schemes [1, 2, 3] shows that the used fuzzy representations of images, as well as the membership functions used to translate images into fuzzy forms, do not have a single mathematical basis. Used fuzzy signs and properties of images built on the representation of images in the form of sets. In this case, there is no possibility of sharing fuzzy technologies with classical algorithms of digital image processing. Some papers describe fuzzy measures, fuzzy integrals used in segmentation and clustering problems as local statistical features defined on the set to which the image[4]. From this, it follows that for the joint use of classical features, measures and algorithms in conjunction with fuzzy imaging technologies, a system of fuzzy signs, measures and distances needed that allows developing new algorithms for digital image processing.

In connection with the preceding, it is proposed to develop a fuzzy mathematical model of a digital image based on its description by the terms of fuzzy sets using fuzzy logic, which is the basis for constructing new algorithms for extracting information about features and image processing algorithms.

2. Fractal Model of Digital Image

The fractal model of the digital image developed and described in [5, 6, 7] implements the system of iterated functions and can be represented by the expression

$$f' = \sum_{k=1}^{n} \left(B_{n_i,m_i}^{r_i}\right)^* [s_i \left(\tilde{w} \left(B_{k_i,l_i}^{d_i}\right)[f]\right) + o_i]. \tag{1}$$

where

$$D = \left(B_{k_i,l_i}^{d_i}\right)[f] \tag{2}$$

is an operator of extracting a domain block with the index d_i from the original image f. The transformation that performs the change in the color characteristics of the domain block and the geometric transformation has the following form

$$D' = s_i \left(w[D]\right) + o_i. \tag{3}$$

After all the transformations are performed on the original domain block, the result is inserted into the image

$$f' = \sum_{i=1}^{M} \left(B_{n_i,m_i}^{r_i}\right)^* [D']. \tag{4}$$

Using 1, 2 and 3, formulate an expression describing the range block of the image through the domain block

$$R_i = D' = s_i \left(w[D]\right) + o_i. \tag{5}$$

For comparison of two images f, h metric used RMS:

$$g(f,h) = \sum_{i,j} \sqrt{\frac{1}{m \cdot n}[f(i,j) - h(i.j)]^2}, \tag{6}$$

where m,n - images size. For each range block of the image R_i, the domain block D_i is selected which best approximates the range block. In other words, there is a pair (R_i, D_j) for which the metric (6) is minimal.

To generate a fractal image code, we used a quadtree of range blocks. The root of the tree is the image itself.

Expression 5 can be represented as follows

$$R_i = D' + \xi_i, \tag{7}$$

$$\xi_i = g(R_i, D').$$

In other words, the fractal model describes the original image using the quadtree. Each leaf of a tree is a range block of the image, which corresponds to a domain block so that the approximation error is minimal:

3. Intuitionistic Fuzzy Set

A fuzzy set A in a finite set $X = x_1, x_2, , x_n$ may be represented mathematically as

$$A = \{(x, \mu_A(x)) \, | x \in X\}$$

where the function $\mu_A(x) \colon X \to [0, 1]$ is the measure of the degree of belongingness or the membership function of an element x in the finite set X.

An Intuitionistic Fuzzy Set A in a finite set X may be mathematically represented as

$$A = \{(x, \mu_A(x), \nu_a(x)) | x \in X\}$$

where the functions $\mu_A(x), \nu_A(x) \colon X \to [0, 1]$ are, respectively, the membership function and the non-membership function of an element x in a finite set X with the necessary condition

$$0 \leqslant \mu_A(x) + \nu_A(x) \leqslant 1.$$

It is clear that every fuzzy set is a particular case of IFS:

$$A = \{(x, \mu_A(x), 1 - \mu_A(x)) | x \in X\}. \tag{8}$$

4. Development of fuzzy fractal model

To use all the features of fuzzy sets is necessary to modify the model (5) and bring it to the fuzzy mind. During the construction of the fractal image code, a set of range blocks, domain blocks and approximation error values are generated:

$$R = \{R_i\}, R_i = \{D_j, \xi_{i,j}\}. \tag{9}$$

Then, using 8, expression 9 can be expressed using a fuzzy set

$$R = \{D_j, \mu_R(D_j), 1 - \mu_R(D_j) | D_j \in D\}. \tag{10}$$

Then the membership function $\mu_R(D_j)$ may be represented as

$$\mu_R(D_j) = \xi_j / \xi_{max},$$

where ξ_{max} the maximum approximation error, which is defined as the difference between absolutely black and absolutely white images:

$$\xi_{max} = 255 / \sqrt{m \cdot n},$$

where m, n - image size.

Let the image f is a graph $G = (V, \bar{U})$, where V is the list of range blocks of the image, $\bar{U}$ is the connections between the range blocks. Then the process of forming a tree of range blocks of the image can be described as follows. There is a range block of zero level R^0, which corresponds to the whole image and $v_1 = R^0$. It contains range blocks of the first level $R_1^1, R_2^1, R_3^1, R_4^1 \in R^0$:

$$v_2 = R_1^1, v_3 = R_2^1, v_4 = R_3^1, v_5 = R_4^1,$$

$$\bar{u_1} = (v_1, v_2), \bar{u_2} = (v_1, v_3), \bar{u_3} = (v_1, v_4), \bar{u_4} = (v_1, v_5) \in \bar{U}.$$

Continuing the formation of the fractal code,a complete tree of range blocks is obtaining, from which the image is constructed. A fragment of the tree of range blocks is shown in Fig. 1.

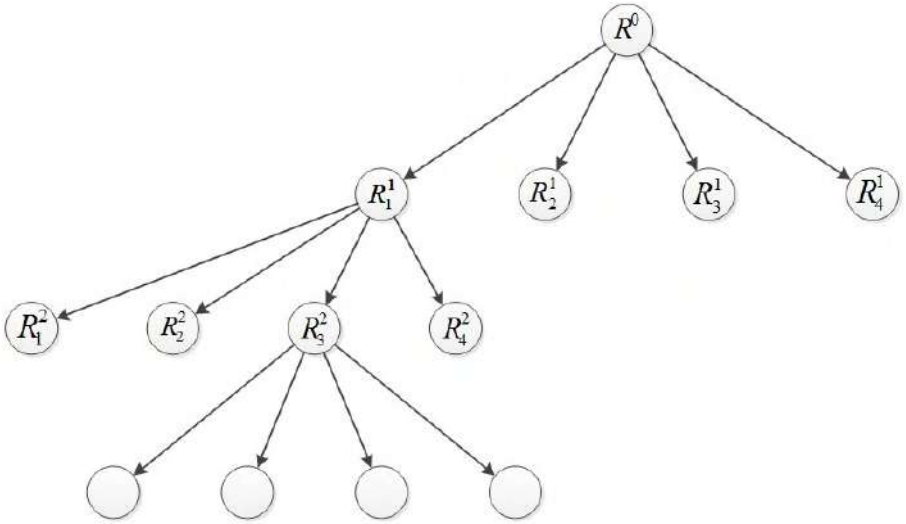


Figure 1. Fragment of the tree of range blocks.

The operator for the formation of a fractal code called a direct fractal transformation:

$$F(f, D) = \Phi, \tag{11}$$

where Φ is the resulting fractal code of the image f by the operator F. An additional parameter of the operator is the list of domain image blocks D involved in the code generation. There is an inverse fractal transformation that forms the image by the fractal code:

$$f' = F^*(\Phi, D) = F^*(F(j, D)). \tag{12}$$

To properly restore an image from its fractal code, it is important that the transformations use the same list of domain blocks D. In the original version, the generated fractal code is full, that means that each vertex of the graph has one parent and four descendants, but the length of the descending chain does not exceed the depth of the quad, which specified at the initial stage of the code construction. If the full tree used as the parameter for restoring the image by the operator 12, then the resulting image f' is similar to the original image.

For obtaining modified images, incomplete trees of range blocks used. To form an incomplete tree, use the operator:

$$\Phi^* = \Gamma(\Phi). \tag{13}$$

The form of the operator Γ depends on the required result. For example, by removing some branches of a fractal tree, you can obtain an image in which the areas for which remote branches respond are smooth, and the remaining areas are clear. Using 10, the threshold operator Γ can be represented as:

$$\Gamma(\Phi) = \{R', D\}, R' = \{R_i|_R(D_j) > \epsilon\}. \quad (14)$$

The operator 16 allows us to use only range blocks whose value of the membership function is greater than a given threshold ϵ.

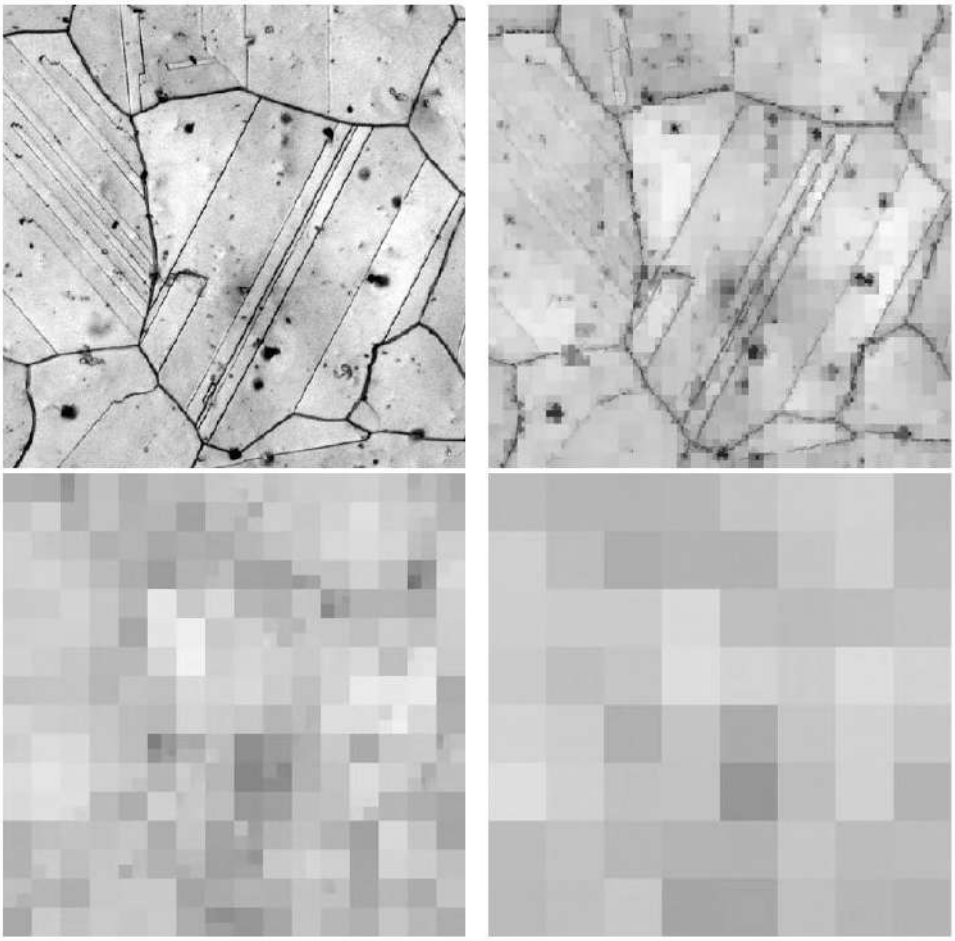


Figure 2. Fractal fuzzy representation of a test image.

To obtain a fractal representation of the image from the fractal code, need to remove the leaves of the tree. The result is a smoothed, fractal image.

Fig. 2 shows an example of such a decomposition of a test image. The resulting set of images allows for a more detailed analysis of the original image. The transition to other scales allows you to move away from small and random details. It is better to reveal the "internal" structure. It is of interest to consider the structural elements allocated at different scales of the representation of the image.

The reason for this approach is the following. Structural elements of the image, such as contours, segments, skeleton, primarily carry all the meaningful information about the original image. At the same time, they have a much more compact representation and are better suited

for analysis, both automatic and subjective, than the original halftone image. In the case of automatic analysis, this is determined by the possibility of a formal description of the structural elements, for example, the parameters of the curve (for the contour, the skeleton) or the area geometric object (for segmentation).

5. Detection of uncharacteristic areas in the image

Most algorithms for digital image processing using fuzzy systems are carried out in three stages: the formation of an image fuzzification, direct processing, and, if necessary, image recovery from an image defuzzification. At the same time, the quality of image processing by such algorithms depends more on the correct choice of the fuzzy representation used [3]. The currently used algorithms for fuzzy representation of images are highly fragmented, i.e. There is no single mathematical device for constructing fuzzy models of digital images that would allow using the mathematical theory of fuzzy sets and fuzzy logic in conjunction with digital image processing methods [4].

An uncharacteristic area in the image is a section that should not be present in a given image, for example, a fracture in the image of the microstructure of the metal, a foreign object on the surface on the aerial photograph, and the like. The task of finding such areas in images is actual in metallography for the search for defects on the surfaces of materials, in radar to search for foreign objects, in medicine for the detection of foreign bodies in the blood and other areas in which image processing is actively applied [5-7].

To detect uncharacteristic blocks in the image, we formed a fractal code using a list of specific blocks instead of domain blocks.

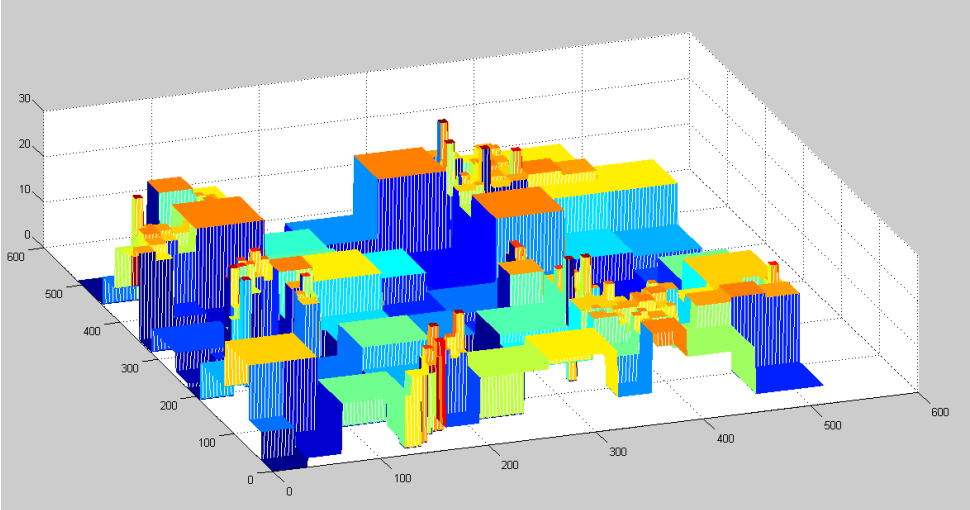


Figure 3. Three-dimensional diagram of errors in the approximation of the test image.

Fractal coding due to the use of approximation implies loss of information, since there is an error in the approximation of rank blocks by domain blocks:

$$R_i = w_i(D_j) + \epsilon_i. \tag{15}$$

In other problems this error can be neglected, because the fractal code of the image is used, without the need for decoding. In this problem, we need to obtain a vector $\{\epsilon\} = \{\epsilon_1, \epsilon_2, ..., \epsilon_k\}, k = 1..N_R$, then we can obtain a two-dimensional function $\Theta(x, y)$:

$$\Theta(x, y) = \mu\{\epsilon_i \mid (x, y) \in R_i\} \tag{16}$$

where $\mu\{\epsilon_i \mid (x, y) \in R_i\}$ - the membership function that determines the amount of fuzzy uncharacteristic of the ranking block R_i in the image.

The result is a three-dimensional diagram showing how well each block of the image approximated by the characteristic areas (Fig. 3). It is possible to establish the presence and location of image sections poorly described using specific areas, by carrying out the threshold processing of the chart

6. Conclusion

The fractal image code represented in the form of a graph G, in which the vertices of V are range blocks R, and the arcs are the nestedness relations of range blocks. At the same time, the structure of the code remains unchanged. This approach changes only the form of representation and the mathematical apparatus of research.

The described approach to the fractal representation of images can be used to form a set of features reflecting the internal structure of the image since small details discarded. Further on, based on the proposed description, algorithms for isolating contours, image segmentation, skeletonisation can be constructed. This support the usability of the proposed algorithm in practical problems of nondestructive testing.

Also, new fractal attributes of digital images offered - specific sites. These signs show which parts of the image are involved in its formation, i.e. according to the self-similarity distribution within the image, it is possible to identify areas using which you can reproduce the image with the highest accuracy. We developed the algorithm for searching for areas of the image that are not characteristic of this class of images. The algorithm based on the fact that uncharacteristic areas of the image poorly approximated by the specific sections of this class of images.

7. References

[1] Mario I and Chacn M 2006 Fuzzy Logic for Image Processing: Definition and Applications of a Fuzzy Image Processing Scheme *Advanced Fuzzy Logic Technologies in Industrial Applications*
[2] Tamalika C 2008 Fuzzy Sets and Their Extensions: Representation, Aggregation and Models *Fuzzy Measures in Image Processing*
[3] Hi Z, Yan H and Pham T 1998 *Fuzzy algorithms: With Applications to Image Processing and Pattern Recognition* (Singapore; New Jersey; London; Hong Kong: Word Scientific) p 225
[4] Bing-Yuan C 2010 Optimal Models and Methods with Fuzzy Quantities Studies in Fuzziness and Soft Computing *Fuzzy Cluster Analysis and Fuzzy Recognition* **248** 117-137
[5] Zhiznyakov A L, Privezentsev D G and Zakharov A A 2015 Using fractal features of digital images for the detection of surface defects *Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications)* **25(1)** 122-131
[6] Privezentsev D G and Zhiznyakov A L 2015 Use of specific image segments in tasks of digital image processing *International Conference Stability and Control Processes in Memory of V.I. Zubov* 659-660
[7] Zhiznyakov A L, Privezentsev D G and Pugin E V 2014 Use of fractal signs of digital images for detection of surface defects *24th International Crimean Conference Microwave and Telecommunication Technology Conference Proceedings* 391-392

Acknowledgments
The reported study was funded by RFBR according to the research project 17-47-330073.

Modeling of Agricultural Spatial Objects with Heterogeneous Dynamically Changing Spatial Structure

R V Brezhnev[1], Yu A Maglinets[1], K V Raevich[1] and G M Tsibulski[1]

[1]FSAEI All-Russia Association of Siberian Federal University, Institute of Space and Information Technologies, Kirenskogo str. 26, Krasnoyarsk, Russia, 660074

Abstract. This article provides research for the models describing the spatial object of Agricultural Land (AL) as observed at medium- and high-spatial resolution satellite images. This object is characterized by variable reflectance features during the vegetation season. These variations are caused both by natural and man-induced environmental effects and by soil cover properties. The vegetation cover growth rates may vary at different areas within the agricultural contour which leads to the creation and development of heterogeneities. The research focuses on the spatial modeling of the agricultural object with heterogeneous dynamically changing spatial structure. A model which allows matching the values of parameters measured at space images with reference and abnormal object behavior has been developed. This model describes the object of research on the level of geometric and relational structures, thereby making it possible to determine spectral and metrical features as part of image heterogeneities, find interrelations between these features within the time limits and map these heterogeneities. The results obtained were tested when monitoring the state and development trends for grain crops at the tested objects in the Sukhobuzimsky district of the Krasnoyarsk Territory.

1. Introduction

The remote monitoring methods are used in agricultural sector to solve a wide range of issue-related tasks, including the assessment of state and utilization of crop rotation fields, soil fertility and soil degradation parameters, state of vegetation cover at plough lands, deposits, hay and grazing lands, forecasting of crop yield and etc. The objects of agricultural monitoring are characterized by variations in time. Variations observed at space images of such objects are related to the natural vegetation stages (change of phenological phases), climatic conditions, and agrotechnical measures (ATM) performed. The type of variations is determined by the properties of soil cover, fertilizers applied and some other factors specified in [1-2]. In some cases, a change in spatial structure of agricultural land is also observed in the form of uneven crop growth within the object contour. Thus, this type of the monitored objects is referred to spatial objects with dynamically variable structure.

General principles for the use of spatial-temporal aspects of dynamically variable spatial objects have been so far put together in the works of B. Claramunt, M. Molenaar, R. Nandal, T. Ott, D. Peuquet, A.Renolen, M. Yuan, Yu. K. Korolev, V.S. Tikunov and etc [3-5]. Event-oriented, 3-domain and History Graph models and etc have been developed. However, these models are aimed at presentation of objects with time-variable boundaries whereas handling of objects with dynamically variable structure is under-developed in the literature.

The object of the present research is a spatial object being vegetable agricultural land with a limited agricultural contour namely the agricultural site [6], further hereon AL or field. This spatial object may be matched with an attribute vector: spectral, textural, metric, topological and etc. During the vegetation season the values of the observed parameters' vector may vary and some attributes may be dynamically added to the formed description as being exhibited at a certain period of time. Furthermore, structural heterogeneity of the object should be taken into account when modeling, moreover the type of heterogeneity is also time variable.

Detection of local heterogeneities of the object and their variation in time at space images allows assessing the uniformity of crop germination capacity, their degradation as well as the performance rate of the planned agrotechnical activities which, in particular, provides for precision farming.

2. Conceptual Model of Agricultural Land Spatial Object

The object of research is characterized by the following types of parameters which are essential for the arrangement of a monitoring system [7]:

1. Spatial confinement (exact geographical references of the object with due account of topological parameters).
2. Name and breed of crop for the current field season.
3. Seasonal vegetation development schedule in the form of changing crop phenological stages.
4. Possible abnormalities (in connection with phenological stages) such as non-germination, soil overwatering, drought, impurities and etc.
5. Man-induced impact as a result of planned and performed agrotechnical activities.
6. Crop characteristics (such as integral state, biomass, leaf surface and etc.).

The abovementioned parameters may be detailed for the specific natural and climatic states and type of vegetation. Thus, Figure 1 shows a set of allowable states of cereal crop, in particular spring wheat, barley and oat correlated with the time intervals (*Tm*) as determined for crop rotation in central regions of the Krasnoyarsk Territory.

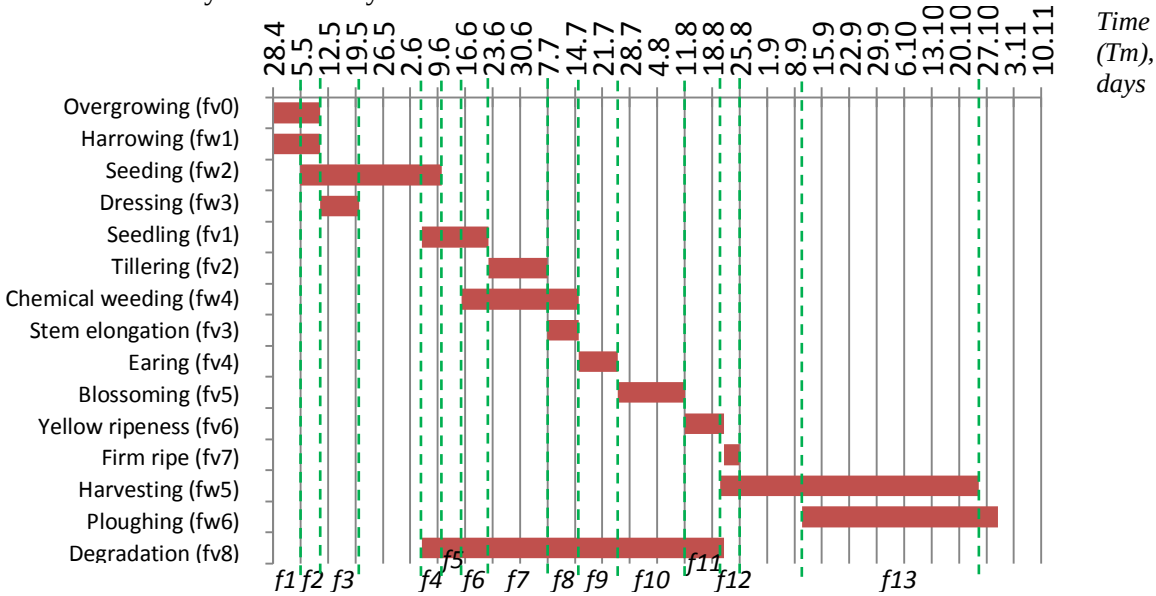


Figure 1. Trends for the agricultural land by the example of spring wheat.

With account of the performed agrotechnical activities the object may be in one of the following f^{wi} states: f^{w1} – harrowing, f^{w2} – seeding, f^{w3} – dressing, f^{w4} – chemical weeding, f^{w5} – harvesting, f^{w6} – ploughing. The crop may also be in certain f^{vi} states which are determined by phenological stages of a certain time sequence. Thus, stages for the cereal crops researched herein describe the following states of the object: f^{v0} – overgrowing, f^{v1} – germination, f^{v2} – tillering, f^{v3} – stem elongation, f^{v4} – earing, f^{v5} – blossoming, f^{v6} – yellow ripeness, f^{v7} – firm ripe stage. A cereal crop may pass to f^{v8} state that is degradation at each stage of vegetation which indicates the

abnormal crop development. When in standard state, the object is presented either by open soil or homogeneous vegetable carpet corresponding to a certain development stage.

Interception of standard time ranges for the object being in one or another state results in heterogeneous structure of the object and requires consideration of the object's composite states which can be analyzed on State-Transition Diagram *G* (Figure 2):

$$G = (F, R), \tag{1}$$

where *F* is a variety of the object's states:

$$F = f^{vi} \cup f^{wj}, \tag{2}$$

R - total state transitions:

$$R = r^1 \cup r^2 \cup r^3, \tag{3}$$

where r^1 is a variety of operators describing a crop development in time (natural change of phenological phases), r^2 is a variety of operators describing abnormal crop development, r^3 is a variety of operators modeling the performance of agrotechnical activities.

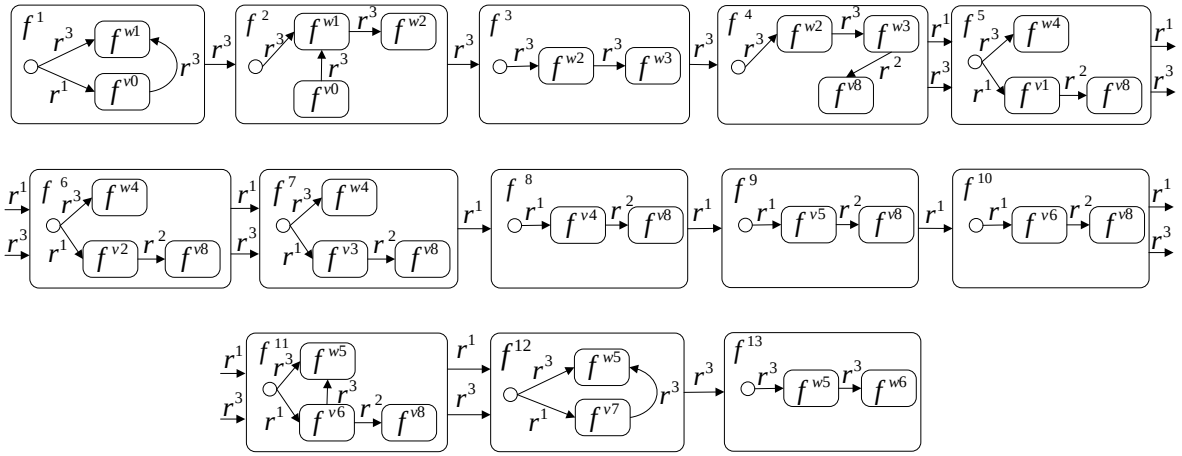
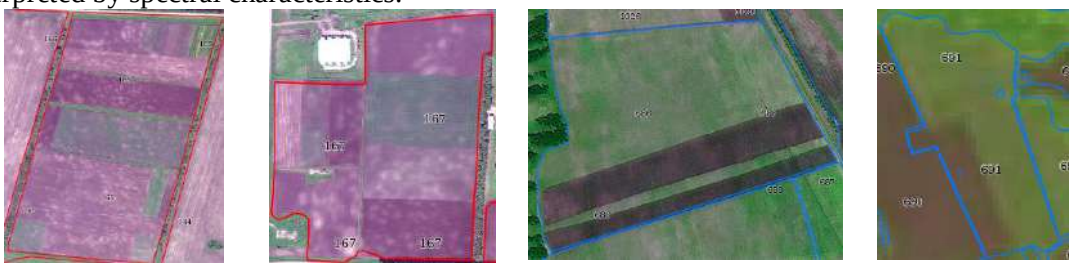


Figure 2. Dynamic model of agricultural land spatial object.

3. Model of a Spatial Object with Heterogeneous Dynamically Variable Structure

The changes in the object's structure have been traced and interpreted with the help of Earth Remote Sensing (ERS). Satellite images (Figure 3) show heterogeneities as a plurality of segments which may be interpreted by spectral characteristics.



a b c d

Figure 3. Heterogeneities within the contours: a, b – Nos 165, 167 as per the high-resolution data WorldView-2 (0.5 m); c – No 686 as per the high-resolution data Spot-6 (6 m); d – No 691 as per the medium-resolution data Landsat-8 (30 m).

Each state is characterized by a set of features which are essential for setting the monitoring task by an end user as well as for allocation and interpretation of states of the object's heterogeneities. The main categories of such features are given below.

1. *Metrical features* including:
   1.1. Geographical coordinates of the object in the set projection $P_{\Pi} = \{(x_i, y_i)\}$, specifying its location. Surface area of polygonal-shaped object calculated as follows:

$$NS = \frac{1}{2}\left|\sum_{i=1}^{n-1} x_i y_{i+1} + x_n y_1 - \sum_{i=1}^{n-1} x_{i+1} y_i - x_1 y_n\right|, \quad (4)$$

where $(x_i, y_i), i = 1, 2, \dots, n$ are a sequence of coordinates of the neighboring vertexes of polygon, $(x_{i+1}, y_{i+1}) = (x_1, y_1)$.

Perimeter of the object:

$$NP = \sum_{i=1}^{n} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}, \quad (5)$$

where $i = 1, 2, \dots, n$ is a number of the contour vertexes.

Thickness of the object:

$$T = 4\pi(NS/NP^2). \quad (6)$$

2. *Spectral features* describe the object in various spectral ranges.

Normalized Difference Vegetation Index (NDVI) has shown good results in Agricultural Land structure classification [8]:

$$NDVI = (NIR - RED)/(NIR + RED), \quad (7)$$

where NIR is a value of spectral channel corresponding to the near IR range with the wavelength of 0.76–1 μm; *RED* is a value of spectral channel corresponding to the visible-light spectrum with the wavelength of 0.6–0.76 μm (red light). When modeling, this Index was compared to other vegetation indexes such as DVI (Richardson and Everitt, 1992), RVI (Jordan, 1969), WDVI (Clevers, 1988), SAVI (Huete, 1988) and etc. with the help of experimental evidence. The above spectral indexes as well as the eigenvalues of NIR and RED spectral channels are informative for vegetation analysis. The developed model is invariant with respect to the choice of a specific index.

3. *Time features* (Tm) specify discrete variations of the object's states.

The image of an object shall be transferred from the iconic level to the level of geometrical and relational structures in order to solve the tasks of heterogeneities' allocation at satellite images and their presentation in the form of heterogeneities' maps as well as to provide information support to the end user. This shall allow calculation of spectral and metrical features of heterogeneities and identification of interrelations between these features and elements of *F* plurality within the set time ranges.

The resulting model of an object with a heterogeneous dynamically changing structure can be presented at the level of geometrical and relational structures as follows:

$$Agricultural\ Land =< ID, P_{\Pi}, NS, NP, T, N, Tm, K, E, \varphi, F >, \quad (8)$$

where *ID* is an identifier (field number), P_{II} is a plurality of the object's coordinates: $P_{II} = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, where x, y are latitude/longitude point coordinates accordingly, *NS* – surface area, *NP* – perimeter, *T* – thickness, *N* – plurality of *NDVI* value ranges corresponding to the specific assessment of the object state, *Tm* – plurality of time ranges for the object state transitions.

In addition to the above features, various semantic data related to the object and specifying the peculiarities of the object development in time must be borne in mind for the object allocation by the end user when setting the monitoring tasks. It includes: set agricultural crop $K = \{k^i\}$ with its own reference model of development in time which is determined, on one hand, by agrotechnical activities identified as a plurality of time-scheduled events $E = \{e^i\}$ and influencing transition of the object state and, on the other hand, by successive phenological development stages $\varphi = \{\varphi^i\}$ identifying the plurality of the object states *F* (2).

The expression (8) describes Agricultural Land object at the level of geometrical and relational structures and allows identification of spectral and metrical features within the localized contours of heterogeneities, setting interrelation of these features in preset time ranges as well as presentation of findings in the form of heterogeneities' map.

4. Experimental Approval

Free satellite images Landsat-8 and Sentinel-2A (VNIR shooting mode) were selected as input data for the experimental localization of heterogeneities. Spectral channels in red (RED) and near IR (NIR)

ranges with wavelengths of 0.63–0.68 and 0.78–0.89 μm accordingly were used. Fourteen scenes were obtained in 2016 and ten scenes in 2017.

The test area within Minderlinskoe educational-research farm in Sukhobuzimsky district of Krasnoyarsk Territory (Figure 4) was chosen for experimental research of the spatial object heterogeneous structure. The chosen area is located in the central part of the Krasnoyarsk Territory. It stretches for 140 km from West to East and for almost 100 km from North to South. The total surface area is 561,260 ha (5,613 km²). Forest-steppe landscape in the western part grades into taiga area on the right bank of the Yenisei River. The survey area the agricultural fields are located at have the following references: between 92.5°E and 94.7°E and between 56.2°N and 56.9°N.

The routine of experiment includes atmospheric correction of RED and NIR channels by DOS method in order to obtain correct values when calculating Normalized Difference Vegetation Index (7); the threshold segmentation by the obtained NDVI channel; correction of the obtained segments to the level of geometry structures into the format of SHP ESRI vector layer and layer importing into PostgreSQL DBMS for relational presentation of the features of heterogeneities; calculation of model characteristics (4-6) and averaged NDVI value for each segment.

Allocation of heterogeneous areas of Agricultural Land and interpretation of their states shall be considered as the task of the satellite image segmentation by the spectral NDVI feature in accordance with the homogeneity criteria which are true for the test objects under consideration with cereal crops at the researched area and which are defined as follows:

$$s(x,y) = \begin{cases} \lambda_0 = \text{background} & with -1 \le f(x,y) < 0{,}025, \\ \lambda_1 = f^{w1}\ or\ f^{w6} & with\ 0{,}025 \le f(x,y) \le 0{,}26, \\ \lambda_2 = f^{w2}\ or\ f^{w5} & with\ 0{,}26 < f(x,y) \le 0{,}43, \\ \lambda_3 = f^{v1}\ or\ f^{v6} or\ f^{v7} & with\ 0{,}43 < f(x,y) \le 0{,}57, \\ \lambda_4 = f^{v2} & with\ 0{,}57 < f(x,y) \le 0{,}65, \\ \lambda_5 = f^{v3} & with\ 0{,}65 < f(x,y) \le 0{,}72, \\ \lambda_6 = f^{v4}\ or\ f^{v5} & with\ 0{,}072 < f(x,y) \le 0{,}85 \end{cases}, \quad (9)$$

where $s(x,y)$ is a segmented image, $\lambda_0, \lambda_1, \dots, \lambda_{m-1}$ are marks of segments corresponding to the preset state, $f(x,y)$ is a function of the image brightness.

The threshold selection procedure during segmentation is based on analysis of statistical data collected starting from 2013 by field NDVI measurements and calculation of correlation dependence with spectral vegetation features as per Landsat-7,8, Sentinel-2A, WorldView-2, Spot-4,6 satellite data.

Therefore, the on-site field investigations of test fields were performed to verify the satellite measurements. The base map of test fields is shown in Figure 4. Four test fields were selected at the test range. These fields are cropped with one culture: with oat (176), barley (155) and wheat (157, 166) crops.

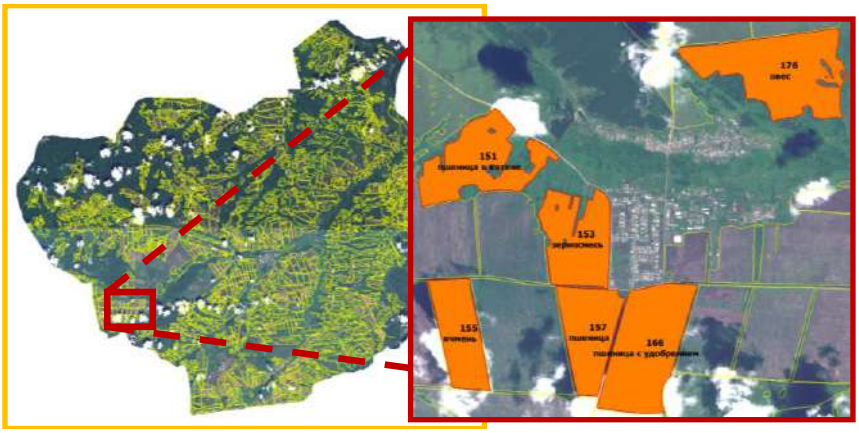


Figure 4. Base map of test fields.

The FI-20A photometrical meter was used for NDVI measurement in order to eliminate atmospheric effects such as clouds or haze. This meter is meant for the day-time operation at the ambient air temperature from 5 to 40°C and relative air humidity up to 80% with no precipitation.

The resultant model allows revealing heterogeneities corresponding to the preset NS dimensions and T form at the geometrical level. This can be used in precision farming when analyzing spot tillage and calculating the required amount of crop protectors and etc. Therewith only segments which state does not correspond to the current phenological stage are taken into consideration. This is done by filtration of localized segments with due account of their possible pertinence τ_F to some actual state $f_A^i \in F$ and of comparison f_A^i with the reference state $f_{\text{н}}^i$ typical for that time.

Determination of state F for heterogeneous area w_i is the dependence function of time Tm and of the average vegetation index $\overline{N}$: $F_{w_i}(Tm,\ \overline{N})$.

The segmented image is worked out to the level of object-relational structure where each area w_i *j-го* of Agricultural Land object with dimensions NS_j^i and form T_j^i which satisfy the given values is space-outlined by a plurality of coordinates $\{P_{\Pi j}^i\}$, is fixed by the unique identifier *ID* specifying the "whole-part" relation with the object itself and shows the actual state f_{Aj}^i. This actual state may or may not correspond to the reference state $f_{\text{н}}^i$ depending on the averaged vegetation index $\overline{N}_j^i$ and time Tm. The reference state is determined by the actual phenological stage φ_j^i or by the agrotechnical activity e_j^i specific for the set crop e_j^i, this being presented in the object-relational model of Agricultural Land heterogeneity:

$$3CXH_j =< ID, \{P_{\Pi j}^i\}, NS_j^i, NP_j^i, T_j^i, \overline{N}_j^i, Tm, k_j^i, e_j^i, \varphi_j^i, f_{Aj}^i >. \quad (10)$$

Segmentation results as per Sentinel-2A data dd. 05.06.2016 are shown in Figure 5-a, b and as per Landsat-8 data dd.15.06.2016 are shown in Figure 5-c, d.

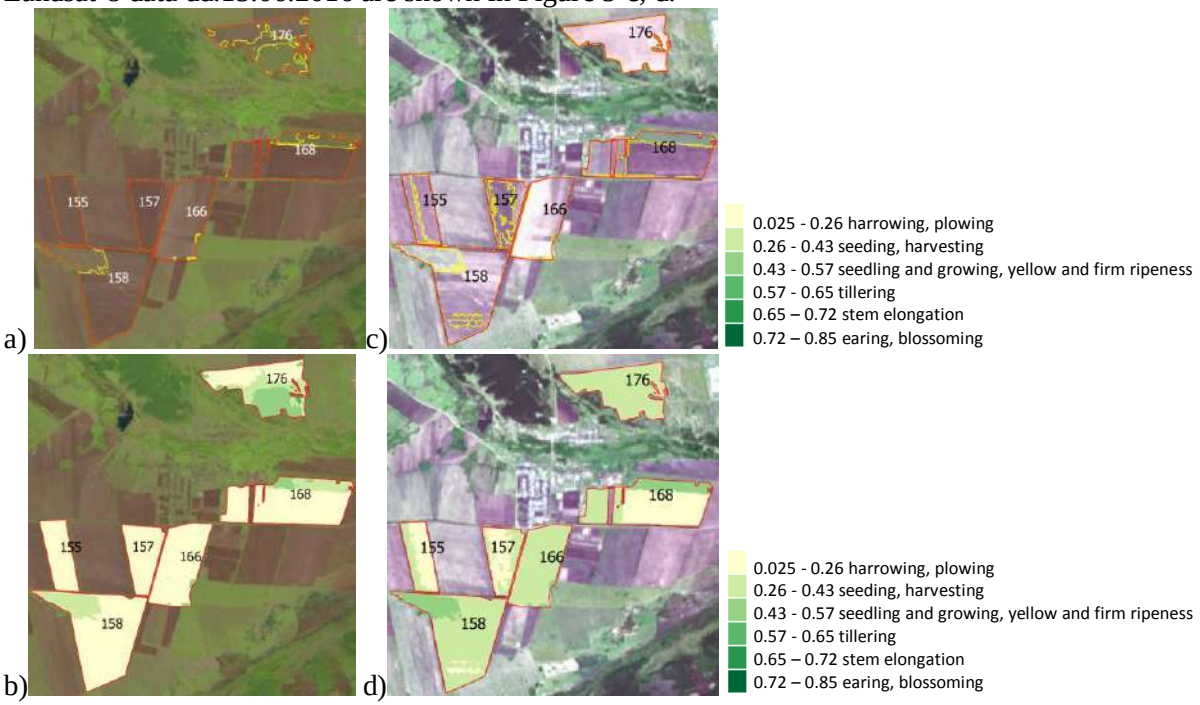


Figure 5. Results of heterogeneities' allocation: a, b – Sentinel-2A dd. 05.06.2016; c, d – Landsat-8 dd. 15.06.2016.

The results received on June 05, 2016, showed that the area of the objects under research was mainly presented by open soil. The situation changed in 10 days that is on June 15, 2016, with the volume of biomass increasing in accordance with the current phenological stage of the cereal crop

development (the seedling stage) and the beginning of the tillering stage which is well shown by a change of the local objects' structure.

Correlation analysis of the spectral survey showed high (0.87-0.99) correlation with the results of measurements in the analyzed satellite images. Twenty to thirty measurements were taken to calculate the average NDVI value for each field. The space coordinates of the measurement points were recorded by GPS-receiver.

5. Conclusion

This article deals with the resultant model of Agricultural Land spatial object with dynamically changing spatial structure. This model allows:

- due consideration of heterogeneous spatial dynamically changing structure of the object;
- calculation of spectral and metrical features of heterogeneities and correlation between these features within the set time range in order to compare the reference f_{H}^{i} and actual f_{A}^{i} states of each heterogeneous area of the object for information support of end-users as regards agricultural monitoring and precision farming;
- provision of allocation results in the form of a heterogeneities' map.

The model is based on software and is built into the hardware-and-software complex of web-GIS for agricultural monitoring which belongs to Siberian Federal University [9, 10] and is available at the following Internet address: http://activemap.ikit.sfu-kras.ru. It co-functions with the software complex in offline mode and allows information support of end-users as regards evaluation of the integral state of Agricultural Lands.

6. References

[1] Raevich K V, Maglinets Yu A and Tsibulskii G M 2016 Intellectual Information System for evaluation of agricultural lands *J. of Sib. Fed. Univ. Engin. & Techn.* **9** 1025-1034
[2] Maglinets Yu A, Raevich K V and Tsibulskii G M 2017 Knowledge-based geoinformation technology for evaluation of agricultural lands *Procedia Engineering* **201** 331-340
[3] Pelekis N 2004 Literature review of spatio-temporal database models *J. The Knowledge Engineering Review* **19(3)** 235-274
[4] Jensen C S 2001 Introduction to Temporal Database Research *CSCI Seminar* **599**
[5] Nandal R 2013 Spatio-Temporal Database and Its Models: A Review *J. of Computer Engineering* 91-100
[6] Dibihin K Yu 2009 Recognition of the elements of the underlying surface of agricultural landfill *Multidisciplinary University as a Regional Center of Education and Science* 587-591
[7] Brezhnev R V and Maglinets Yu A 2017 Information Support Technique for Solving Agricultural Land Monitoring Tasks Based on Earth Remote Sensing Data *J. of Sib. Fed. Univ. Engin. & Techn.* **10(6)** 819-827
[8] Rouse J W 1973 Monitoring vegetation systems in the Great Plains with ERTS *ERTS Symposium NASA* **SP-351(I)** 309-317
[9] Brezhnev R V, Maglinets Yu A, Maltsev E A, Perfilev S E, Sidorov A Yu, Tsibulskii G M and Shokol A S 2012 Software-and-technological infrastructure of information support for solving territorial administration tasks *J. of Sib. Fed. Univ. Engin. & Techn.* **5(3)** 340-352
[10] Maglinets Yu A, Tsibulskii G M and Noskov M V 2016 Modern trends to build up the regional ERS systems *J. of Sib. Fed. Univ. Engin. & Techn.* **9(7)** 1012-1018

Context-based method for lossless compression of RGB and multispectral images

A V Borusyak[1], P A Pakhomov[1], D Yu Vasin[1] and V E Turlapov[1]

[1]Lobachevsky State University of Nizhni Novgorod, Prospekt Gagarina 23, Nizhni Novgorod, Russia, 603952

Abstract. We consider the problem of compression of RGB and multispectral images by context-based methods. The algorithm' logic allows for its examination by using the example of full-color images as a particular case of multispectral images. The image-forming channels are divided into two groups: main and additional (detecting) channels. A distinguishing feature of the main channels is a significant correlation between neighbors. A number of variants of prediction from the adjacent channel for the main and additional channels for lossless image compression were considered. In the experiment on a series of images of different contents, the proposed algorithm showed a superior compression ratio in comparison with the popular WinRar, 7z, PNG archivers for all prediction variants. The leader among popular compression methods, JPEG-LS, was surpassed in the record configuration 2b on the image from the Landsat series by 40%. We expect to continue research on a wider sample of images and to use this algorithm to compress multispectral images with a greater number of channels.

1. Introduction

Context-based modeling is an important step in high-performance lossless data compression. Serious advantages offered by high compression degree enable prediction based on the model of matching (coincidence) of contexts. These advantages were successfully demonstrated using the Prediction by Partial Matching (PPM) method published in 1984 [1] to solve the task of text compression. In 2005, the PPM method was significantly improved in [2] by mixing several contexts with weights that change during the method execution (based on machine learning methods). In this method, the model of each context independently estimates the probability and confidence that the next data bit will be 0 or 1. The forecasts are further weighed (the sum of the weights is 1), the weights are corrected by the prediction success criterion. Open source code (www.mattmahoney.net/dc) software (PAQ8) is implemented in the method. This software demonstrated a high rating in several independent tests.

The wide use of this approach to images compression began around the early 2000s, but it was used primarily for binary images containing mostly text and lines [3]. It offered a compression improved by 14% compared with the analog and a 25% better performance. In 2002, a parallel algorithm for this method was developed [4]. In 2001, the context-based method was applied in the development of a new video coding standard for entropy estimation in the coding procedure using binary adaptive arithmetic coding technology, which increased the coding rate by 35% [5].

Of essential importance in the application of the context-based approach is the difference between images and texts consisting in the presence of noise in the images. Therefore, the result obtained in [6] in 2008 was very important for image compression: the prediction method for images with the help of

Prediction by Partial Approximate Matching (PPAM) was presented. Unlike the PPM modeling method that uses exact contexts, PPAM introduces the notion of approximate contexts. Thus, PPAM models the probability of encoding a symbol based on its previous contexts, and contextual occurrences as a result are considered in an approximate manner. The method demonstrated competitive lossless compression and good performance when compressing images that have repeating areas with similar characteristics.

However, the use of the context method for compressing color and multichannel images has remained complicated and ambiguous for a very significant reason: the effective definition and use of contexts for such images is a complex task, since in essence it is the compression of three or more images simultaneously. Nevertheless, in 2011, the publication [7] explored the prospects of using PAQ family methods in combination with machine learning (ML) methods for simple color images and for lossy compression. A number of problems were identified: 1) PAQ can be applied only to one-dimensional sequences, and the expansion for several sequences is not trivial (even in the case of identification of chicken carcass parts); 2) the authors were unable to construct parametric models of typical image contexts, which required for PAQ methods a huge storage capacity. In all four test images used to compare the PAQ-ML method with JPEG and JPEG2000, it was superior to JPEG2000 both in terms of the compression ratio and the quality of the compressed image. The method proposed by the authors showed a significant change in color amounting to the distortion of the palette, while JPEG2000 maintained the ratio of color channels in the local context, was able to locally parameterize the change of this ratio and thus proved to be the winner.

It is also of interest to study the possibilities of using context-based compression methods for color (RGB) and multispectral images of Earth remote sensing (ERS). In the general case, ERS images are multi-channel, i.e. each pixel in the image is specified by the channel value vector. The early compression algorithms included, as a rule, independent operations on individual sample matrices, which were the matrices of the original image channels, or one, two or all the three RGB channels assigned to represent them. Therefore, the publications at that time primarily considered algorithms for processing single-channel (halftone) images, which are basic for implementing all compression methods. More recent publications are related to the compression of hyperspectral images, where the hierarchical compression method for both hyperspectral images (HSI) and for ERS as a whole occupies one of the leading positions [8], [9], [10]. In [9], the following statistical characteristics of the HSI are given:

• the difference between the maximum and minimum brightness gradations reaches thousands and tens of thousands times; such images cannot be converted to "byte images";

• components are very dependent; intercomponent correlation is extremely high (above 0.95 for 85.2% of the pairs of neighboring components);

• most components have high intracomponent correlation (above 0.85 for 87.4% of all components).

In what follows, we will be guided by these considerations.

The hierarchical multiscale representation is based on the results of a number of previous studies. It serves to solve not only the problem of ERS data compression, but also several other problems at the same time, such as the problem of compact storage and high-performance adaptive (in terms of permissible losses and the observer's position) visualization of the terrain surfaces with a controlled value of distortions [11]. In this paper, we use a multi-scale wavelet representation of elevation data and a JPEG2000 encoder to compress 8-bit quantized height differences between their predicted and accurate values. This approach can be used without any significant changes to compress any channel (including the reference channel) of multi- and hyperspectral images.

The high correlation of most of the neighboring HSI channels allows us to apply context-based compression methods at a new level and to use the previous high-correlation channel as the context for the current channel, which has been successfully realized and investigated in publications [9], [10]. The high correlation of the HSI channels has made it possible to bring the level of their lossless compression to the values of the order of 4-5.

Unfortunately, it has not been possible so far to achieve this level for the compression of multispectral data, because the high correlation of neighbors is not a rule for such type of data. When

considering the problem of compressing multispectral data, we will assume the channels of multispectral images to be unequal in terms of their information role in the summary image. One of the channels will be taken as the main (reference) channel, while the others will be used as: 1) special contrast channels for detecting objects of interest; 2) complementary channels, highly correlated with the reference channel (if any). For example, in the RGB image of an oasis in the desert, the yellow sand will be determined by almost identical maps of the red (reference) and green (complementary) channels, and the blue water will be determined by the contrast blue (water-detecting) channel. A similar situation would arise if we were to shift the infrared and ultraviolet (detection) channels to the visible region around the green (reference) channel. Obviously, for evolutionary reasons, the red or green channels are more acceptable to us as support channels. We will take one of them as a reference channel in a "conditionally RGB" image to be compressed, which is quite close to the method of "common reference" spectral components for compressing hyperspectral images [9]. The context-based method for compressing RGB and multispectral images proposed below is the development of an algorithm for adaptive compression of indexed and color images with the use of context modeling [12-14].

2. The algorithm for lossless compression of RGB images

Compression of each pixel is performed channel-by-channel. First, the color component responsible for the red color is compressed, next, the value of the color component of the green color is encoded, and then the value of the blue component is encoded. For each channel, its own context is formed. The structure for context storage is identical to the structure of the algorithm for indexed images [13,14]. For each channel, individual context models of the following orders are used: 6,4,2,1,0. The full-color probability coder (FPC) compression algorithm [14] is as follows: 3 separate keys of the current context and 3 independent forests of AVL-trees are used to store context models (for each of the RGB channels). The following actions are performed in the cycle:

- A consecutive pixel is extracted from the input image file as a current one;
- The maximum-order context (MOC) is formed as the current context: the contexts Cont1, Cont2, Cont3 of the maximum order are formed sequentially for the red, green and blue channels, respectively, as an array of unsigned one-byte integers storing the previous values of the corresponding channel of the current pixel.
- The procedure of channel-by-channel coding of the current pixel value in the current context is performed (for details, see [13]);
- If it is not possible to evaluate and encode the current channel value in the current context, since this value is encountered in the current context for the first time, a lesser order context is formed and this context becomes current, thus a return to point 3 occurs. This continues until the current value of the color component is encoded, which is guaranteed by the fact that occurrence counters for all pixel values in the context of the smallest (zero) order are initially assigned the value of unity. The descent to a lesser order context is realized by applying the exclusion technique, which allows, in case of departure to the contexts of a smaller order m, to exclude from consideration all the values of the pixel occurrence counters that are contained in the context model of the order r, $0<m<r<R$, since none of them is an encoded pixel value.
- The context model (as the MOC) is updated in accordance with the code of the current pixel.
- If not all the pixels are coded, the transition to step 1 occurs, otherwise the encoding is completed.

3. Using the link between channels in context-based image compression

To increase the compression ratio, images were studied to determine the existence and possibility of using the relationship of color channel values.

Experiments have been performed to compare compression of full-color images in three modes:

- Independent channel coding: only the R channel values are used to form the R channel context; only the G channel values, for the context of the G channel; only the B channel values, for the context of the B channel.

- Partial relationship between the channels: for the context of the R channel, only the values of the R channel are used; for the G channel, the values of the R and G channels; for the context of the B channel, the values of all the channels: R, G and B.
- Complete relationship between the channels: for the contexts of each of the channels R, G, B, the values of all the channels R, G, B are used.

In full-color images, the color component values of neighboring pixels often have close or similar values. In most cases, there is a gradual color change from one pixel to another. It is known that for most ERS images, the histograms of the difference between the channel values of adjacent pixels in the line-by-line readout is a normal distribution with zero mean and small sigma. This allows us to use, instead of channel values, the difference between its values for the current and previous pixels. The same increments can be used if necessary to form the contexts: the first value in each line of the context is considered an increment of 0. Since increments can also be negative, it is advisable to apply a linear transformation. As a result, the calculation of increments occurs according to the formula:

$$rn_{i,j}^{q} = (g + (\alpha_{i,j}^{q} - \alpha_{i-1,j}^{q}))\%g,$$

where $rn_{i,j}^{q}$ is the new normalized value of the difference between the current $\alpha_{i,j}^{q}$ and the previous $\alpha_{i-1,j}^{q}$ values of the current color channel, q is the color channel number, g is the number of channel gradations (in full-color images it is 256), % is the modulo operation. The difference $rn_{i,j}^{q}$ is only used as the value of the current element being encoded; for the context elements, the initial values of the color components are used.

The FPC algorithm for full-color image compression is described in detail in [14].

In this algorithm, it is possible to use instead of the channel value the increments between its values for the current and previous pixels Experiments have been performed to compare compression of full-color images in three modes:

a) without using increments between neighboring pixels
b) using increments only for pixel values, and channel eigenvalues for the context.
c) using increments for both the pixel value and the context elements.

4. Results of experiments

To determine the most effective way of context formation, experiments were conducted to compress images of different size and contents, as shown in Tables 1-2. Table 3 contains thumbnail images used for experimental tests.

Table 1 shows the source file sizes and compression ratios for widely used algorithms, such as WinRar, 7z, PNG, JPEG-LS. The JPEG-LS algorithm is chosen for comparison, since it is closer to the proposed algorithm in terms of the compression method used and has a greater compression ratio for most images compared to JPEG2000 in the lossless compression mode. The compression ratio of the entire set of files is taken as the sum of the volume of all source files divided by the sum of the volumes of compressed files. In terms of the compression of the sum of files, the leader is JPEG-LS with the result of 2.18, while for the compression of the landsat.bmp file from the Landsat series, the best performance was shown by the 7z archiver with a result of 2.51 versus 1.98 for JPEG-LS.

Table 2 presents the comparative results separately for cases 1-3 described in Section 3 for the variants a, b and c. As a result of the experiments, it was established that approach No.2b with a partial relationship between RGB channels and using increments for pixel values only, and channel values proper for the context, turned out to be the most effective one. Approach No.3 showed a reduced compression ratio in comparison with approach No.1, Approach No.2 on the average was slightly more successful than No.1. Variant 2b is 3.5% better than 1a and 1b, more than 2% better than JPEG-LS for the whole set of files, and more than 40% better for compression of images from the Landsat series (and also 12% better for this image than 7z). It should be noted that the high compression ratio is achieved due to the greater use of computing resources (270 seconds for the compression of the landsat.bmp file compared to 55 seconds for WinRar, 100 seconds for 7z, 15 seconds for PNG, 5 seconds for JPEG-LS). We have established that the approach using the channel difference as the

values is especially effective in predicting when the correlation between the channels is taken into account.

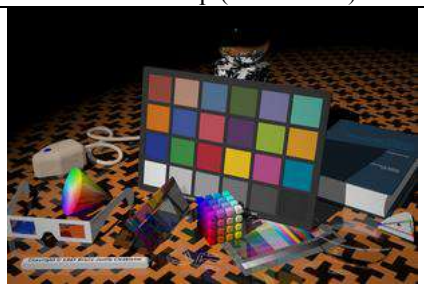
Table 1. The sizes of test images and compression ratios for commonly used compression algorithms.

		Compression ratio			
File \ algorithm	Size (bytes)	WinRar	7z	PNG	JPEG-LS
30057A.bmp	186786486	1,84	1,72	1,37	2,21
artificial.bmp	18874422	10,64	16,54	11,28	12,07
big_tree.bmp	83101254	1,57	1,35	1,15	1,84
energy_bliss.bmp	3932214	4,50	2,56	1,97	4,72
landsat.bmp	46129734	2,10	2,51	1,93	1,98
The entire set of files	338824110	1,89	1,77	1,43	2,18

Table 2. Compression ratios for the variants 1a-3c of the proposed algorithm.

file \ algorithm	1a	1b	1c	2a	2b	2c	3a	3b	3c
30057A.bmp	2,04	2,04	2,03	2,11	2,16	2,12	1,96	2,10	2,03
artificial.bmp	13,38	13,40	12,65	17,39	19,08	17,65	16,97	19,23	17,18
big_tree.bmp	1,83	1,83	1,75	1,63	1,76	1,70	1,46	1,72	1,66
energy_bliss.bmp	2,44	2,46	2,62	2,85	3,98	4,29	2,74	3,72	4,15
landsat.bmp	2,82	2,81	2,39	2,76	2,82	2,42	2,69	2,71	2,32
The entire set of files	2,16	2,16	2,10	2,14	2,23	2,14	1,97	2,16	2,07

Table 3. Thumbnails of images being compressed.

30057A.bmp (8063x7721)	artificial.bmp (3072x2048)	big_tree.bmp (6088x4550)
		
energy_bliss.bmp (1280x1024)	landsat.bmp (3792x4055)	
		

5. A detailed description of the approach used

Let us consider in detail variant 2b. The context formation is clearly illustrated in Figure 1. The colors correspond to the channels R - red, G - green, B - blue.

Only the R values from the previously processed pixels are used as the context $Cont_1$ for the red channel value (R). For the context of the green channel $Cont_2$, in addition to the value of the green color component (G) of the previously processed pixels, the R values from the previously processed and current pixels are also used. For the context values of the blue channel $Cont_3$, in addition to the values of the blue color (B) of the previously processed pixels, the values R and G of the previously processed and current pixels are also used.

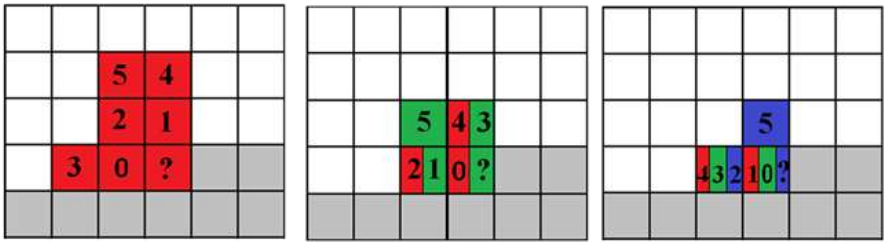


Figure 1. Contexts of the R, G, B channels for the full color coding algorithm (FPC).

To increase the compression ratio, the increment between the corresponding channels for the current and previous pixels is used instead of R, G and B of the current pixel. With this approach, it is possible to use efficiently the correlation between color channels. In the software implementation, each of the three contexts, $Cont_1$, $Cont_2$, $Cont_3$, is a class that includes the active context length and an array of N unsigned 1-byte integers that store the previous values of the corresponding channel for the current pixel, where N is the number of pixels in the context. The contexts $Cont_1$, $Cont_2$, $Cont_3$ are formed according to the formula (1).

$$Cont_{\mu}(rn_{i,j}^{q}) = \begin{cases} i_0 = i - t_k + b_k \\ j_0 = j - k \\ t_k = p - k; \\ k = 0,1,\ldots p_{\mu}; p_{\mu=1} = 2; p_{\mu=2,3} = 1 \\ b_0^{\mu=1} = 0,1,2,\ldots,t_0 - 1; b_1^{\mu=1} = 0,1,2,\ldots 2t_1; b_2^{\mu=1} = 0; \\ b_0^{\mu=2} = -1,0,0; b_1^{\mu=2} = 0,0,1; \\ b_0^{\mu=3} = -1,-1,0,0,0; b_1^{\mu=3} = 0; \end{cases} \quad (1)$$

$$i = 0,1,2\ldots m-1;\ j = 0,1,2\ldots n-1$$

$$q_{\mu=1} = 1; q_{\mu=2} = 1,2,1,2,1,2; q_{\mu=3} = 1,2,3,1,2,3$$

$$\mu = \{1,2,3\}$$

where *m, n are* the width and height of the compressed image, i_0, j_0 are the coordinates of the individual context pixels, $rn_{i,j}^{q}$ is the current encoded element, p is the height of the context in pixels, q is the channel number (1-red, 2-green, 3-blue), μ is the number of the channel to be coded.

6. Conclusion

The problem of compression of RGB and multispectral images by context-based methods has been considered. The logic of the algorithm has been examined using the example of full-color images as a particular case of multispectral images. The channels that form the image are divided into two groups: main and additional (detecting) channels. A distinguishing feature of the main channels is a significant correlation between neighbors. Variants of prediction from the adjacent channel for the main and additional channels for lossless image compression have been considered. In the experiment on a series of images of different contents, the proposed algorithm showed a superior compression ratio in comparison with the popular WinRar, 7z, PNG archivers for all the considered variants of context-based prediction. The leader among the popular archivers, JPEG-LS, was surpassed by our algorithm in the record configuration 2b (and also in the non-record configurations 1a, and 1b) on the image from the Landsat series by 40% with a compression ratio of 2.82 versus 1.99. The best results were demonstrated with the approach, when one channel was used as a reference (master) channel and was compressed independently, and each subsequent channel used the values of pixels from previous channels to form the context. The compression ratio was further increased by using 1-byte increments for channels instead of channel values, while the use of channel values in the context was preserved. We expect to continue research on a wider sample of images and to use this algorithm to compress multispectral images with a greater number of channels.

7. References

[1] Cleary J, and Witten I 1984 Data compression using adaptive coding and partial string matching IEEE *Transactions on Communications* **32(4)** 396-402
[2] Mahoney M 2005 Adaptive weighing of context models for lossless data compression *Florida Tech. Technical Report* **CS-16**
[3] Ageenko E 2000 *Contex-based Compression of Binary Images* (University of Joensuu, Finland) p 120
[4] Ageenko E 2002 Context-based compression of binary images in parallel *Journal Software - Practice & Experience* **32(13)** 1223-1237 DOI 10.1002/spe.480
[5] Marpe D 2001 Video Compression Using Context-Based Adaptive Arithmetic Coding *International Conference on Image Processing* **D-10587** 558-561 DOI: 10.1109/ ICIP.2001.958175
[6] Zhang Y and Adjeroh D A 2008 Prediction by partial approximate matching for lossless image compression *IEEE Trans. Image Process.* **17(6)** 924-935
[7] Knoll B and de Freitas N 2011 *A Machine Learning Perspective on Predictive Coding with PAQ* (University of British Columbia, Vancouver, Canada)
[8] Gashnikov M V, Glumov N I and Sergeev V V 2010 A hierarchical compression method for space images *Automation and Remote Control* **3** 147-161
[9] Gashnikov M V and Glumov N I 2014 Hierarchical compression for hyperspectral image storage *Computer Optics* **38(3)** 482-488
[10] Gashnikov M V and Glumov N I 2016 Onboard processing of hyperspectral data in the remote sensing systems based on hierarchical compression *Computer Optics* **40(4)** 543-551 DOI: 10.18287/2412-6179-2016-40-4-543-551
[11] Yusov E and Turlapov V 2008 JPEG2000-based compressed multiresolution model for real-time large-scale terrain visualization *Int. Conf. on Computer Graphics and Vision, Proceedings* 8 (Access mode: www.graphicon.ru/html/2008/proceedings/English/S8/Paper_1.pdf)
[12] Borusyak A V and Vasin Yu G 2015 Algorithm for compression of indexed images using context-based modelling *Proceedings of the 9th Open German-Russian Conference on Image Recognition and Understanding* 60-62
[13] Borusyak A V and Vasin Yu G 2016 Development of the algorithm for adaptive compression of indexed images using context-based modelling *Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications* **26(1)** 4-8
[14] Vasin Yu G and Borusyak A V 2016 Compression of large-format images by means of statistical coding using context-based modeling *Proceedings of International Scientific Conference Situational centers and class 4i information-analytical systems for monitoring and security tasks* 274-278

Acknowledgements
This work was supported by a grant from the Russian Science Foundation No. 16-11-00068.

3d synthetic aperture radar image

A N Leukhin[1], A A Rozentsov[2], V I Bezrodnyy[1], A A Voronin[1], D Yu Karasev[2] and N A Kokovihina[1]

[1] Mari State University, Lenin square 1, Yoskar-Ola, Russia, 424000
[2] Volga State University of Technology, Lenin square 3, Yoskar-Ola, Russia, 424000

Abstract. Synthetic aperture radar (SAR) is a coherent active microwave imaging method. In remote sensing it is used for mapping the scattering properties of the Earth's surface in the respective wavelength domain. The algorithms for the formation of 3D radar images in multi-position interferometric systems for remote sensing of the Earth are considered. Examples of reconstruction of the relief map for systems with one and two transmit antenna are presented.

1. Introduction

The application of interferometric data processing to obtain information about the terrain and its changes, implementation of high resolution (1-3 m) regimes have become the main trends in the development of modern radar systems for space observation. Such processing of space-based synthetic aperture radar`s data includes the following steps: synthesis a pair of complex radar images of the same surface region, their spatial overlap with the formation of an interferogram; phase noise filtering on the obtained interferogram; deployment of the phase of the interferogram and its full geocoding (recalculation of the values of the expanded phase in the values of the relief heights and the transition from the flight coordinate system to any cartographic projection).

The purpose of this paper is to review the methods of interferometric data processing to obtain 3d synthetic radar image of a surface model.

It is necessary to first describe the basics of 3d SAR imaging and explain its features.

A second antenna is installed on an aircraft at a certain distance from the first antenna in order to enable forming a three-dimensional map of the underlying surface mode in synthetic aperture radar. As a rule, the spacing can be carried out either in height or in a horizontal plane perpendicular to the direction of flight.

There are several options for the operation of SAR in 3D mode [1]:

1. Radiation is performed through both antennas by turns.
2. Radiation is made through one antenna (for example, A_1), and receiving is performed through two antennas A_1 and A_2.
3. On the radiation, the transmitter operates on one of its own antennas (A_3), the receivers have two own antennas A_1 and A_2.

3D images of surface are analysed by hypertrace transformation [2].

Let's consider the basic geometrical relations necessary for calculation of objects` height in SAR 3D mode.

2. "Two transmitters, two receivers" system

Figure 1 shows the basic geometric equations for the 3D SAR system, for the case when the aircraft is equipped with two transmitter and two receivers, operating through its own spaced from each other antenna.

Using the basis of the Pythagorean theorem we can write:

$$\begin{cases}(H_1-h)^2+r_1^2=R_1^2\\(H_2-h)^2+(r_1+d)^2=R_2^2\end{cases} \quad (1)$$

Values of the heights H_1 and H_2 of the antennas and antenna spacing d in the horizontal plane can be considered known while SAR is working. Also known and sloped range R_1 and R_2. You need to find the horizontal distance r_1 and the height of the object h. Before solving the equation (1) we can first calculate r_1 and then determine h or vice versa. The solution to the system is determined by the ratio between the values d and $|H_1-H_2|$. If $|H_1-H_2|>d$ you choose the first method, otherwise the second. Two solutions are obtained when you use each method. A preference in favor of a solution that gives value r_1 least different from the values of the horizontal range, designed for zero height $r_1^{(0)}=(R_1^2-H_1^2)^{1/2}$, i.e. $\left|r_1-r_1^{(0)}\right|\to\min$ [3].

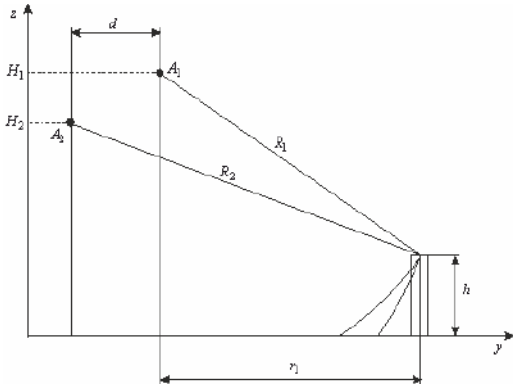


Figure 1. Basic geometric equations in "two transmitters, two receivers" system.

Possible solutions of the equation (1) are defined by equations (2) and (3):

$$\begin{aligned}r1=&-(H1^2\cdot d+H2^2\cdot d+R1^2\cdot d-R2^2\cdot d+d^3+2\cdot H1\cdot H2\cdot d+\\&+(H1^2-2\cdot H1\cdot H2+H2^2-R1^2+2\cdot R1\cdot R2-R2^2+d^2)^{1/2}\cdot\\&\cdot(2\cdot H1\cdot H2-H1^2-H2^2+R1^2+2\cdot R1\cdot R2+R2^2-d^2)^{1/2}\cdot\\&\cdot(H2-H1))\cdot\frac{1}{2\cdot(H1^2-2\cdot H1\cdot H2+H2^2+d^2)};\\h=&-\frac{H2^2-H1^2+R1^2-R2^2+d^2+2\cdot r1\cdot d}{2\cdot H1-2\cdot H2}\end{aligned} \quad (2)$$

$$\begin{aligned}h=&-(H1\cdot d^2+H2\cdot d^2+H1^3+H2^3-H1\cdot H2^2-H1^2\cdot H2-\\&-H1\cdot R1^2+H1\cdot R2^2+H2\cdot R1^2-H2\cdot R2^2+\cdot d\cdot\\&\cdot(H1^2-2\cdot H1\cdot H2+H2^2-R1^2+2\cdot R1\cdot R2-R2^2+d^2)^{1/2}\cdot\\&\cdot(2\cdot H1\cdot H2-H1^2-H2^2+R1^2+2\cdot R1\cdot R2+R2^2-d^2)^{1/2})\cdot\\&\cdot\frac{1}{2\cdot(H1^2-2\cdot H1\cdot H2+H2^2+d^2)};\\r=&-\frac{R1^2-R2^2+d^2+(H2-h)^2-(H1-h)^2}{2\cdot d}\end{aligned} \quad (3)$$

On the practice ranges R_1 and R_2 are known with an item resolution accuracy, so the calculations of value h and r_1 use value $\Delta R = R_2 - R_1$, calculated on the basis of the measurement result of the phase difference

$$\Delta R = \frac{\Delta\varphi^{(w)}}{4\pi}\lambda = \frac{\varphi_2 - \varphi_1}{4\pi}\lambda, \quad (4)$$

where φ_1 and φ_2 - phases of the signals received from the first and second antennas, respectively; λ - wave`s length.

Because the value of the phase shift is in the range $[0,2\pi)$, as a rule, there is ambiguity of the measurement values ΔR. To fix it uses the "unwrap" phase. Consider possible approaches to its implementation [4].

The first approach is based on the preliminary construction of the dependence of phase shifts to the zero level, depending on slant range to the antenna A_1 [5]:

$$\Delta\varphi^{(0)}(R_1) = 4\pi\Delta R^{(0)}(R_1)/\lambda = 4\pi((H_2{}^2 + ((R_1^2 - H_1^2)^{1/2} + d)^2)^{1/2} - R_1)/\lambda \quad (5)$$

Then the value of the phase shift $\Delta\varphi^{(u)}(R_1)$ used to calculate the difference between the sloping distances, determined from the relationship:

$$\Delta\varphi^{(u)}(R_1) = \Delta\varphi^{(w)} + trunc\left(\frac{\Delta\varphi^{(0)}(R_1)}{2\pi}\right)2\pi \quad (6)$$

The value of the slant range R_2 used for the calculation h and r_1 is determined from the relation:

$$R_2 = R_1 + \frac{\Delta\varphi^{(u)}}{4\pi}\lambda \pm 2\pi m \quad (7)$$

To determine the value m, you can use the following approach. Typically, the height difference between adjacent pixels is relatively small and we can assume that the height of the object in some neighborhood is constant. Presented according to m estimates of the altitude differences in the neighboring pixels have a pronounced minimum, and calculations show that this minimum is achieved when the value of the parameter $\hat{m}$ corresponding to the true values of the altitude and slant range. The algorithm of the calculation value $\hat{m}$ is the following:

1. Around the current image point with coordinates x_0, y_0, set the gate;
2. Sets the range of values m: $m = m_{\min}..m_{\max}$
3. For each point in the gate with coordinates x, y, a values $r_{x,y}$ and $h_{x,y}$ are calculated with current value m;
4. Calculate total measurement error of the heights and horizontal distances. Terms $(y_1 - y_0)$ and $(y_2 - y_0)$ take into account the current offset of pixels in the horizontal range:

$$\Delta h_m = \sum_{x1=x0-dx}^{x0+dx}\sum_{y1=y0-dy}^{y0+dy}\sum_{x2=x0-dx}^{x0+dx}\sum_{y2=y0-dy}^{y0+dy}\left|h_{x1,y1} - h_{x2,y2}\right|$$
$$\Delta r_m = \sum_{x1=x0-dx}^{x0+dx}\sum_{y1=y0-dy}^{y0+dy}\sum_{x2=x0-dx}^{x0+dx}\sum_{y2=y0-dy}^{y0+dy}\Delta y_{x1,x2,y1,y2} \quad (8)$$
$$\Delta y_{x1,x2,y1,y2} = \left|\left(r_{x1,y1} - (y_1 - y_0)\right) - \left(r_{x2,y2} - (y_2 - y_0)\right)\right|$$

5. As a result, selects the value $\hat{m}$ at which

$$\Delta h_{\hat{m}} \to \min, \ \Delta\kappa_{\hat{m}} \to \min \quad (9)$$

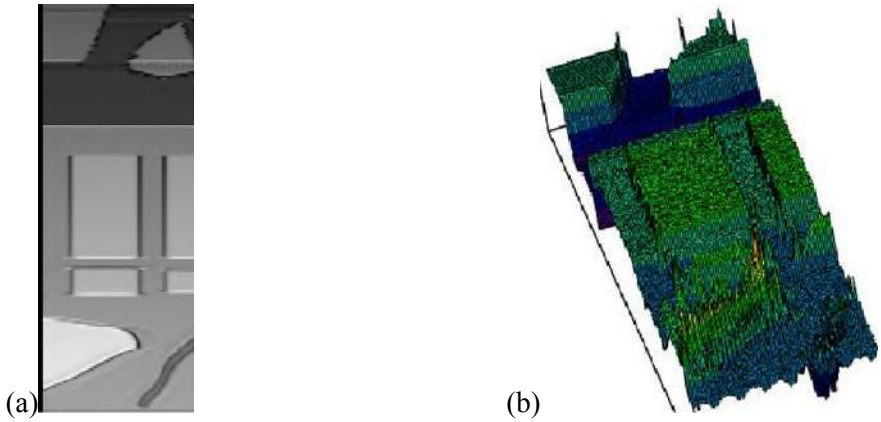
Because of spacing antennas, the resulting images have some mutual shift, and its magnitude will depend on the range. In this regard, before calculating the elevation, you must perform the mutual correction of the shifts of image elements. For this calculate values of distances R_2 corresponding to

the range R_1 on the zero level and overwrite the elements of the second array with the ranges R_2 in cells that correspond to values in the range R_1 [6]:

$$J_{R1,x}^{(кор)} = J_{R2,x}^{(2)}$$
$$R_2 = (H_2{}^2 + ((R_1^2 - H_1^2)^{1/2} + d)^2)^{1/2} \quad (10)$$

where $J_{R2,x}^{(2)}$ - image from the second antenna.

Figure 2 shows an example of the recovery bump maps for the considered case when the following system parameters: $H_1 = H_2 = 10m$, $d = 5m$ the maximum altitude of the relief 5m.



(a) (b)

Figure 2. An example of the recovery of the terrain surface. a – synthesized image, b – 3D relief of the surface.

3. "Two transmitters, one receiver" system

Figure 3 shows the basic geometric equations for the 3D SAR system, for the case when an aircraft equipped with one transmitter operating, for example, via an antenna A_1, and two receivers, operating through its own separated antennas A_1 and A_2.

In this system a signal between antennas A_1 and A_2 the point on the object surface takes place in two ways $S_1 = 2R_1$ and $S_2 = R_1 + R_2$ respectively. In the simulation of the hologram samples of the signal in the first image is recorded with a pixel corresponding to the distance R_1, and a second hologram pixel $\frac{R_1 + R_2}{2}$.

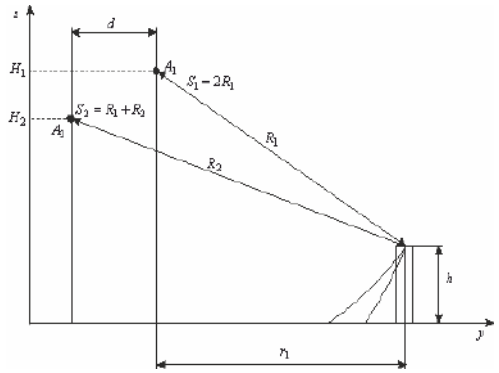


Figure 3. Basic geometric equations in "two transmitters, one receiver" system.

The value of the path S_2 is determined based on the known path S_1 and phase difference due to antenna spacing:

$$S_2 = S_1 + \frac{\Delta\varphi}{2\pi}\lambda m, m = \ldots,-2,-1,0,1,2,\ldots \quad (11)$$

Knowing the values of the parameters S_1, S_2, H_1, H_2, d, you can find value and a horizontal range r_1:

$$\begin{gathered} D2 = (-d^2 \cdot (-d^2 \cdot S1^2 - 2 \cdot H1^2 \cdot S2^2 - 2 \cdot H2^2 \cdot S2^2 + 2 \cdot S1 \cdot d^2 \cdot S2 + S2^4 - H2^2 \cdot S1^2 - \\ - 4 \cdot H1^3 \cdot H2 - 2 \cdot d^2 \cdot S2^2 - H1^2 \cdot S1^2 + d^4 + 2 \cdot H2^2 \cdot d^2 + H1^4 + H2^4 + 2 \cdot H1^2 \cdot d^2 + \\ + 6 \cdot H1^2 \cdot H2^2 + 2 \cdot S1 \cdot S2 \cdot H1^2 - 2 \cdot S1 \cdot S2^3 - 4 \cdot H1 \cdot H2 \cdot d^2 - 4 \cdot H1 \cdot H2^3 + 4 \cdot H1 \cdot H2 \cdot S1^2 - \\ - 4 \cdot H1 \cdot H2 \cdot S1 \cdot S2 + 2 \cdot H2^2 \cdot S1 \cdot S2 + 4 \cdot H1 \cdot H2 \cdot S2^2 + S1^2 \cdot S2^2))^{0.5} \\ h = \frac{-H1 \cdot S1 \cdot S2 + H2 \cdot d^2 - H1^2 \cdot H2 + H1^3 + H2^3 + H1 \cdot d^2 - H2 \cdot S2^2}{2 \cdot (H1^2 + H2^2 + d^2 - 2 \cdot H1 \cdot H2)} + \\ + \frac{H1 \cdot S2^2 + H2 \cdot S1 \cdot S2 - H1 \cdot H2^2 - D2}{2 \cdot (H1^2 + H2^2 + d^2 - 2 \cdot H1 \cdot H2)} \end{gathered} \quad (12)$$

As in the first case, measuring the phase shift may be ambiguous. The dependence of phase shifts on the zero level from the slant range to the antenna A_1 is described by the expression:

$$\Delta\varphi^{(0)}(R_1) = 2\pi \frac{(H_2{}^2 + ((R_1^2 - H_1^2)^{1/2} + d)^2)^{1/2} - R_1}{\lambda} \quad (13)$$

Then the value of the phase shift $\Delta\varphi^{(u)}(R_1)$ is used to calculate the difference between the sloping distances is determined from the equation (6). Path S_2 is used for the calculation h and r_1 is determined from the expression:

$$S_2 = 2R_1 + \frac{\Delta\varphi^{(u)} \pm 2\pi m}{2\pi}\lambda \quad (14)$$

The value m is chosen by the equation (9).

Still from the spacing antennas, the elements of the resulting images can have some mutual shift. To compensate for this shift is necessary to perform the image correction according to the relation [7]:

$$J^{(кор)}_{R_1,x} = J^{(2)}_{\frac{R_1+R_2}{2},x}, \quad (15)$$

where R_2 is computed as in equation (10).

Figure 4 shows an example of the recovery bump maps for the considered case when the following system parameters $H_1 = H_2 = 10m$, $d = 5m$ the maximum altitude of the relief 5m.

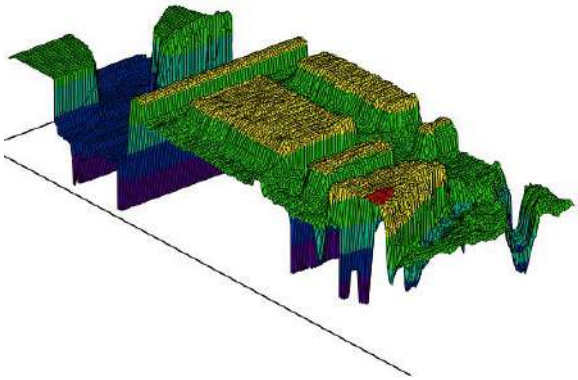


Figure 4. Synthesized 3D image of the surface.

4. Conclusions

Solution of a problem of restoration of a landscape`s in a synthetic aperture radar at various configuration of a reception-transmitting path is considered in this work. An original phase unwrapping algorithm based on joint minimization of the estimating error of object`s height and

sloped range in neighboring pixels of the image is proposed. Examples of the restored 3D images are presented.

5. References

[1] Richard B, Philipp H 1998 Synthetic aperture radar interferometry *Inverse Problems* **14(4)** R1-R54
[2] Fedotov N, Syemov A and Moiseev A 2016 Analysis of conditions that influence the properties of the consructed 3d-image *Computer Optics* **40(6)** 887-894 DOI: 10.18287/2412-6179-2016-40-6-887-894
[3] Kobernichenko V and Sosnovsky A 2012 Interferometer data processing space radar imagery of high resolution *Physics of Wave Processes and Radio Engineering Systems* **15(3)** 75-83
[4] Vidal-Pantaleoni A, Rafael O and Miguel F 1999 A Comparison of Phase Unwrapping Techniques in Synthetic Aperture Radar Interferometry *IEEE International Geoscience and Remote Sensing Symposium* 1354-1356
[5] Osmanoglu B and Dixon T 2011 On the importance of path for phase unwrapping in synthetic aperture radar interferometry *Applied Optics* **50(19)** 3205-3220
[6] Richards M 2007 A Beginner's Guide to Interferometric SAR Concepts and Signal Processing *IEEE A & E Systems Magazine* **22(9)** 5-29
[7] Sosnovsky A and Kobernichenko V 2012 Phase unwrapping algorithms investigation in digital elevation maps generation using space-based InSAR *Izvestiya of the Higher Educational Establishments of Russia. Radio electronics* **7** 84-92

Acknowledgments
The work is executed at financial support of the Ministry of Education and Science of the Russian Federation, project No. 2.2226.2017/Project Part and project No. 2.9140.2017/Basic Part. The work is performed under financial support of Russian Found of Basic Research, research project No. 15-07-99514.

Modeling of polarization SAR system

D V Karasev[1], A N Leukhin[2], A A Voronin[2] and V I Bezrodny[2]

[1]Volga State University of Technology, Lenin Sq. 3, Yoshkar-Ola, Russia,
[2]424000 Mari State University, Lenin Sq. 1, Yoshkar-Ola, Russia 424000

Abstract. We have implemented a combined mathematical model for polarization SAR. It allows simulating reflective objects on the background of the earth. We define the characteristics of underlying surface as an approximation of real experimental data. Value of backscattered field dependent on the angle of incidence. We get polarization characteristics of metal objects from the geometry configuration. It is based on the methods of physical optics and the physical theory of diffraction.

1. Introduction

Mathematical modeling plays important role in the creation of Synthetic Aperture Radar (SAR). It allows investigating the influence of observation conditions in a wide range of parameters. Moreover, it allows us to develop new algorithms in lack of experimental data.

A promising method of increasing the informativeness of radar images is polarimetric data. Polarimetric SAR gives us extra information about properties of the underlying surface. This information makes it possible to classify structures on the image even if objects have similar shape or brightness on the image. This classification finds its application both in the military and civilian industry [1].

Computational complexity and amount of data of the SAR imaging are quite big. Often it is very hard to create real-time processing possible. A lot more data required to simulate radio wave propagation. For each point in radio hologram, we should calculate RCS for every object and the underlying surface. Precise methods such as Finite Element Method (FEM) or Finite Difference Time Domain (FDTD) method unsuitable in this case. Approximate methods needs ratio $\lambda > 3 \cdot l$ while precise $\lambda > 10 \cdot l$ where l — length of the largest element of mesh.

In this paper, we discuss the raycast methods that take into account the polarization characteristics. Such methods do not reflect the exact solutions of the wave equations but provide a good approximation.

2. Decomposition of the scattering matrix

There are several methods for the scattering matrix decomposition. They represents scattering matrix as a linear combination of the matrices. Each of them corresponding to the basic scattering mechanisms [2].

One such method is the representation of the scattering matrix as a sum of the Pauli matrices:

$$S = \frac{a}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{b}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \frac{c}{\sqrt{2}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \frac{d}{\sqrt{2}} \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix}. \quad (1)$$

The first term corresponds to the single scattering without change of polarization. The second term corresponds to a double reflection in which one of the orthogonal components changes sign. The third term represents the scattering-on dihedral reflector oriented at an angle of 45 degrees to the vertical. When wave reflected from such a reflector, the polarization changes to an orthogonal one. In the case of backscattering, the Pauli basis will include only the first three matrices.

Another form of representation of the scattering matrix is the Krogager decomposition. This decomposition consists of three elements. First one is equivalent of scattering on a sphere. The second one represents a dihedral corner reflector. The third element represents helix. For the last two types, the matrices depend on the orientation angle θ of the reflector.

The scattering matrix for this expansion is as follows:

$$S = k_{\mathrm{s}} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + k_{\mathrm{d}} \cdot \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} + k_{\mathrm{h}} \cdot \mathrm{e}^{\mp \mathrm{i}2\theta} \cdot \begin{bmatrix} 1 & \pm\mathrm{i} \\ \pm\mathrm{i} & 1 \end{bmatrix}. \tag{2}$$

The coefficients k_{s}, k_{d}, and k_{h} determine the contribution of the corresponding scattering mechanisms.

Both Pauli and Krogager decomposition allow a visual assessment of the geometry. Each decomposition reflects the degree of heterogeneity of the surface. It may be useful to distinguish between natural and artificial objects.

Another method provide the analysis based on the coherence matrix T. Elements of this matrix are calculated by transformations of scattering matrix S.

$$T = \frac{1}{2} \begin{bmatrix} (S_{\mathrm{HH}} + S_{\mathrm{VV}}) \cdot (S_{\mathrm{HH}} + S_{\mathrm{VV}})^* & (S_{\mathrm{HH}} + S_{\mathrm{VV}}) \cdot (S_{\mathrm{HH}} - S_{\mathrm{VV}})^* & 2(S_{\mathrm{HH}} + S_{\mathrm{VV}}) S_{\mathrm{HV}}^* \\ (S_{\mathrm{HH}} - S_{\mathrm{VV}}) \cdot (S_{\mathrm{HH}} + S_{\mathrm{VV}})^* & (S_{\mathrm{HH}} - S_{\mathrm{VV}}) \cdot (S_{\mathrm{HH}} - S_{\mathrm{VV}})^* & 2(S_{\mathrm{HH}} - S_{\mathrm{VV}}) S_{\mathrm{HV}}^* \\ 2S_{\mathrm{HV}}(S_{\mathrm{HH}} + S_{\mathrm{VV}})^* & 2S_{\mathrm{HV}}(S_{\mathrm{HH}} + S_{\mathrm{VV}})^* & 4S_{\mathrm{HV}} S_{\mathrm{HV}}^* \end{bmatrix}. \tag{3}$$

The coherence matrix has three positive eigenvalues λ_1, λ_2, λ_3. In [3] it was suggested to use the relation:

$$P_j = \frac{\lambda_j}{\lambda_1 + \lambda_2 + \lambda_3}, \quad j = 1, 2, 3; \tag{4}$$

with the use of which a parameter was introduced, called the scattering entropy

$$H = -\sum_{j=1}^{3} P_j \log_3 P_j. \tag{5}$$

Scattering entropy represents the degree of randomness of the scattering, its values lie between 0 and 1. The value $H = 0$ corresponds to the perfect single reflection mechanism, and the value $H = 1$ — complete diffuse scattering.

Coherence matrix can be reduced to diagonal form by transformation

$$T = U \cdot \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \cdot U^{-1}, \tag{6}$$

$$U = \begin{bmatrix} \cos\alpha_1 & \cos\alpha_2 & \cos\alpha_3 \\ \sin\alpha_1 \cos\beta_1 \exp(\mathrm{i}\delta_1) & \sin\alpha_2 \cos\beta_2 \exp(\mathrm{i}\delta_2) & \sin\alpha_3 \cos\beta_3 \exp(\mathrm{i}\delta_3) \\ \sin\alpha_1 \cos\beta_1 \exp(\mathrm{i}\gamma_1) & \sin\alpha_2 \cos\beta_2 \exp(\mathrm{i}\gamma_2) & \sin\alpha_3 \cos\beta_3 \exp(\mathrm{i}\gamma_3) \end{bmatrix} \tag{7}$$

Based on the matrix U, the average angle can be calculated

$$\alpha = \sum_{j=1}^{3} P_j \alpha_j, \tag{8}$$

which characterizes the dominant scattering mechanism. A value of $\alpha = 0°$ corresponds to isotropic scattering on the surface, $\alpha = 45°$ — dipole scattering, and the value $\alpha = 90°$ — the double reflection.

Thus, methods of analyzing the coherence matrix allow us to divide natural objects into clusters with different scattering mechanisms.

3. Modeling of underlying surface SAR

The dimensions of the underlying surface are usually much larger than objects dimensions. It means that the surface needs an own approach to calculate scattered field.

We use a polygonal model of the surface. Each polygon has finite dimensions and corresponds with large-scale irregularities as presented at figure 1. In this case, the size of the facets should be less than the resolving power of the radar [4].

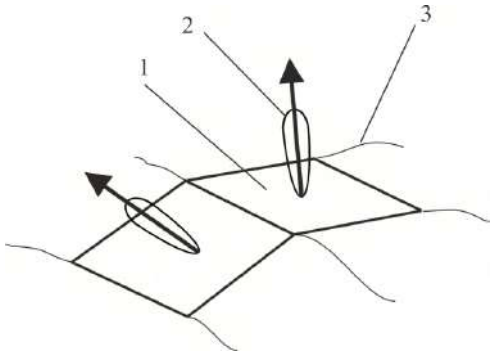


Figure 1. Faceted model of radar reflection by the earth's surface: 1 — facet; 2 — local backscattering diagram; 3 — surface.

The signal reflected from the surface is the sum of the signals from all the irradiated facets. The signal from individual facet has its own amplitude and its arbitrary phase. The relative position and backscattering diagram determine signal characteristics.

We have created a database of different types of reflecting areas. Each type has its own material on the 3D model of a scene. Type of material connects the physical characteristics with the particular facet. An incident angle on backscattering diagram defines specific Radar Cross Section (RCS) value. Each material has three backscattering diagrams for each polarization. At this moment we do not take into account anisotropy of cross polarization. So we assume that VH polarization equals HV polarization.

For the backscattering diagram, we took the experimental data given in [5, 6, 7]. In these works, RCS value is a mean value of the particular material. Thus it is a statistical approach to simulate underlying surface properties. In range [0; 90] degrees specific RCS is interpolated.

Examples of interpolated values of specific RCS in X-band are represented in figure 2.

At angles close to the vertical scattering for most surfaces will be close to the mirror reflection and the highest values of specific RCS. At the angles close to the horizontal, the backscattering will be very small. At intermediate values of the slip angle, the specific RCS, expressed in dB, varies with increasing slip angle according to a law close to linear.

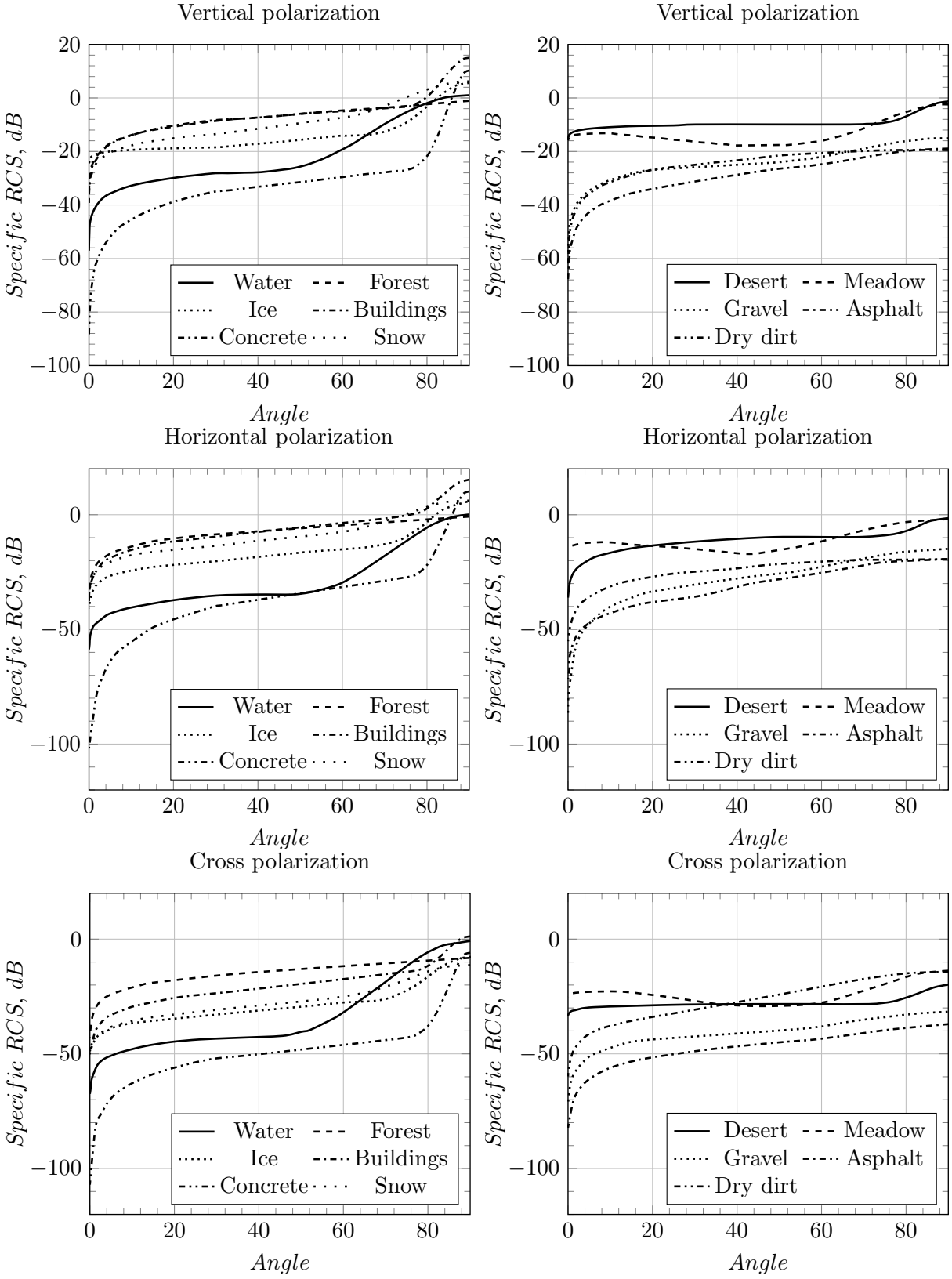


Figure 2. Dependence of specific RCS various surfaces of slip angle for different polarizations.

Experiments using polarimetric SAR given in [8], show that for a phase difference between reflected waves with matched polarization and cross-polarization has a uniform distribution in the interval $[0, 2\pi]$ for any distributed targets and, does not contain information about the target.

In contrast, the phase difference of the reflected radio waves on the matched polarizations depends both on the wavelength and the angle of incidence. Also, its influence has shape and dimensions, surface roughness, and material properties of the object.

Upon reflection from the relatively smooth surfaces of the phase difference will be close to zero. In the case of double reflection, for example from buildings or tree trunks, the phase difference will be close to 180°. In scattering from an inhomogeneous medium, for example from vegetation, the phase difference can vary from 0 to 180°. In some cases, there may be joint effect of these scattering mechanisms.

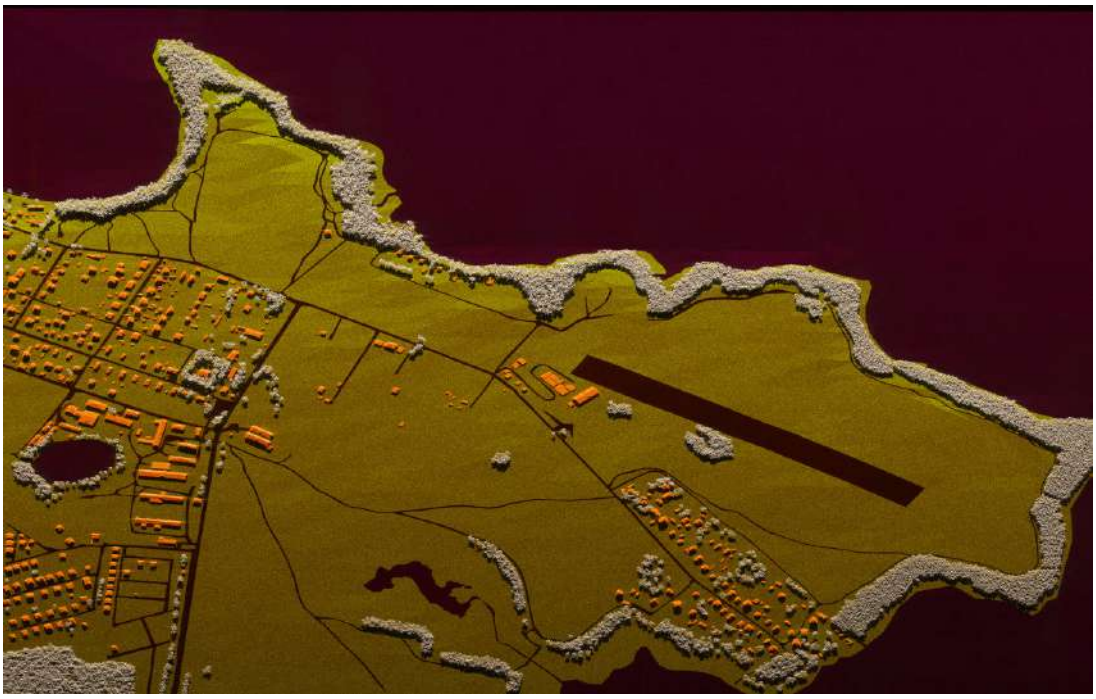


Figure 3. Radar image in pseudo colors.

Using the created model of radar signal reflections were simulated process of producing a radar image. Results are shown in figure 3. Different material types have their own pseudo colors: grass — yellow-green; houses — orange; trees — white; water — dark purple.

When generating radio holograms, the following model parameters were set:

- Flight altitude: 5 km;
- Flight path: linear;
- Resolution: 30 cm;
- Scanning angle: 45°;
- Modulation: phase;
- Carrier frequency: 10 GHz.

The simulated scene has the following dimensions: width 2700 m; length of 2500 m; height 40 m. The average height of the houses is one floor.

If a person tries to investigate each channel of the pseudo-color image as independent radio image, it will be hard to find a difference between VV and HH polarization. But anyway they do exist. The more obvious difference in comparison to cross-polarization. So, for the most comfortable representation combined image in pseudo-colors is needed. There are many ways to combine vertical, horizontal and cross polarization data, it may depend on the purpose of the SAR.

4. Modeling of metal objects

Often, of particular interest are metal objects created by man, especially for a military purpose. Due to the high conductivity, metal objects effectively reflect radio waves, and in this case, the most interesting is the dependence of the RCS on the shape of the object, and also on the angle of the survey.

The model under discussion is a composition of different approximating methods based on the raycasting technique. The basic idea of composition proposed in a number of papers in which authors described relationships for the calculation of the scattered field taking into account the polarization and bounces of the rays [9, 10, 11, 12]. Key methods described below.

William B. Gordon created far field approximation of the Kirchhoff formula for a field scattered on a metal plate of an arbitrary form usually is given by some surface (double) integral. This double integral can be reduced to a linear integral estimated around the boundary. Moreover, if the boundary is a polygon, this integral can be reduced to a finite sum [13]. Polygon approximation of 3D shape allows us to use the fastest way to calculate scattered field.

Ufimtsev Petr Yakovlevich in his book [14] studies the diffraction of electromagnetic waves on bodies large in comparison with the wavelength. Approximate and strict methods are discussed. The results obtained give light to nature of such phenomena as Fresnel diffraction, shadow radiation, depolarization backscattering, the process of formation of edge waves, and so on.

Shyh-Kang Jeng proposes in his paper [15] a method to compute the near-field RCS and Doppler spectrum of a target when the distances to the antennas are comparable to the target size. By dealing with a small piece of the target surface at a time, the transmitting antenna, and the receiving antenna are in the far-field zone of the small piece of the induced currents. The electromagnetic field produced by this small piece of induced currents can be written as a spherical wave. Sum up all spherical waves produced by every small piece of induced currents and we can obtain the total scattered field at the receiving antenna.

In figure 4 we used color pallet with some scaling coefficients as follows: red — HH polarization, blue — VV polarization, green — VH + HV polarization.

When generating radio holograms, the following model parameters were set:

- Flight altitude: 5 km;
- Flight path: linear;
- Resolution: 12.5 cm;
- Scanning angle: 45°;
- Modulation: chirp;
- Carrier frequency: 10 GHz.

Each image oriented as follows: slope range from left to right, azimuth from top to bottom. Every object has the size of 2 m: width, height, length and diameter are equal to 2 m.

It is clearly seen that the red and blue component dominate the majority of images. The brightest cross-polarization component is present in the image of the sphere and the cylinder. A significantly smaller amount of the green component is represented in the image of the cube. At the corner reflector, the bright cross-polarization component is present only in the center. In a square cross polarization plane missing almost completely.

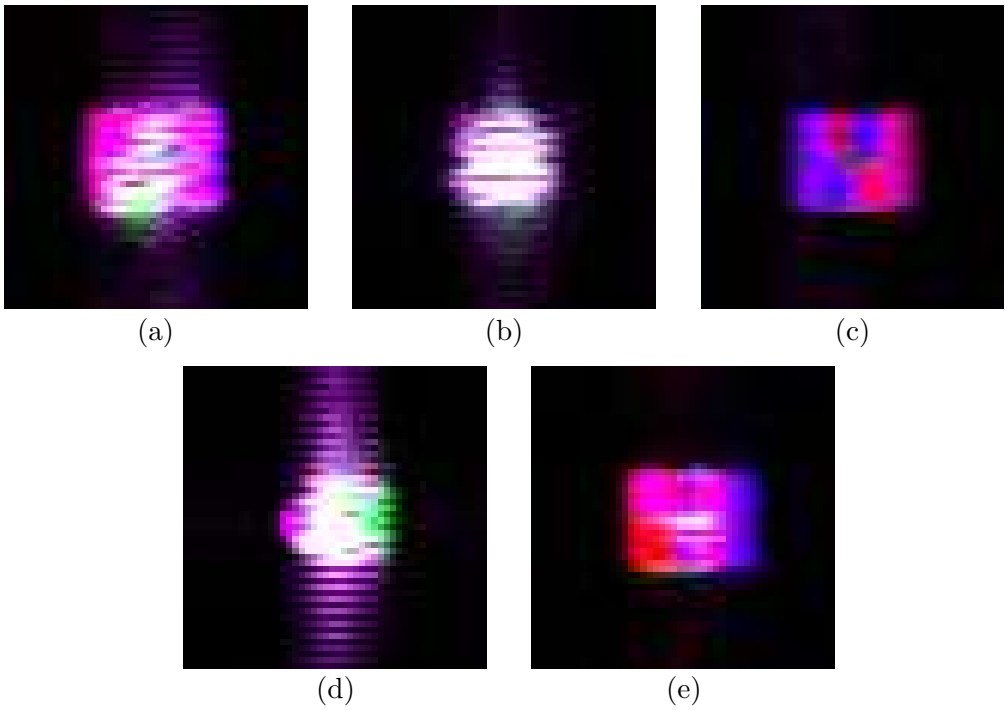


Figure 4. Radar image of the test objects in pseudo colors: (a) cube, (b) sphere, (c) plane, (d) cylinder, (e) corner reflector.

5. Conclusions

We implemented a mathematical model and SAR visualization using a combined technique. On the one hand, a statistical approach we used to model the underlying surface. This allows us to model large areas of view. On the other hand, for different m etallic objects, we produce a more accurate calculation based on physical optics and the physical theory of diffraction. T he a pproach t o m odeling d escribed i n t his p aper a llows o btaining t est d ata for the development of image segmentation algorithms. Figure 4 shows the dependence of the polarization characteristics on the shape of the object. Such kind of information can be useful in tasks of target recognition.

6. References

[1] Afanasyev A A and Zamyatin A V 2017 *Computer Optics* **41(3)** 431-440 DOI: 10.18287/2412-6179-2017-41-3-431-440
[2] Cloude S R and Pottier E 1996 *IEEE transactions on geoscience and remote sensing* **34** 498-518
[3] Cloude S R and Pottier E 1997 Application of the h/a/α polarimetric decomposition theorem for land dassification *Proc. SPIE* **3120** 132-143
[4] Baskakov A I, Zhutiaeva T S and Lukashenko Y I 2011 *Locating methods of research objects and environments* (Moscow: Academia) (in Russian)
[5] King C and Moore R K 1973 *A survey of terrain radar backscatter coefficient measure program* (Tech. Rep. The University of Kanzas, Center for Research Inc.) **2** 243
[6] Skolnik M 2008 *Radar Handbook* (New York: McGraw Hill)
[7] Katz I and Spetner L M 1960 *Journal of Research of the National Bureau of Standards* **64** 485
[8] Oh Y, Sarabandi K and Ulaby F T 2002 *IEEE Transactions on Geoscience and Remote Sensing* **40** 1348-1355

[9] Borzov A B, Suchkov V B and Sokolov A V 2004 Journal Of Radio Electronics (in Russian)
[10] Borzov A B, Suchkov V B, Shakhtarin B I and Sidorkina Y A 2014 *Journal of Communications Technology and Electronics* **59** 1356-1368
[11] Akhiyarov V V, Borzov A B, Suchkov V, Shakhtarin B I and Sidorkina Y A 2015 *Journal of Communications Technology and Electronics* **60** 1297-1304
[12] Suchkov V B 2013 *Systems and means of communication, television, and broadcasting* (in Russian)
[13] Gordon W B 1975 *IEEE Transactions on Antennas and Propagation* **23** 590-592
[14] Ufimtsev P Y 2009 *Theory of edge diffraction in electromagnetics* (California: The Institution of Engineering and Technology)
[15] Jeng S K 1998 *IEEE Transactions on Antennas and Propagation* **46** 551-558

Acknowledgments
The work is executed at financial support of the Ministry of Education and Science of the Russian Federation, project No. 2.2226.2017/Project Part and project No. 2.9140.2017/Basic Part.

Methods and applications for controlling the correctness of physical exercises performance

V L Rozaliev[1], A I Vybornyi[1], Y A Orlova[1] and A V Alekseev[1]

[1]Volgograd State Technical University, Lenina Avenue 28, Volgograd, Russia, 400005

Abstract. This document contains the description of the program for the control of the correct physical exercises performance implementation using Microsoft Kinect, the method of the comparison between live motions performed by the user and recorded motions, description of the testing of the program and also information about different approaches for gesture recognition.

1. Introduction

Nowadays the automation is used in many areas, including sport and physical culture. Most of the modern consoles have motion sensors. It gives developers big opportunities. The application of the new technologies and methods may be interesting and popular among people who care about their health, but do not have time for the gym and gamers who are interested in new experience. It is possible that in the near future people will give up going to the gym and hiring personal trainers and will use a virtual coach instead, doing physicals exercises at home in front of their consoles.

Program that is being developed may be useful both for the people who are recovering from injuries and for the people who just don't have enough time to go to the gym. The main objective of the program is to compare a predetermined sequence of the human movement with actual human movement captured via Kinect. The program should allow user to train at home controlling the way exercises are performed by the user and reporting to user about the mistakes he makes.

2. Different approaches for gesture recognition using MS Kinec

2.1. Hidden Markov Models

The method described by Jonathan Hall uses a Markov chain or a Markov Model. It is a typical model for a stochastic sequence of a finite number of states. These states are defined based on observations or data and these observations are essential for gesture recognition. In this approach, the observation data used are sequential 3D points (x, y, z) of Joints. A physical gesture can be understood as a Markov chain where the true states of the model S = s1, s2, s3,...,sN define the 3D position of Joints for each state. A gesture is recognized based on the states as well as the transition between these states. These states are hidden and hence this type of Markov model is called a Hidden Markov Model (HMM). At each state an output symbol O = o1, o2, o3,...,oM is emitted with some probability, and one state transitions to another with some probability. The emission and transition probabilities are learned while training the model with known gesture data and these values are stored in the emission and transition matrices. Each trained model can then be used to determine the probability with which a

given gesture appears in test data. In the manner described above, trained HMMs can be used to recognize gestures.[4, 10]

2.2. Gesture Service with Windows SDK

Gesture Service for Kinect project considers gestures to be made up of parts. In our context, parts refer to key poses of an Exercise and gesture refers to a sequence of key poses or in other words the complete Exercise. Each part of a gesture is a specific movement that, when combined with other gesture parts, makes up the whole gesture. Recognizing gesture parts are not sufficient to recognize a gesture. This is due to the fact that transitions between gesture parts play a crucial role in a gesture. To incorporate transitions, the method considers three results that a gesture part can return when it checks to see if it has been completed or not. The state of the gesture part is set to "Fail" if the user moved in a way that was inconsistent with the gesture part. The state of the gesture part is set to "Succeed" if the user performed a part of the gesture correctly and the system will automatically check for the next part of the gesture. Finally, the state of the gesture part is set to "Pausing" if the user is transitioning to the next gesture part. It indicates that the user did not fail the gesture but did not perform the next part either.

The overall system comprises of three classes, namely, gesture controller, gesture and gesture part. The method uses a Gesture Controller to control the transition between gesture parts and updates the state of the gesture part.[4, 8]

2.3. Gesture Service with Windows SDK

The Kinect Space provides a tool which allows everybody to record and automatically recognize customized gestures using the depth images and skeleton data as provided by the Kinect sensors. This method is very similar to the Hidden Markov Model as discussed before. The software observes and comprehends the user interaction by processing the skeleton of the user. The unique analysis routines allow to not only detect simple gestures such as pushing, clicking, forming a circle or waving, but also to recognize more complicated gestures as, for instance, used in dance performances or sign language. In addition it provides a visual feedback how good individual body parts resemble a given gesture. The system can be easily trained for recognizing a gesture without writing any code. [4]

2.4. Kinect SDK Dynamic Time Warping (DTW) Gesture Recognition

Kinect SDK Dynamic Time Warping Gesture Recognition project allows developers to include fast, reliable and highly customizable gesture recognition in Microsoft Kinect SDK C-sharp projects. It uses Dynamic time warping (DTW) algorithm for measuring similarity between two sequences which may vary in time or speed. It uses skeletal tracking but the drawback with this software is that it currently supports only 2D vectors and not 3D. The software includes a gesture recorder that records the user's skeleton and trains the system. The recognizer software then recognizes the gestures that have been trained by the user. [5, 6]

2.5. Neural Networks

Neural networks are also one of the most commonly used and effective methods of gesture recognition. This is reflected in various works and sources of different authors. [7, 9] In most cases, the latter have a "pattern" character (reference gesture / movement, with which a coincidence must occur).

The models for the gesture recognition have been constructed by using ten different Neural Networks (NN), one for each gesture, which have been trained providing a set of feature sequences of the same gesture as positive examples and the remaining sequences of other gestures as negative examples. Each NN has an input layer of 480 nodes corresponding to the feature vectors Vi for 60 consecutive frames, an hidden layer of 100 nodes and an output layer of one node trained to produce 1 if the gesture is recognized and zero otherwise. The Backpropagation Learning algorithm has been applied and the best configuration of hidden nodes has been selected in an heuristic way after several experiments. At the end of the learning phase, in order to recognize a gesture a sequence of features is provided to all the 10 NNs and the one which returns the maximum value is considered the winning

gesture. This classification procedure gives a result also when a gesture does not belong to any of the ten classes. For this reason a threshold has been introduced in order to decide if the maximum answer among the NN outputs has to be assigned to the corresponding class or not. [7]

3. Analysis methods of the control of the correct physical exercises performance

To implement the program it was decided for now to use angle comparison method. It wasn't mentioned in the section before, but the method is described below. The implementation of the program for the control of the correct physical exercises performance was divided into two main phases: recording phase and comparison phase.

3.1. Recording phase

In the first phase we needed to record the exercise and save it to file for the further use as a movement with a perfect form.

A few approaches have been tried for the recording. Firstly, we were trying to save only the coordinates and types of the joints. But this approach was too inefficient. Another suggested method was serialization and in the end it was decided to use it.

Serialization is the process of converting an object into a stream of bytes in order to store the object or transmit it to memory, a database, or a file. Its main purpose is to save the state of an object in order to be able to recreate it when needed.

The object is serialized to a stream, which carries not just the data, but information about the object's type, such as its version, culture, and assembly name. From that stream, it can be stored in a database, a file, or memory. [1]

Thus, with the use of the serialization the collection of frames with the data about the skeleton is saved to the file. This approach is good because we don't need to divide skeleton data and take only particular parts of it. Instead, the collection of frames is saved and each frame contains the whole information about the skeleton including coordinates, joints, type of joints, positions, orientation etc.

After we got the file with the information about the exercise we needed to read this file and process the data. To do it we used deserialization (the reverse process to the serialization).

3.2. Comparison phase

In Now that we have saved motion data about the exercise that will be considered as a standard, we need to compare the standard with the user motion.

The basic idea is to compare the angles of the joints in the recorded motion standard with angles of the joints in the user motion. So, everything comes down to the calculation of the angles between vectors and comparison of these angles. In our case parts of the human body may be considered as vectors. Thus we need to know the values of the angles of the joints in the recorded motion standard and in the user motion. These angles are calculated with the use of the same method. [2]

Method takes 4 arguments: skeleton data and 3 arguments which represent the types of the joints. For example, in addition to the skeleton data it could be three joint types - JointType.ShoulderCenter, JointType.ShoulderLeft, JointType.ElbowLeft. If we pass these arguments in our method the angle will be calculated in the left shoulder joint (JointType.ShoulderLeft). Combining the joints, we can get two vectors that have one common point, and this point will be the joint in which we are calculating the angle. Coordinates of vectors are obtained, vectors are normalized, cross product and dot product are calculated and then the angle between vectors is calculated using Atan2 method. This method returns the angle whose tangent is the quotient of two specified numbers.

The angles of the joints in the user motion are calculated every frame. For the motion standard angles are calculated in advance and stored in list. [10]

Because it's almost impossible to repeat the motion standard with 100 percent accuracy program takes into account little errors in the user motion. At this point the error is 15 degrees in the value of the angles. Also, program compares current frame of the user motion not only with the one frame of the motion standard but with the 16 closest frames (8 previous and 8 subsequent) because user may do exercise a little faster or a little slower than standard motion demands. Angles of the joints in the

recorded motion standard and angles of the joints in the user motion are compared in the loop. If the difference between angles is less than 15 degrees it means that user is doing motion correctly.

To calculate the accuracy of the particular repetition in a particular exercise using described above algorithm the result of the comparison (whether the motion has been done correctly or not) is saved and stored. And the results of the comparison are stored for every joint involved in the exercise. When motion standard "ends" (all the frames of the recorded motion has been played) the analysis of the results starts. For the analysis the percent of the exercise performance correctness is calculated. The amount of frames in which the motions of the user were correct is divided by the total amount of frames in this exercise. It is done for every joint involved in the exercise. And after this the conclusion is made about whether this repetition may be considered as the correct or not. If the percent of correctness for each joint is above 85 and the arithmetic mean of all the percents is above 90 then this repetition may be considered as a correctly performed.

To analyze whether the motion is done by the user correctly or not we also used production rule system. All the exercises were divided into groups by the joints that are used in the exercises. And the production rules were set for each group. The example of the production rules for one of the exercises (overhead squats) is described below.

The next joints are used in this exercise: Shoulders, Elbows, Spine, Hips, Knees.

Production rules for this exercise:

Rule 1: IF (exercise = jumping jack) OR (exercise = squats) OR (exercise = overhead squats) OR (exercise = hip raises) THEN (compare angles in shoulder joints = yes) AND (compare angles in elbow joints = yes) AND (compare angles in spine = yes) AND (compare angles in hips = yes) AND (compare angles in knees = yes)

Rule 2: IF (difference between angles in shoulder joints < 15) THEN (result of the shoulder joints comparison in the current frame = true)

Rule 3: IF (difference between angles in shoulder joints >= 15) THEN (result of the shoulder joints comparison in the current frame = false)

Rule 4: IF (difference between angles in elbow joints < 15) THEN (result of the elbow joints comparison in the current frame = true)

Rule 5: IF (difference between angles in elbow joints >= 15) THEN (result of the elbow joints comparison in the current frame = false)

Rule 6: IF (difference between angles in spine < 15) THEN (result of the comparison in spine in the current frame = true)

Rule 7: IF (difference between angles in spine >= 15) THEN (result of the comparison in spine in the current frame = false)

Rule 8: IF (difference between angles in hips < 15) THEN (result of the comparison in hips in the current frame = true)

Rule 9: IF (difference between angles in hips >= 15) THEN (result of the comparison in hips in the current frame = false)

Rule 10: IF (difference between angles in knee joints < 15) THEN (result of the knee joints comparison in the current frame = true)

Rule 11: IF (difference between angles in knee joints >= 15) THEN (result of the knee joints comparison in the current frame = false)

Rule 12: IF (percent of correctness in each joint >= 85) AND (average percent of correctness >= 90) THEN (repetition is counted = true)

Rule 13: IF (percent of correctness in each joint < 85) AND (average percent of correctness < 90) THEN (repetition is counted = false)

The screenshots of the program while user is doing the exercise are shown on the figures 1 and 2.

3.3. Combining comparison with recognition

To increase the accuracy of the exercises performance control we decided to combine the comparison method described in the section above with the method described in the section 2.2. This combination has brought better and more accurate results in a context of movement control and also allowed user to do exercises with their own speed and pace without need to repeat exactly after the program standard.

Figure 1. Exercise is performed correctly.

Figure 2. Exercise is performed incorrectly.

In general, the algorithm can be described as follows. Each frame user live data received via Skeleton Stream using Kinect camera is processed by obtaining information about the position of the joints, as well as by calculating the angles in the required joints. Then, the obtained data is compared with the recorded data from reference movement.

To obtain the current comparison result of a part of the motion, the relative positioning of the joints in the user's movement is compared first. If the result of the comparison is positive, that is, the relative position of the joints is the same, there is another comparison. Now the comparison takes place in the joints of the reference movement and the user's movement. And in the case of their coincidence, the current part of the movement is considered correctly executed.

After obtaining the result of the angles and the position of the joints comparison, based on this result, the current state of motion is checked and updated. In case the result indicates that the movement is true, the transition to the next part of the movement or the end of the movement occurs. Otherwise, joints with mismatches in comparison phase are defined, and depending on the result, it is concluded whether the exercise is performed incorrectly or not. If the result of the comparison is equal to "failed", or the number of frames in the current movement exceeded the maximum allowed, the exercise is considered to be performed incorrectly. If the result of the comparison is "uncertain", then the user still has the opportunity to perform the exercise correctly. At the same time, in both cases the joints, in which there are errors in the performance, are determined and marked during the exercise execution. The algorithm is shown on the figures 3 and 4.

4. Testing of the program

The program has been tested by 10 users and each of them tested all 10 exercises represented in the program. The results of the testing are shown in the table I. Columns with the numbers from 1 through 10 represent 10 users. If the user was able to do a few repetitions of the exercise in the table this exercise was marked with +. If the user wasn't able to do even one repetition of the exercise in the table this exercise was marked with -. If the user was able to do a few repetitions of the exercise with some difficulties in the table this exercise was marked with +-. The program has been tested with two methods. The one with gesture recognition in it and the one without it. The results of the testing are shown in the tables below.

Without gesture recognition method exercises such thrusters and hip raises caused the biggest difficulties among users. It means that we need to make these exercises easier (for example, by decreasing the percent of correctness threshold value) or replace them with the other exercises.

It is also worth mentioning that most of the users needed some time to get used to the way the exercises should be performed. Perhaps, decreasing the percent of correctness threshold value of the exercises would be the good idea not only for those movement that caused some difficulties among users but also for the other exercises. But with the combination of the gesture recognition and movement comparison methods the major part of users participating in the experiment didn't have difficulty with almost all of the exercises. Summing up the results of the testing we may say that users handled the performance of most of the exercises quite well and the method with combination of the gesture recognition and movement comparison works a lot better than the one without recognition.

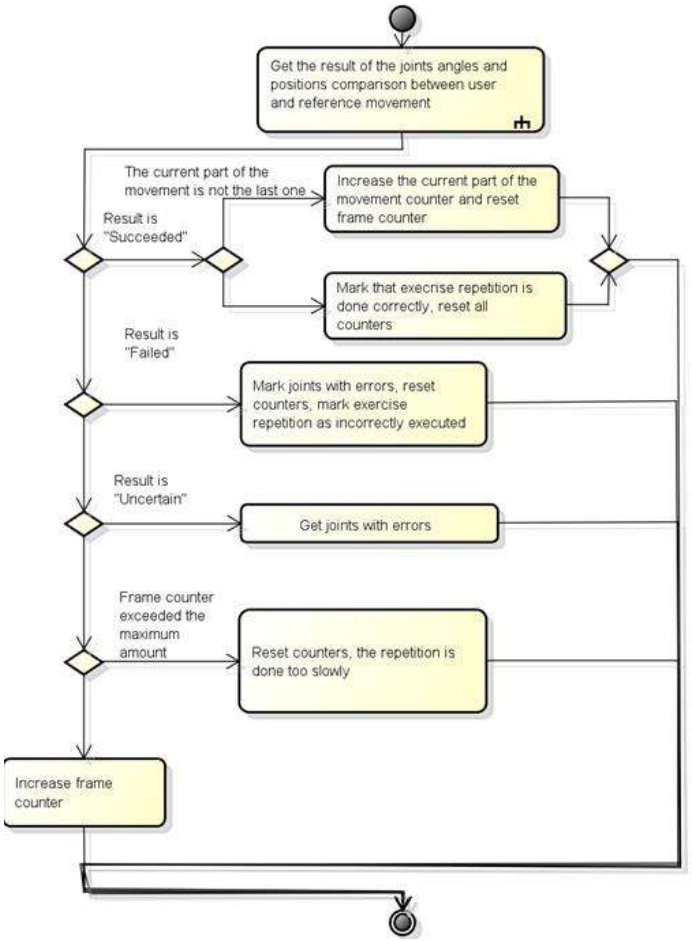


Figure 3. Part of the movement processing.

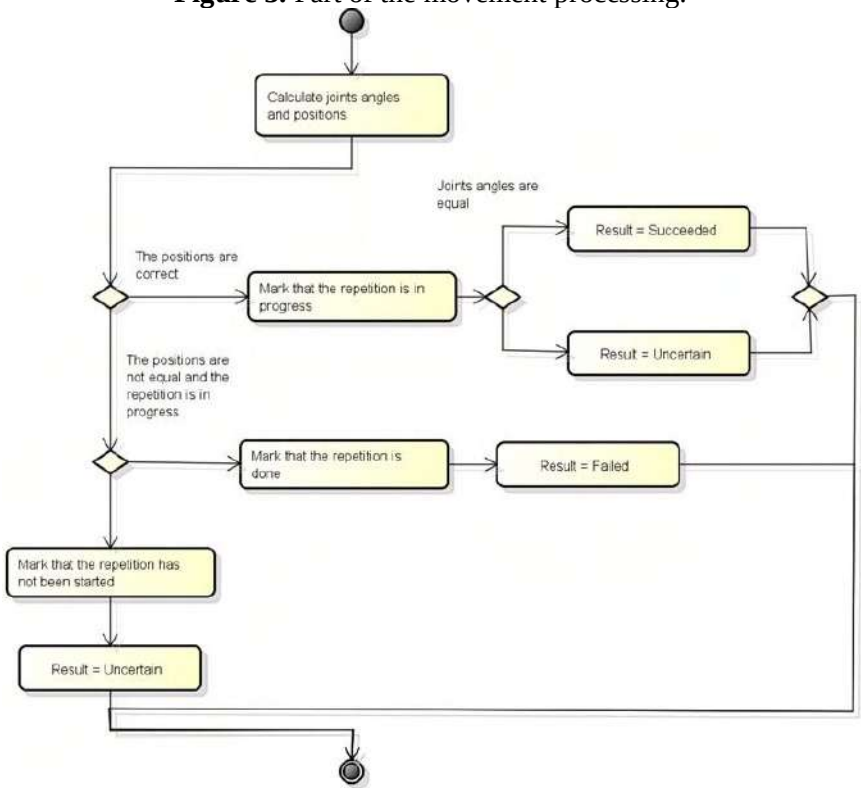


Figure 4. Getting the result of the joints positions and angles comparison.

Table 1. Results of the testing. Method without gesture recognition.

Exercise	№1	№2	№3	№4	№5	№6	№7	№8	№9	№10
Hand raises	+	+	+	+	+	+	+	+	+	+
Elbow rotation	+	+	±	+	+	+	+	+	±	+
Military press	+	+	+	+	+	±	+	+	+	+
Jumping Jack	+	+	±	±	+	±	+	+	+	±
Side bend	+	±	±	±	±	±	+	±	+	±
Squats	+	±	±	±	±	±	±	±	+	±
Overhead squats	+	+	±	±	±	+	±	±	+	+
Thrusters	±	–	–	–	–	–	–	–	±	–
Side lunges	+	+	+	+	+	+	+	+	+	+
Hip raises	±	–	–	–	–	–	–	–	–	–

Table 2. Results of the testing. Method with gesture recognition.

Exercise	№1	№2	№3	№4	№5	№6	№7	№8	№9	№10
Hand raises	+	+	+	+	+	+	+	+	+	+
Elbow rotation	+	+	+	+	+	+	+	+	±	+
Military press	+	+	+	+	+	+	+	+	+	+
Jumping Jack	+	+	+	+	+	±	+	+	+	±
Side bend	+	+	+	±	+	+	+	±	+	±
Squats	+	+	+	+	+	+	+	+	+	±
Overhead squats	+	+	+	±	+	+	+	±	+	+
Thrusters	+	±	+	±	+	+	+	+	+	–
Side lunges	+	+	+	+	+	+	+	+	+	+
Hip raises	+	±	+	±	+	+	+	+	±	±

5. Conclusion

This document contains the description of the program for the control of the correct physical exercises performance implementation using Microsoft Kinect.The implementation of the program was divided into two main phases: recording phase and comparison phase.In the first phase we had to record the human-motion and save it into a file for later processing. A few approaches (on how to read and save those data from tracked human skeleton) have been tried here. The successful approach that is used now is a serialization – saving the collection of skeleton frames into a data structure in binary format.

The second phase is a comparison between live motions performed by the user and recorded motions with the combination of gesture recognition. The main idea is to calculate recorded motion's joint angles and user's joint angles, compare them, considering a little error and then with the use of the production rule system analyze performed exercise to know whether the motion was correct or not.

At the moment there are ten exercises represented in the program, involving different joints and muscle groups. Program gives user the feedback about the performance of the exercises by marking the joints in which user make mistakes with red colour. Also, for more detailed information about the accuracy of the repetition performance user can open the output file that contains the percents of the accuracy for every joint in every exercise in every repetition.

6. References

[1] *MSDN - Microsoft Developer Network: Serialization (C and Visual Basic)* (Access mode: https://msdn.microsoft.com/ruru/library/ms233843.aspx)

[2] Hemed A 2012 Motion Comparison using Microsoft Kinect: FIT3036 *Computer Science* 27

[3] Orlova Y A, RozalievV L and Shpirko A A 2013 Automation of the control of the physical exercises performance for rehabilitation using Microsoft Kinect *Physical Education and Sports Training* **1** 53-58

[4] Ravi A 2013 Automatic Gesture Recognition and Tracking System for Physiotherapy *Electrical Engineering and Computer Sciences* (Technical Report No. UCB/EECS) p30

[5] *Codeplex - open source project hosting: Kinect SDK Dynamic Time Warping* 2011 (Gesture Recognition)
[6] D'Orazio T, Attolico C, Cicirelli G and Guaragnella C 2014 DTW Algorithm: Mining of gene expression time series with dynamic time warping techniques *A Neural Network Approach for Human Gesture Recognition with a Kinect Sensor* 741-746
[7] *MSDN - Microsoft Developer Network: Gesture service for the Kinect with the windows SDK* 2011 (MCS UK Solution Development)
[8] Tang A, Lu K, Wang Y, Huang J 2013 A Real-time Hand Posture Recognition System Using Deep Neural Networks *ACM Transactions on Intelligent Systems and Technology* **9(4)** 23
[9] Ghahramani Z 2012 An Introduction to Hidden Markov Models and Bayesian Networks *International Journal of Pattern Recognition and Artificial Intelligence* 25
[10] Kopenkov V N, Myasnikov V V 2016 Development of an algorithm for automatic construction of a computational procedure of local image processing, based on the hierarchical regression *Computer Optics* **40(5)** 713-720 DOI: 10.18287/2412-6179-2016-40-5-713-720

Acknowledgments
The work is partially supported by the Russian Foundation for Basic Research (16-07-00407, 16-07-00453, 16-47-340320, 18-07-00220 projects).

Investigation of hyperspectral image pixel signatures by the empirical mode decomposition method

P A Pakhomov[1], A V Borusyak[1] and V E Turlapov[1]

[1]National Research Lobachevsky State University of Nizhny Novgorod, Gagarin Avenue 23, Nizhny Novgorod, Russia, 603950

Abstract. The signature of hyperspectral image (HSI) pixels and their decomposition into empirical modes (EM) and low-frequency residuals are investigated. On the basis of estimates related to the EM-decomposition method, the possibility of switching from a 2-byte representation of the values of the HIS-signature to a 1-byte one is examined using the example of the Moffett Field from the AVIRIS spectrometer. It is revealed that the localization of the minimum window sizes for the first EM is correlated with the localization of the significant influence of the atmosphere; the first low-frequency residues have a fairly high correlation coefficient with the signature and the first 2 of them and their EM are most interesting for use; 50 of the 224 HIS-channels are noisy and can be excluded from consideration; EM with practically no loss of accuracy can be reduced to a 1-byte representation. The management of the classification capabilities of signatures by changing the threshold value of the correlation coefficient with the sample, as well as the application of the 1st and 2nd low-frequency residues in place of the signature, was studied. Classification capabilities of signatures in a 1-byte representation are almost equivalent to a 2-byte one, which makes it possible to put a signature with 1-byte representation as the object of compression. For the wavelet decomposition of the HSI data array, in combination with a 1-byte representation, a nearlossless compression ratio of 6.65 is obtained.

1. Introduction

Hyperspectral image (HSI) compression is the central task in the processing of such images. The first technological stage, where this task was of vital importance, was the transmission of Earth remote sensing (ERS) data to Earth, since it was not practicable to store and fully process such large amounts of data onboard a spacecraft/aircraft. Thus, in [1] published in 1997, the algorithm of Context-Based Adaptive Lossless Image Coding (CALIC) was proposed. This algorithm was two-dimensional, i.e. it compressed images channel-by-channel and, therefore, it was called 2D-CALIC. By that time, it had been shown that the classified context adaptive prediction implemented in the adaptive selection of adaptive predictors (ASAP) and adaptive combination of adaptive predictors (ACAP) methods yielded good results. The ACAP method was equipped with three-dimensional predictors obtained by learning based on land data. Both methods were applied in the well-known Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), where a data density of 6-5 bit/pixel was achieved. Another popular approach is based on the development of nonlinear predictors. This was implemented in the LOCO-I algorithm, standardized as JPEG-LS [2], and in the above-mentioned 2D-CALIC [1], which was soon upgraded by the authors to 3D and was called 3D-CALIC [3]. In the 2004 publication [4], an optimized version of the 3D-CALIC method, called M-CALIC (M-multiband), was proposed. It was based on the multi-

channel nature of context-based prediction. The transition from the lossless compression algorithm to the near-lossless one in the M-CALIC method was performed in accordance with [5] by applying the sampling of the values with a step $(2\delta+1)$, where δ is the permissible rounding error for pixel values.

The choice of the preferable type of compression very much depends on the applications where the HSI thus stored will be used. The papers on the interpretation of hyperspectral aerospace measurements for diagnosing the state of natural and technogenic objects, coastal waters, and crop areas [6], [7], [8] underline the importance of a qualitative improvement in the channel resolution in the HSI for precise classification of the type and state of the objects being observed, and note that these problems cannot be solved by means of multispectral images alone. At the same time, the high correlation of neighboring channels cannot be seen as a reason for excluding one of these channels as uninformative because the entire small fraction of the differing information is necessary to accurately determine the object of interest or its state.

One of the first approaches in the development of compression algorithms for HSI is to optimize the number of spectral channels while preserving the information value of hyperspectral imaging. This approach is reflected both in domestic [9] and in foreign sources [10], [11]. By the early 2000s, two main lines of research were defined to address the problem of optimizing the number of HSI channels [10]: 1) *Feature Extraction*; 2) *Band Selection*. With the possible exception of [12], these two lines of research have in common the possibility of detecting the most informative channels at the initial stage by the projective optimization method. In this method, the optimal approach is to project the initial HSI X of the size $n \times N$ (n is the number of channels, N is the number of pixels per channel) into a new HSI Y of a smaller size $m \times N$ ($m < n$), which maximizes the projection index $J=J(Y)$, where $Y = A^T X$, and A is an $n \times m$ matrix. Usually, the value of the index is estimated through the spectral data variance by the principal component method [9], [11], and the matrix A is formed from the columns (eigenvectors) of the covariance matrix of the original data. Sometimes, after applying the principal component method, the method of independent components, well-proven on non-Gaussian distributions, can be applied [9],[13].

The divergence of information J between two pixels of the HSI was chosen as a criterion in [14] for the synthesis of images reflecting the spatial distribution of the amplitude ratios of like pixels in different channels. Also of interest is the integral estimate for the amount of information proposed in [9] taking into account the signal-to-noise ratio.

In a series of works of 2008-2013 [15], a three-stage algorithm for lossless HSI compression was proposed, including:

1) Taking into account the relationship between HSI channels by calculating the correlation, constructing a linear predictor of the next channel value from the previous one, and forming arrays of deviations of the value of the next channel from the predicted one.
2) Forming an auxiliary data structure for storing unique pairs of groups of element values in a 1-byte representation, as well as pointers to these groups.
3) Compression of the data obtained after the transformations by means of a standard entropy algorithm by processing the generated auxiliary data structures.

An approximately 40% gain in compression was achieved compared to JPEG-LS. It was established that about 45% of the gain was obtained by using a 1-byte representation of deviations from the predicted values; 15-26% of the gain was obtained due to the reordering of the channels during compression.

In a series of papers published in 2009-2016 [16],[17],[18],[19] a method of hierarchical lossless compression for the purpose of storing HSI was constructed. The following requirements for the method were formulated: 1) the possibility of compression of multi-channel images; 2) quick access to fragments of compressed images at various scales; 3) low computational complexity of decompression; 4) strict error control; 5) high efficiency in "no-error" mode and with small errors; 6) use of interchannel dependencies; 7) quick access to the specified components of compressed images at various scales; 8) compression of 16-bit images.

The following dependencies on the channel number are investigated and used for compression: 1) the correlation coefficient with the next channel; 2) the channel average E; 3) the difference between

the minimum and maximum in the channel; 4) channel variance *D*. A method of "common reference components" for channel groups in combination with the "sliding approximation" method within the group is proposed, which permits to independently compress groups of neighboring channels. The best achieved value of lossless compression ratio was about 3.6; the recommended number of hierarchy levels was 4. The nature and the range of value changes in the channel $\Delta X_i = \max_k(x_{ki}) - \min_k(x_{ki})$ were shown depending on the channel number *i*. Thus, for the image "Urban and Mixed Environment" (SpecTIR spectrometer), the maximum value of the changes ΔX was about 32000, while for the "Cuprite-1" image (AVIRIS) - about 12000. In this case, the change curves ΔX_i, $E(X_i)$, $D(X_i)$ look more informative for object classification than the correlation ratio curve.

The results of the research into the methods of lossless HSI compression as of 2013 were fixed in the relevant standards [20], [21]. In particular, the values of the lossless compression ratio of the order of 4-5 were stated as very high figures for HSI.

The research performed during the last decade into the role of noise in the compression of HSI, including its role in near-lossless compression and compression with losses, is covered in a series of publications [22], [23], [24]. It was found that: 1) the noise in the hyperspectral images is signal-dependent and is almost uncorrelated spatially; 2) the noise parameters in adjacent channels are generally close, although the overall range of variation of the signal-to-noise ratio over hyperspectral images is very wide; 3) in almost all hyperspectral images the quality of approximately 80-85 percent of the total number of channels is close to ideal, that is, the PSNR is close to or above 35 dB.

In this situation, it would be of interest to explore some universal methods capable of analyzing the signatures of both individual pixels and their integral forms, in terms of classification with respect to object and background or for detection the presence of noise, as the classification of the object and the background or for the presence of noise detection.

2. Description of the empirical mode decomposition method

In our opinion, one can use for the above purposes the method of empirical mode decomposition (EMD). The empirical mode decomposition method was published in 1998 [25], then it was adapted in 2008 for images [26] and is currently applied for a number of image processing tasks [27]. In particular, it can also be applied for hyperspectral images to decompose the signature of an image pixel, similar to a signal, into intrinsic modes having a space/time-varying periodicity. Relying on this property, we want to isolate the highest-frequency part of the signature, which by its nature can represent noise, and also allocate channels (channel groups) that determine individual attributes of the object class specified by the sample pixel.

The decomposition into empirical modes is based on the following assumptions for the signal: 1) the signal has at least two extrema: one maximum and one minimum; 2) the characteristic time scale is determined by the interval between the extrema; 3) if the data is completely devoid of extrema (trend), but may contain inflection points, then the signal can be divided into parts in order to reveal the extrema.

The time interval between successive extrema is taken as the determination of the time scale for the intrinsic oscillation mode, since it not only provides a much higher resolution, but it can also be applied to the data with a non-zero mean (positive or negative values, without zero crossings). In our case, the axis of the channel numbers plays the role of the time axis.

The mode extraction method, called *sifting*, is described as follows [25]. The decomposition method uses envelopes built on local maxima and minima separately. For this purpose, the local extrema of the $f(t)$ signal are identified, and all maxima are interpolated by the cubic spline line as the upper envelope $U(t)$. In the same way, the lower spline envelope $L(t)$ is constructed on the minima. The mean between the upper and lower envelope $R(t)=(U(t)+L(t))/2$ receives the low-frequency residue status and is used for further transformations as $f(t)$, and the difference of the functions $f(t)$ and $R(t)$ receives the status of the first empirical mode $\varphi_1(t)$.

The empirical mode is a function with the following properties: 1) the number of the function extrema on the interval considered differs by no more than one from the number of zero crossings of this function; 2) the half-sum of the upper and lower envelope of the function is close to zero.
As a result of the decomposition of the signal $f(t)$ into empirical modes, we get:

$$f(t) = r(t) + \sum_{i=1}^{N} \varphi_i(t), \quad (1)$$

where $\varphi_i(t)$ are empirical modes, $r(t)$ is the trend residue. The first modes contain high-frequency components of the signal, and the last modes and the residue are low-frequency components.

Fast decomposition algorithm

We use the fast adaptive decomposition method proposed in [26]. Since we are considering a signature, the signal is one-dimensional: it depends only on the channel number. Accordingly, the algorithm is reduced to a one-dimensional version, and it is also simplified in comparison with [26] and [27] while preserving the idea of the original method [25]. Namely, the operations of constructing envelopes (upper and lower), and then, calculation of the points of the mean curve $R(i)$ between them, are replaced by smoothing (averaging) over a symmetric window of width w.

Figure 1. General view of the Moffett Field HIS.

Algorithm steps:

1. Assign the signature of the k-th pixel as the signal $f(i)$,
2. Initialize the size w of the processing window with a value of 3 (w = 3), the number of the empirical mode is q=0.
3. Calculation of low-frequency residue

$$R(i) = \frac{1}{w} \sum_{j=i-w/2}^{i+w/2} f(j), \quad (2)$$

the values corresponding to the positions of the window that go beyond the limits of the signature are taken to be equal to the edge value of the signature.

4. Construct an empirical mode:

$$q = q + 1, \quad \varphi_q(i) = f(i) - R(i) \quad (3)$$

5. Find and enter into the arrays pU and pL all the points of local extrema (minima and maxima) of the current empirical mode $\varphi = \varphi_q$ within the window W of width w with the center in the current channel that satisfy the conditions:

for the array pU $\quad \varphi(i) > \varphi(j), \forall j \in W_w(i)$, (4)

for the array pL $\quad \varphi(i) < \varphi(j), \forall j \in W_w(i)$, (5)

where $W_w(i)$ is a window of width w with the center in the channel i = 1,…,n.

6. If the number of maxima in pU or the number of minima in pL is less than 2, then R cannot be further decomposed and the process terminates.

7. For each local maximum, find the distance $d_{\max}$ along the channel axis to another nearest maximum, for each local minimum, to another minimum $d_{\min}$ and then, take the smallest of them:

$d = \min(d_{\min}, d_{max})$, and update the window size $w = 2\lceil d/2 \rceil + 1$ (6)

8. Set $f(i) = R(i)$ and perform steps 3-7.

Figure 3 shows an example of signal decomposition into empirical modes.

3. Experimental research

The initial data are freely available hyperspectral images obtained with the use of the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). Figure 1 shows the general view of the Moffett

Field HSI in three channels: 29th - R = 638.2nm; 19th - G = 540.6nm; 11th - B = 462.8nm. This sample HSI shows various types of terrain, including water bodies, green vegetation and urban areas. As it was mentioned above, when constructing and analyzing compression algorithms for almost any data, the question arises about the presence of noisy data to which noise was added during the registration process or those data that were distorted as a result of transmission through the communication channels. The volume of such data can be significant and it can affect the resulting compression ratios. In some cases, information noise can be filtered in one way or another by data processing algorithms. Therefore, it is quite important to remove this noise before the compression of the original image. The chosen test sample is stored in the ENvironment for Visualizing Images format (ENVI) and has 2 bytes of information per pixel in each channel. Its width is 752 pixels, and its height is 1924 pixels. In total, the chosen HSI has 224 channels (about 10 nm wide each), which cover the wavelength range from 0.365 μm to 2.497 μm. To investigate the correlation of images of adjacent HSI channels, a linear Pearson correlation coefficient was calculated between the current channel and the next one (223 pairs):

$$r_{XY} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}}. \quad (7)$$

The result of the calculation is shown in Table 1.

Table 1. Distribution of the values of the Pearson correlation coefficient for adjacent HSI channels.

Range	Number of channels	In percentages, %
0,9999-0,99999	32	14.34
0,999-0,9999	91	40.80
0,99-0,999	48	21.52
0,9-0,99	40	17.93
0,3-0,9	12	5.83

As can be seen from Table 1, over 94% of all channels have the correlation coefficient exceeding 0.9, and over 76% of channels have the correlation coefficient exceeding 0.99. This suggests that the neighboring channels are usually very similar to each other, despite their modulation by the influence of the atmosphere. Channels in which the correlation coefficient with the adjacent channel is lower than 0.99, mainly fall within the wavelength ranges (near 1.4 μm, 99-128 channels of our HIS, and 1.9 μm, 153-166 channels) of almost complete absorption of light by a mixture of water vapor and carbon dioxide. Approximately the same high absorption there is in the band above 2.5μm (channels 218-224). In this HSI it is about 50 channels (or 22% of this HSI), the contents of which are almost impossible to use because of the high noise level. If necessary, these channels can be deleted without loss to solve problems of classification of objects of the image.

3.1. An example of a decomposition of a signature into empirical modes and low-frequency residuals. Investigation of the possibilities

Figure 2. The pixel chosen for constructing empirical modes.

As standard for HSI storage is used 2 bytes per pixel per channel. The range of values of this type of data is significantly wider than that of universally used formats for storing visual information. One of an important reason for the introduction of the 2-byte accuracy of HSI is the high accuracy of modern spectrometers, which is achieved through special calibration methods [28]. At the same time, a fairly authoritative source [29] claims that already because of the influence of the atmosphere on the measurements of the spectrometer, it is difficult to talk about achieving accuracy higher than 2-5%. This makes relevant the question of how many bits are needed to store one pixel in the HSI channel to exclude losses for future classification of HSI-objects, and to determine their state.

The application of the method of empirical modes will be considered using the water signature as an example. The pixel of the water sample is marked in red in Figure 2. Figure 4 shows an important for further analysis

part of the result of the decomposition of the water signature into empirical modes and low-frequency residuals (R). The complete decomposition of the signature contains 15 empirical modes and a trend.

Table 2 shows a number of dependencies on the number of the empirical mode (EM) for the first eight modes: 1) for the window size *w*; 2) for the correlation coefficient of the original signature with its corresponding low-frequency residue; 3) for maximum deviation of the empirical mode from zero (Max) in absolute units of a 2-byte representation and in percentage to the maximum value of this signature and the maximum value throughout the HSI. Values in percent are given for comparison with an estimate of the achievable accuracy of HSI in 2-5% given by D.Landgrebe (1999).

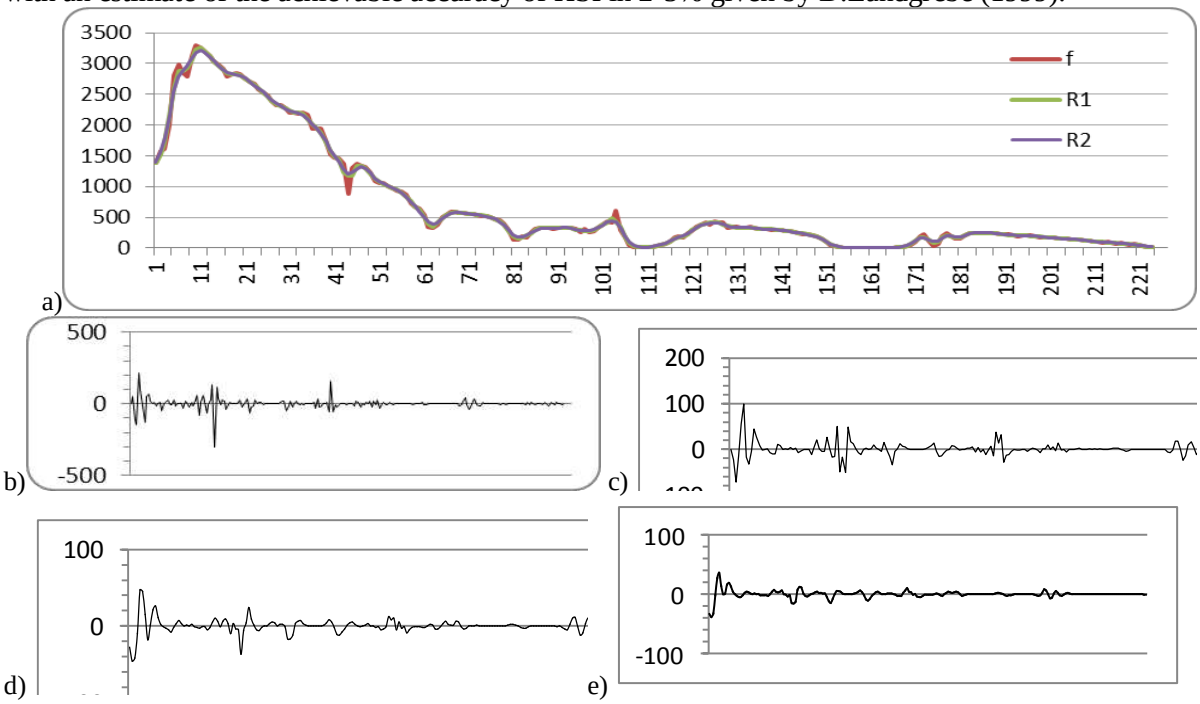


Figure 3. Decomposition of the signature into empirical modes and low-frequency residuals: a – signature (*f*), first (R1) and second (R2) low-frequency residuals; b-e – empirical modes 1-4.

The values of the Pearson correlation coefficient between signature and low-frequency residues are quite high: about from 0.999 to 0.99 (see table 2). Specifically, the significance of these values for the classification of HSI objects will be considered below, but on the whole this indicates that the signature in the correlation estimates can be replaced by at least its first low-frequency residues corresponding to the required accuracy.

Table 2. The change in the window size and in the maximum of the empirical mode (EM) with the its number up to 8 steps, the correlation coefficients of the signature S with its low-frequency residuals R1-R8.

Empirical Mode (EM) Number	1	2	3	4	5	6	7	8
Window size *w*	3	3	3	3	3	3	7	10
Correlation coefficient r_{SR}	0,9991	0,9987	0,9982	0,9979	0,9976	0,9973	0,9950	0,9906
Max of the EM	306	99	47	36	26	21	267	215
In [%] to Max of the signature	9.5	3.1	1.5	1.1	0.8	0.7	8.3	6.7
In [%] to Max of HSI	4.1	1.3	0.6	0.5	0.3	0.3	3.6	2.9

In the first 6 steps of the decomposition, the algorithm automatically selects a minimal blur window with a size of 3 channels. This may indicate the presence of noise, at least in part of the channels of our HSI. There are certain successes in correcting the influence of the atmosphere on the hyperspectral image [30], [31], but for channels where the signatures are close to zero, the correction is practically

useless. From the point of view of the compression of the HSI, it is advantageous to detect noisy channels and remove them from consideration. Let's consider, whether assignment of the minimum size of the window can serve as the detector of noisiness of the channel.

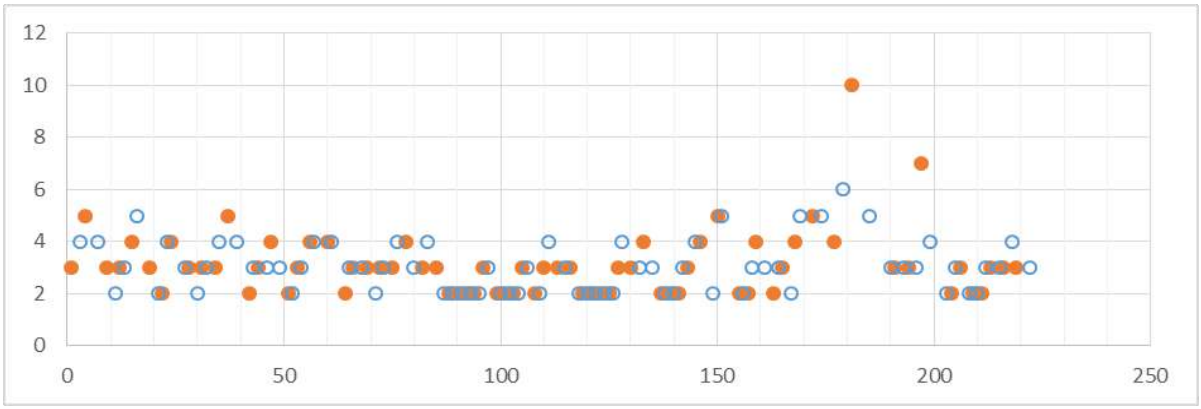


Figure 4. Distance (d) between adjacent extrema of the 1-st empirical mode: maxima (points with fill) and minima (point without fill), via the minimum of which the window size w is determined. The horizontal axis is the channel number.

To do this, consider the distances between the current and next extremums (between two adjacent maxima, dmax, or minima, dmin), by which and according (6) the window size w is determined along the axis of the channel numbers. The zone of noisy channels should give a smaller local window size than the local window size in the "clean" channels zone. The minimum possible value of d, for which w = 3, is 2. Figure 4 shows the distances between extrema for the first empirical mode.

Figure 5. Samples of objects for classifying.

As can be seen from the graph, the presence of the values d = 2 is not limited to intervals 99-128, 153-166, although in these intervals they are the most. The appearance of d = 2 beyond the limits of these intervals correlates well with the windows of the influence of the atmosphere, but do not lead to a completely noisy channel. As a result, via the values of d=2, we can only select the channels of suspects for noise and channels that are substantially modulated by the influence of the atmosphere.

3.2. Classification capabilities of signatures and their low-frequency residues on the basis of values of the correlation coefficient

As is known, each unique object of HSI has an individual signature. The difference between image objects is manifested in the difference in the behavior of signatures in some sequence of channels.

Let us consider the possibilities of classifying objects of the HSI via the signature of the object sample, using for this purpose different values of the Pearson correlation coefficient. Let us also consider how the transition to a 1-byte representation of the signature or the use of its low-frequency residuals instead of the signature will affect the classification capabilities. Let's select several pixels-samples of objects belonging to different objects (Figure 5). Sample 1 represents vegetation (tree crowns); sample 2 - concrete (from the airport runway); sample 3 - soil sample (Soil1) in the airport area; Samples 4 and 5 are water samples (Water1, Water2), but sample 4 looks like a darker one; sample 6 - soil sample (Soil2) from the field. Figure 7 shows the signatures that are constructed as sample averages from signatures that have a correlation coefficient with a sample signature of at least 0.999. We calculate the correlation coefficient of the selected signature-samples with the signatures of all other pixels of the HSI. We shall establish the thresholds to be investigated for classifying an object based

on the Pearson correlation coefficient with the sample signature: {0.9; 0.99; 0.999; 0.9999}. To the desired class we will include all the signatures that have a correlation coefficient with a sample greater than or equal to a given threshold. In addition to the binary classification (object / background) of the 2-byte representation of the HSI, a binary classification was performed on the first and second low-frequency residue, as well as on the signature in a 1-byte representation. Then the difference in the classification results was calculated.

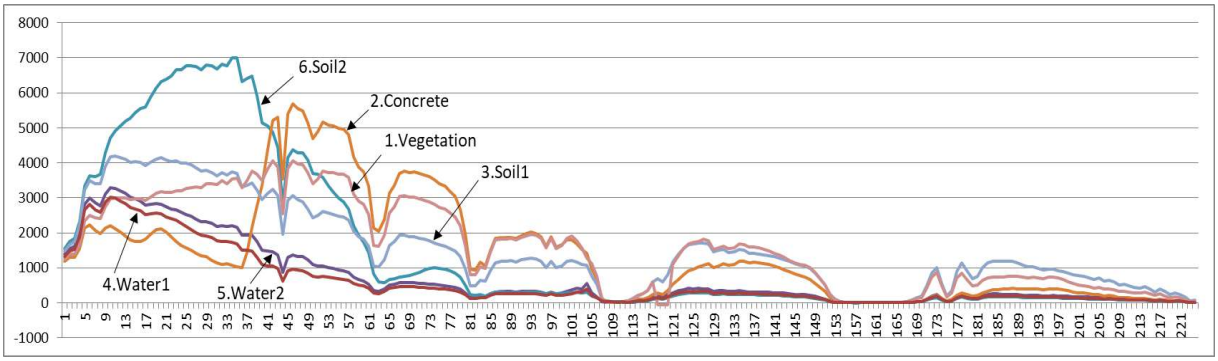
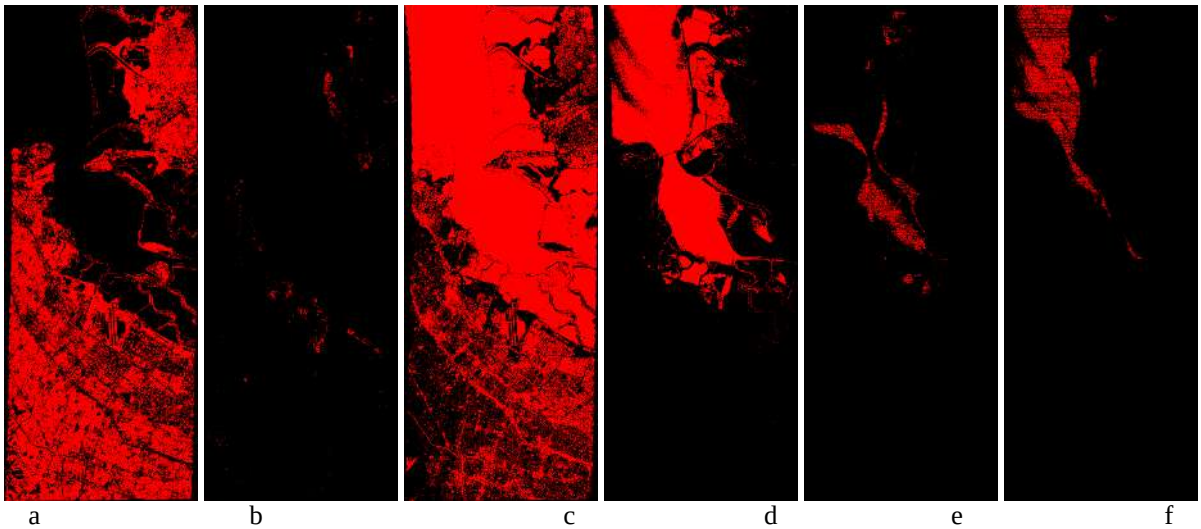


Figure 6. Signatures of samples of objects marked for classifying objects in Figure 5.

Examples of the classification of the initial HSI for different thresholds for the correlation coefficient are shown in Figure 7. The results of using in classifying the first or second low-frequency residue instead of the signature sample are shown in Table 3 (only for the 0.9 threshold). Comparative characteristics of the classification results for all signature samples in a 2-byte and 1-byte is given in Table 4 (only for the 0.999 threshold). These data show that the classification by the value of the correlation coefficient with signature sample can be used as a classification tool. In a number of cases, the threshold value of the correlation coefficient can play a generalizing role, as shown in Figure 7 (a and c). The threshold (T) of the correlation coefficient equal to 0.9 for the signature of sample 1 (Vegetation) made it possible to highlight almost the entire forest park and urban area, excluding building and road surface, and gave a classification mask that is practically complementary to the classification at T = 0.9 for sample 5 (Water2). The threshold T = 0.9 for sample 5 (Water2) captured besides water also the fields adjacent to the coastline, road surfaces, urban buildings. For the threshold T = 0.99 (Figure 7, b), we see in the Vegetation class only a very small number of objects. And at T = 0.999 - only 6 pixels corresponding to the sample. For sample 5, Water2, the threshold T = 0.99 (Figure 7, d) gives already almost the entire surface of the water of the gulf and the flowing rivers, as well as the flood areas adjacent to the shoreline. The threshold T = 0.999 (Figure 7, e, f) already completely separates the objects specified by different samples of water 4 and 5. The value of the threshold at 0.999 in all cases makes it possible to classify objects with signature shapes almost completely similar to the sample signature. The threshold value of 0.9999 guarantees the selection of objects completely identical to the sample or with the similar signature shapes, but allowing deviations in values not exceeding 2-5%. Undoubtedly, the classification problem for the selected 6 samples can be solved on the basis of estimates [16]: $\Delta X_i, E(X_i), D(X_i)$, where i is the channel number. Moreover, in the application of these estimations, which are intervalwise, one, two or three intervals will suffice, in which the signatures are clearly ranked by the value of the estimates.

Let us further consider the results of using, instead of the sample, the signature of its first or second low-frequency residue (see Table 3). Visually at T = 0.9 and T = 0.99, the difference in classification by signature and low-frequency residuals is observed as some smoothing and narrowing along the boundary of objects, increasing with increasing threshold. At T = 0.99, the detection density of the object area can also decrease (loss from 3% density, to Water1, and up to 30% - to Vegetation). At T = 0.999, significant losses occur in the object.



a b c d e f

Figure 7. Examples of classification by object samples at different threshold values (T) of the correlation coefficient: a) sample 1 – Vegetation, T=0.9; b) sample 1 – Vegetation, T=0.99; c) sample 5 – Water2, T=0.9; d) sample 5 – Water2, T=0.99; e) sample 5 – Water2, T=0.999; f) sample 4 – Water1, T=0.999.

Table 3. Classification of objects by the threshold of the correlation coefficient 0.9 for signatures of samples S in a 2-byte representation and their low-frequency residues R1 and R2.

N	Object	Sample coordinates (x;y)	Pixels in Class			Mismatched pixels			
			via S	via R1	via R2	via R1		via R2	
						pcs	%	pcs	%
1	Vegetation	442; 1332	566744	558819	557542	8169	1.4	9846	1.7
2	Concrete	447; 1275	1018521	1014854	1015257	15303	1.5	21956	2.2
3	Soil1	426; 1351	120603	116374	115699	4229	3.5	4904	4.1
4	Water1	209; 470	767970	759931	756722	8061	1.0	11310	1.5
5	Water2	315; 775	873712	864921	861536	8797	1.0	12206	1.4
6	Soil2	540; 630	294865	283321	280976	11544	3.9	13889	4.7

3.3. Ability to switch from a 2-byte representation of signatures to a 1-byte

The possibility of switching from a 2-byte representation of signatures to a 1-byte one was evaluated on the basis of a visual comparison of the results of a binary classification and a quantitative evaluation of the difference of the classification masks. A 1-byte representation is obtained from a 2-byte division of double byte values by 32 with rounding. In all cases of visual comparison, the difference was not noticeable. The difference in masks in all cases remained within 7.5% (the worst case for the Water2 sample, see Table 4). The unmatched pixels in all cases were distributed over the area of the classified object in proportion to the density of the object.

Correlation coefficients between the signatures in the 2-byte and 1-byte representations for the samples: Vegetation, Concrete, Water1, Water2, Soil1, Soil2. Their values were 0.99997, 0.99998, 0.99994, 0.99995, 0.99998, 0.99999 respectively. The correlation coefficient decreases with increasing divisor. For example, for the Water2 sample, when the divisor is changed in the sequence {16, 32, 64}, the correlation coefficient with the 2-byte signature, while remaining high, changes as follows: 0.99998; 0.99995; 0.99980.

This suggests that it is possible to switch to calculations with a 1-byte representation, if necessary, saving the divisor for backwards compatibility with a 2-byte representation.

Table 4. Classification of objects at the correlation threshold of 0.999 for sample signatures in a 2-byte and 1-byte representation.

N	Object	Sample coordinates (x;y)	Pixels in class (2-byte)	Pixels in class (1-byte)	Mismatched pixels	
					pcs	%
1	Vegetation	442; 1332	6	6	0	0
2	Concrete	447; 1275	430	429	7	1.6
3	Soil1	426; 1351	920	904	20	2.2
4	Water1	209; 470	60972	56418	4564	7.5
5	Water2	315; 775	27176	26453	801	2.9
6	Soil2	540; 630	831	826	5	0.6

3.4. Improving the performance of calculations in the HSI classification and evaluating the possibility of the signature-based HSI compression

In order to improve the performance of the classification, an algorithm is implemented on the graphics processor. The working time of the classification algorithm based on the correlation coefficient with the sample signature is defined as the average (for 10 passes) total classification time for all 6 selected signatures. Instead of the sample signature, its low-frequency residue can be used. The measured classification time for the central (CPU) and graphics (GPU) processors is shown in Table 5.

Table 5. Performance achieved on CPU and GPU for HSI classification via value of correlation coefficient.

Hardware	Number of Core	Clock frequency, MHz	Time (sec)
CPU: AMD Phenom II X4 955	4	2500	257. 07
GPU: NVIDIA GeForce GTX 970	1664 (CUDA)	1152	0.31

From Table 5, it can be seen that the use of a graphics processor makes it possible to carry out the classification procedure 836 times (3 orders of magnitude) faster than on a central processor. In this case, we can talk about the operation of the algorithm in real time, which makes its use in desktop data processing applications quite realistic. The performance for the same CPU of the empirical mode decomposition was 0.91 seconds per thousand signatures.

In the interest of evaluating the capabilities of signature-based compression, an attempt is made to increase the number of zero or close to zero values in a compressible array by applying wavelet decomposition. Based on the length of the signature array, 7 steps of the Haar transformation are used for the signature wavelet decomposition of the HSI in the 1-byte representations. After that, the source 2-byte HSI and the result of the Haar transformation were archived by ZIP (without losses) with the maximum compression ratio (Table 6). The compression ratio for 1-byte signatures was 6.65. Since we replaced the double-byte representation with a 1-byte representation, we will treat this compression as nearlossless.

Table 6. Compression of the entire HSI data array.

File size, MB	Without wavelet decomposition	1-byte signature with wavelet decomposition
Source file	648,18	324,09
Archived file	421,12	97,48

4. Conclusion

The signature of hyperspectral image (HSI) pixels and their decomposition into empirical modes (EM) and low-frequency residuals are investigated. On the basis of estimates related to the EM-decomposition method, the possibility of switching from a 2-byte representation of the values of the HIS-signature to a 1-byte one is examined using the example of the Moffett Field from the AVIRIS spectrometer.

It is revealed that: 1) the localization of the minimum window sizes for the first EM is correlated with the localization of the significant influence of the atmosphere; 2) the first low-frequency residues have a fairly high correlation coefficient with the signature; 3) the greatest interest for the

decomposition of the signature is represented by one or two first EM and the corresponding low-frequency residue; 4) the 1st and 2nd modes on the significant part of the channel axis are close to zero and can be reduced to a 1-byte representation practically without loss of accuracy; 5) 50 of the 224 HIS-channels are noisy and can be excluded from consideration.

The management of the classification capabilities of signatures by changing the threshold value of the correlation coefficient with the sample, as well as the application of the 1st and 2nd low-frequency residues in place of the signature, was studied. Classification capabilities of signatures in a 1-byte representation are almost equivalent to a 2-byte one, which makes it possible to put a signature with 1-byte representation (as eight senior digits) as the object of compression.

The classification procedure is implemented on the GPU, which accelerated its execution more than 800 times, to fractions of a second.

Classification according to the samples using [16] on the basis of estimates ΔX_i, $E(X_i)$, $D(X_i)$, where i is the channel number, can additionally reduce the number of channels necessary for classification. If the estimates are applied in a series of channel intervals, one or three intervals may be sufficient, in which the signatures are clearly ranked by the values of the estimates.

For the wavelet decomposition of the HSI data array, in combination with a 1-byte representation, a nearlossless compression ratio of 6.65 is obtained.

For the future works it is interesting to investigate also the approaches based on the machine learning methods like publications [19],[33],[34],[35].

5. References

[1] Wu X and Memon N 1997 Context-based, adaptive, lossless image coding *IEEE Trans. Commun. Apr.* **45** 437-444

[2] Weinberger M J, Seroussi G and Sapiro G 2000 The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS *IEEE Trans. Image Processing* **9** 1309-1324

[3] Wu X and Memon N 2000 Context-based lossless interband compression - Extending CALIC *IEEE Trans. Image Processing* **9** 994-1001

[4] Magli E, Olmo G and Quacchio E 2004 Optimized onboard lossless and near-lossless compression of hyperspectral data using CALIC *IEEE Geoscience and Remote Sensing Letters* **1(1)** 21-25

[5] Wu X, Memon N and Sayood K 1995 A Context-Based, Adaptive, Lossless/Nearly-Lossless Coding Scheme for Continuous-Tone Images *ISO/IEC JTC 1/SC 29/WC 1* **202**

[6] Kozoderov V V, Kondranin T V, Kazancev O, Bobylev V I, Scherbakov M V, Borzyak V V, Dmitriev E V, Egorov V D, Kamencev V P, Belyakov A Yu and Loginov S B 2009 Processing and interpretation of hyperspectral aerospace measurement data for remote diagnostics of natural and technogenic objects *Earth Research from Space* **2** 36-54

[7] Chaban L N, Vecheruk G V and T S 2009 Gavrilova. Investigation of the possibilities of land cover classification using hyperspectral imagery in thematic processing packages of remote sensing data *Proceedings of MIPT* **1(3)** 171-180

[8] Chaban L N, Vecheruk G V, Kondranin T V, Kudryavtsev S V and Nikolenko A A 2012 Modeling and thematic processing of images identical to video data of ERS hyperspectral equipment under development *Current Problems of Earth Remote Sensing from Space* **9(2)** 111-121

[9] Popov M A and Stankevich S A 2006 Methods for optimizing the number of spectral channels in problems of ERS data processing and analysis *Contemporary Problems of Earth Remote Sensing from Space* **1** 106-112

[10] Jiménez L and Landgrebe D A 1999 Supervised Classification in High Dimensional Space: Geometrical, Statistical, and Asymptotical Properties of Multivariate Data *IEEE Transactions on Systems, Man, And Cybernetics-Part C: Applications and Reviews* **28(1)** 39-53

[11] Arzuaga-Cruz E, Jimenez-Rodriguez L O and Velez-Reyes M 2003 Unsupervised Feature Extraction and Band Subset Selection Techniques Based on Relative Entropy Criteria for Hyperspectral Data Analysis *Proc. SPIE* **5093** 462-473
[12] Lee C and Langrebe D A 1993 Feature Extraction Based on Decision Boundaries *IEEE Trans. on Pattern Analysis and Machine Intelligence* **4(15)** 388-400
[13] Myasnikov E V 2017 Hyperspectral image segmentation using dimensionality reduction and classical segmentation approaches *Computer Optics* **41(4)** 564-572 DOI: 10.18287/2412-6179-2017-41-4-564-572
[14] Nakariyakul S and Casasent D 2004 Hyperspectral feature selection and fusion for detection of chicken skin tumors *Proc. SPIE* **5271** 128-139 DOI: 10.1117/12.517443
[15] Zamyatin A V and Sarinova A Zh 2013 An algorithm for compressing hyperspectral aerospace images with the account of inter-band correlation *Applied Informatics* **5(47)** 35-42
[16] Gashnikov M V and Glumov N I 2014 Hierarchical compression for hyperspectral image storage *Computer Optics* **38(3)** 482-488
[17] Gashnikov M V and Glumov N I 2016 Onboard processing of hyperspectral data in the remote sensing systems based on hierarchical compression *Computer Optics* **40(4)** 543-551 DOI: 10.18287/2412-6179-2016-40-4-543-551
[18] Gashnikov M V, Glumov N I, Kuznetsov A V, Mitekin V A, Myasnikov V V and Sergeev V V 2016 Hyperspectral remote sensing data compression and protection *Computer Optics* **40(5)** 689-712 DOI: 10.18287/2412-6179-2016-40-5-689-712
[19] Vorobiova N S, Sergeyev V V and Chernov AV 2016 Information technology of early crop identification by using satellite images *Computer Optics* **40(6)** 929-938 DOI: 10.18287/2412-6179-2016-40-6-929-938
[20] Auge E, S´anchez J E, Kiely A, Blanes I and Serra-Sagrista J 2013 Performance impact of parameter tuning on the CCSDS-123 lossless multi- and hyperspectral image compression standard *Journal of Applied Remote Sensing* **7(1)** 16
[21] Blanes I, Magli E and Serra-Sagrisat J 2014 A tutorial on Image Compression on Optical Space Imaging Systems *IEEE Geoscience and Remote Sensing Magazine* **2(3)** 8-26
[22] Abramov S K and Lukin V V 2015 Problems of automating the processing of hyperspectral remote sensing images *Aerospace Engineering and Technology* **6(123)** 101-110
[23] Abramova V V, Abramov S K and Lukin V V 2015 Multistage Iterative Method for Blind Evaluation of Mixed Noise Characteristics on Images *ITS* **6(1)** 8-14
[24] Lukin V, Abramov S, Ponomarenko N, Uss M, Zriakhov M, Vozel B, Chehdi K and Astola J 2011 Methods and automatic procedures for processing images based on blind evaluation of noise type and characteristics *SPIE Journal of Applied Remote Sensing* **5(1)** 27 DOI: 10.1117/1.3539768
[25] Huang N E, Shen Z, Long S, Wu M C, Shih H H, Zheng Q, Yen N-C, Tung C C and Liu H H 1998 The Empirical Mode Decomposition and Hilbert Spectrum for nonlinear and non-stationary time series analysis *Proc. R. Soc. London A* **454** 903-995
[26] Bhuiyan S M A, Adhami R R and Khan J F 2008 A novel approach of fast and adaptive bidimensional empirical mode decomposition *IEEE International Conference on Acoustics, Speech and Signal Processing* 1313-1316
[27] Guryanov F and Krylov A 2017 Fast medical image registration using bidirectional empirical mode decomposition *Signal Processing: Image Communication* 1-6
[28] Introduction to Hyperspectral Imaging: Tutorial *MicroImages Inc.* (Access mode: http://www.microimages.com/documentation/Tutorials/hyprspec.pdf)
[29] Podlipnov V V and Skidanov R V 2017 Calibration of an imaging hyperspectrometer *Computer Optics* **41(6)** 869-874
[30] Landgrebe D 1999 Information Extraction Principles and Methods for Multispectral and Hyperspectral Image Data *Information Processing for Remote Sensing* 3-38
[31] Denisova A Y and Myasnikov V V 2016 Atmospheric correction of hyperspectral images using small volume of the verified data *Computer Optics* **40(4)** 526-534 DOI: 10.18287/2412-6179-2016-40-1-526-534

[32] Denisova A Y, Juravel Y N and Myasnikov V V 2016 Estimation of parameters of a linear spectral mixture for hyperspectral images with atmospheric distortions *Computer Optics* **40(3)** 380-387 DOI: 10.18287/2412-6179-2016-40-3-380-387
[33] Sirota A A and Dryuchenko M A 2015 Generalized image compression algorithms for arbitrarily-shaped fragments and their implementation using artificial neural networks *Computer Optics* **39(5)** 751-761 DOI: 10.18287/0134-2452-2015-39-5-751-761
[34] Sergeev V V and Yuzkiv R R 2016 A parametric model for the autocorrelation function of space hyperspectral data *Computer Optics* **40(3)** 416-421 DOI: 10.18287/0134-2452-2016-40-3-416-421
[35] Kopenkov V N and Myasnikov V V 2016 Development of an algorithm for automatic construction of a computational procedure of local image processing, based on the hierarchical regression *Computer Optics* **40(5)** 713-720 DOI: 10.18287/2412-6179-2016-40-5-713-720

Acknowledgements
This work was supported by a grant from the Russian Science Foundation (project No. 16-11-00068).

Interpolation for differential and hierarchical compression of multidimensional signals

S A Denisov[1] and M V Gashnikov[1]

[1]Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. A comparative research of interpolation algorithms for hierarchical and differential methods of signal compression is performed. The hierarchical method of signal compression is based on hierarchical grid interpolation. The differential method of signal compression is based on differential pulse-code modulation (DPCM). Versions with maximum error control are used for both methods. Computational experiments are performed in natural test signals. The dependence of the compression ratio on the maximum error and on the standard error is given for both methods.

1. Introduction

To date, digital systems contain a huge amount of multidimensional information. First of all, these are hyperspectral data [1-2], the results of remote sensing [3], as well as video, charts, diagrams, drawings, photographs, etc. Such a variety of information requires a huge size of space to store this information on servers. Accordingly, to reduce the size of stored signals, effective algorithms for signal compression are required.

There is large number of algorithms for signal compression [4-11]. The compression method JPEG [8] is currently the most common method for signal compression. This method is based on a discrete cosine transformation [9]. The JPEG-2000 method [10] is more effective [11]. This method is based on wavelet. But this method is much less common.

Both of these methods are based on transformation the signal into a certain space of coefficients. This allows you to achieve high compression ratios. However, error control in the transformed space is difficult. If it is necessary to control the error, other methods of signal compression are promising. First of all, these are differential [4-5] and hierarchical [12-13] methods of signal compression. These methods allow you to control the maximum error [14].

In addition, the computational complexity of these methods is much less, since there is no transformation into a spectral space. These methods also have many other advantages (noise immunity, buffer memory management in real-time systems, etc.). Therefore, the task of investigating such compression methods is topical.

Differential and hierarchical compression methods decorrelate a signal in the same way. Transformation to a difference representation of the signal occurs. The difference signal is calculated as the difference between the original and interpolated samples of the signal. The difference between differential and hierarchical compression methods lies in interpolation algorithm. In differential methods, a progressive scan is used, and the signal samples are interpolated on the basis of previous samples. In hierarchical methods, the signal is interpolated based on more resampled versions of the same signal. The question of which interpolation algorithm is more effective requires research.

In this paper, computational experiments are performed comparing differential and hierarchical methods of signal compression with each other. The research is performed in the ordinates "error/compression ratio". The compression method of JPEG is used as a basis for comparison.

2. Differential compression of signals

Differential compression of signals is also called differential pulse-code modulation (DPCM) [4-5]. Here is a brief description of this compression method (see Figure 1). The multidimensional signal $\mathbf{X} = \left\{ x\left(\vec{c}\right) \right\}$ is processed in the order of some scan ($\vec{c}$ is the vector of the signal's arguments).

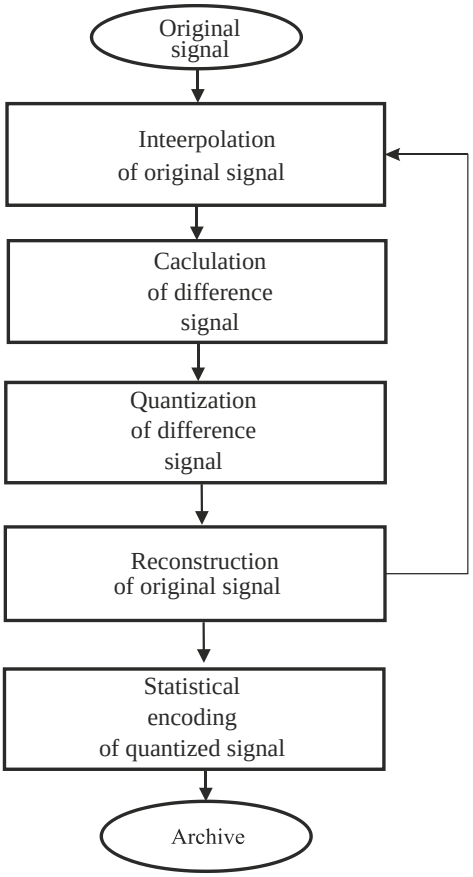


Figure 1. Differential compression of signal.

First, for each sample, an interpolated value is calculated

$$\hat{x}\left(\vec{c}\right) = I\left(\left\{ \bar{x}\left(\vec{k}\right) \right\}\right), \tag{1}$$

where $I(..)$ is some interpolator, and $\left\{ \bar{x}\left(\vec{k}\right) \right\}$ are the previous (already processed) signal samples. These samples have already been compressed and decompressed.

After interpolation, the sample of difference signal is calculated:

$$f(\vec{c}) = x(\vec{c}) - \hat{x}\left(\vec{c}\right). \tag{2}$$

If the original signal has a high correlation, the variance of the difference signal is much smaller than the variance of the original signal. This makes it possible to compress the difference signal much more than the original signal. To increase the compression ratio, the difference signal is quantized using the quantization function $Q(..)$:

$$q\left(\vec{c}\right) = Q\left(f\left(\vec{c}\right)\right). \tag{3}$$

The entropy [4] of the quantized signal $q\left(\vec{c}\right)$ is much smaller than the entropy of the difference signal $f\left(\vec{c}\right)$, so the quantized signal is compressed much more. In this paper, a uniform scale is used for quantization. In this situation, the actual quantizer can be written in the form:

$$q = Q(f) = sign\,(f)\left[\frac{e_{\max} + |f|}{2e_{\max} + 1}\right], \tag{4}$$

where [..] denotes the integer part of a number, and $sign(f)$ is calculated as follows:

$$sign(f) = \begin{cases} 1, & если\ f > 0 \\ 0, & если\ f = 0 \\ -1, & если\ f < 0 \end{cases}$$

The use of this quantizer allows us to control the maximum error e_{max} [5] for each signal sample:

$$e_{\max} = \max_{\vec{c}} \left| f\left(\vec{c}\right)\right| = \max_{\vec{c}} \left| x\left(\vec{c}\right) - \overline{x}\left(\vec{c}\right)\right|. \tag{5}$$

The quantized signal $q\left(\vec{c}\right)$ is processed by a statistical encoder and stored in the archive.

After the quantization, the signal sample is restored:

$$\overline{x}\left(\vec{c}\right) = \widehat{x}\left(\vec{c}\right) + \left(1 + 2e_{\max}\right) q(\vec{c}), \tag{6}$$

where $\overline{x}\left(\vec{c}\right)$ are the reconstructed signal samples. Exactly the same values will be obtained after decompression. With compression, these values are needed for interpolation (1) of subsequent samples. The use of reconstructed signal samples rather than the original signal samples for interpolation (1) makes it possible to ensure that the interpolated values are the same at the stages of compression and decompression.

3. Hierarchical compression of signals

Hierarchical compression of signals is based on a special hierarchical representation of the signal [12-13]. This representation of the signal is similar to a quad tree. A multidimensional signal $\mathbf{X} = \left\{x\left(\vec{c}\right)\right\}$ is represented as a set of hierarchical levels $\mathbf{X}_l$ (see also Figures 2-3):

$$\mathbf{X} = \bigcup_{l=0}^{L-1} \mathbf{X}_l,$$

$$\mathbf{X}_{L-1} = \left\{ x_{L-1}\left(\vec{c}\right)\right\},$$

$$\mathbf{X}_l = \left\{ x_l\left(\vec{c}\right)\right\} \setminus \left\{ x_{l+1}\left(\vec{c}\right)\right\},\ l < L-1,$$

where L is the number of hierarchical levels $\mathbf{X}_l$, $\{x_l\left(\vec{c}\right)\}$ is the signal resampled with step 2^l:

$$x_l\left(\vec{c}\right) = x\left(2^l \vec{c}\right). \tag{7}$$

This hierarchical representation of the signal makes possible a sequential compression of the hierarchical levels. We compress the hierarchical levels in the following sequence: $\mathbf{X}_{L-1}$, $\mathbf{X}_{L-2}$, .., $\mathbf{X}_2$, $\mathbf{X}_1$, $\mathbf{X}_0$. The hierarchical level $\mathbf{X}_{L-1}$ is stored in the archive without any changes, since the data size of this level is very small. The samples of all other hierarchical levels are interpolated on the basis of samples of more resampled hierarchical levels:

$$\widehat{x}_l(\vec{c}) = I\left(\bigcup_{k=l+1}^{L-1} \overline{X}_k\right), \tag{8}$$

where $\hat{x}_l(\vec{c})$ are interpolating values, $I(..)$ is interpolation function, and $\overline{X}_k$ are the hierarchical levels that have already been compressed and decompressed.

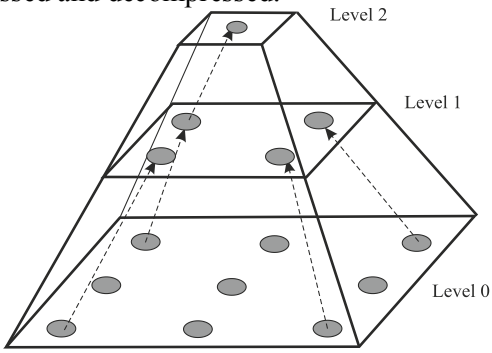


Figure 2. Hierarchical representation of a two-dimensional signal.

After the interpolation, the difference signal is calculated:

$$f_l\left(\vec{c}\right) = x_l\left(\vec{c}\right) - \hat{x}_l\left(\vec{c}\right). \tag{9}$$

Then this difference signal is quantized by means of some quantization function $Q(..)$:

$$q_l(\vec{c}) = Q\left(f_l(\vec{c})\right). \tag{10}$$

The quantized signal $q_l\left(\vec{c}\right)$ is processed by a statistical encoder and stored in an archive. The quantizer (4) with uniform scale is used in this paper, which allows us to control (5) the maximum error $e_{\max}$ for each signal sample.

After quantization, the signal sample recovery is performed:

$$\overline{x}_l\left(\vec{c}\right) = \hat{x}_l\left(\vec{c}\right) + \left(1 + 2e_{\max}\right) q_l(\vec{c}), \tag{11}$$

where $\overline{x}\left(\vec{c}\right)$ are the reconstructed signal samples. These samples are subsequently used to interpolate (8) the samples of less resampled hierarchical levels of the signal.

3		1		2		1		3
1		1		1		1		1
2		1		2		1		2
1		1		1		1		1
3		1		2		1		3

Figure 3. Samples numbers of hierarchical levels of a two-dimensional signal (empty cells correspond to the hierarchical level number zero).

4. Interpolation of digital signals during compression

With differential and hierarchical compression, the already processed samples are used for interpolation. To reduce computational complexity, simple averaging functions are used in this case [4-5]. To simplify the explanation, we describe interpolation functions for the case of a two-dimensional signal $\mathbf{X} = \left\{x\left(\vec{c}\right)\right\} = \left\{x\left(m,n\right)\right\}$.

During differential compression, the following interpolator is used in this work [15-17]:

$$\hat{x}\left(m,n\right) = \frac{1}{2}\overline{x}(m-1,n) - \frac{1}{2}\overline{x}(m-1,n-1) + \overline{x}(m,n-1).$$

With hierarchical compression, the following expression was used to interpolate samples with indices of the form $(2m+1,2n+1)$:

$$\hat{x}_l(2m+1,2n+1)=\left[\frac{1}{4}\left(\bar{x}_{l+1}(m,n)+\bar{x}_{l+1}(m+1,n)+\bar{x}_{l+1}(m,n+1)+\bar{x}_{l+1}(m+1,n+1)+1\right)\right], \quad (12)$$

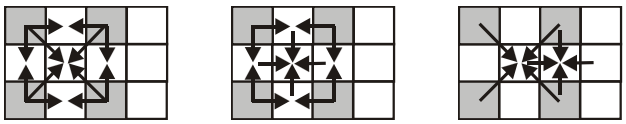
where [..] denotes the integer part of the number, $\{x_l(m,n)\}$ is a signal resampled in 2^l times:

$$x_l(m,n)=x\left(2^l m,2^l n\right). \quad (13)$$

Samples with indices $x_l(2m+1,2n)$ and $x_l(2m,2n+1)$ are interpolated as follows:

$$\hat{x}_l(2m+1,2n)=\left[\frac{1}{2}\left(\bar{x}_{l+1}(m,n)+\bar{x}_{l+1}(m+1,n)\right)\right]. \quad (14)$$

The interpolator (13-14) is shown in Figure 4a. Next to Figures 4b-4c two more hierarchical interpolators are shown in the same notation.



(a) (b) (c)

Figure 4. Interpolators with hierarchical compression.

The described averaging interpolators for hierarchical and differential compression are very simple and have low computational complexity.

5. Experimental research of interpolation algorithms of signals

The considered interpolators were investigated within the framework of the appropriate methods of signal compression. For this research, computational experiments were performed. Nature test signals were used in computational experiments. Some examples of decompressed test signals are shown in Figures 5-6.

(a) (b) (c)

Figure 5. Decompressed signals for differential compression (the maximum error e_{max} is 2, 10, 20).

(a) (b) (c)

Figure 6. Decompressed signals for hierarchical compression (the maximum error e_{max} is 2, 10, 20).

The algorithms were compared in the coordinates "error/compression ratio". As a basis for comparison, the JPEG method was used. In this case, the maximum (4) error e_{max} was used as an error

between the original and decompressed signals. We also used the root mean square (RMS) error as a measure of this error:

$$e_{RMS} = \sqrt{\frac{1}{S}\sum_{\vec{c}}\left(x\left(\vec{c}\right) - \overline{x}\left(\vec{c}\right)\right)^2}, \tag{15}$$

where S is the number of signal samples, $x\left(\vec{c}\right)$ is the original signal, $\overline{x}\left(\vec{c}\right)$ is the decompressed signal.

Typical results of computational experiments are shown in Figures 7-8. Based on the results of computational experiments, the following conclusions were made.

1. Compression ratio K of hierarchical compression and differential compression is greater than the compression ratio of the JPEG method with small RMS error. With large RMS error, the compression ratio of the JPEG method is better.
2. The maximum error e_{max} hierarchical compression is less than the maximum error of the differential compression and the maximum error of the JPEG method.
3. RMS error e_{RMS} of hierarchical compression is less than the RMS error of differential compression.

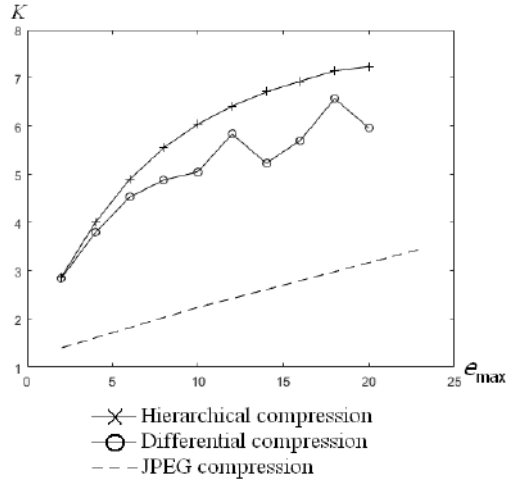


Figure 7. Dependences of the compression ratio K on the maximum error e_{max}.

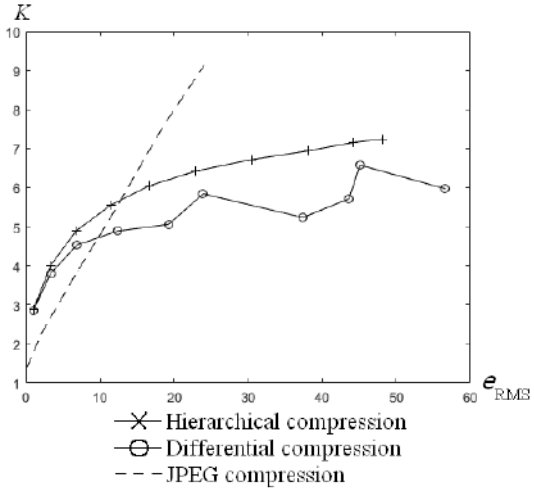


Figure 8. Dependences of the compression ratio K on the RMS error e_{RMS}.

6. Conclusion

Differential and hierarchical methods of multidimensional signals compression are considered. Algorithms of signals interpolation for differential and hierarchical signal compression are considered. Algorithms of interpolation and signal compression are implemented as software. Computational experiments were performed to investigate signal interpolation algorithms within the framework of

signal compression methods. Computational experiments were performed comparing hierarchical and differential methods of signal compression on a set of natural test signals. As a basis for comparison, the JPEG compression method was used. The results of the computational experiments are shown in the coordinates "error/compression ratio". As a measure of error, RMS and maximum errors are used. The conditions under which hierarchical compression has an advantage over differential compression and the JPEG method are described.

7. References

[1] Chang C 2013 *Hyperspectral Data Processing: Algorithm Design and Analysis* (New York: Wiley Press) p 1164
[2] Chang C 2007 *Hyperspectral data exploitation: theory and applications* (New York: Wiley-Interscience) p 440
[3] Borengasser M 2004 *Hyperspectral Remote Sensing – Principles and Applications* (CRC Press) p 128
[4] Sayood K 2012 *Introduction to Data Compression* (The Morgan Kaufmann Series in Multimedia Information and Systems) p 743
[5] Salomon D 2007 *Data Compression. The Complete Reference* (Springer-Verlag) p 1118
[6] Gupta V, Sharma A and Kumar A 2014 Enhanced Image Compression Using Wavelets *International Journal of Research in Engineering and Science (IJRES)* **2** 55-62
[7] Woon W M, Ho A T S, Yu T, Tam S C, Tan S C and Yap L T 2000 Achieving high data compression of self-similar satellite images using fractal *Proc. of IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS)* 609-611
[8] Wallace G 1991 The JPEG Still Picture Compression Standard *Communications of the ACM* **34(4)** 30-44
[9] Plonka G and Tasche M 2005 Fast and numerically stable algorithms for discrete cosine transforms *Linear Algebra and its Applications* **394(1)** 309-345
[10] Li J 2003 Image Compression: The Mathematics of JPEG-2000 *Modern Signal Processing* **46** 185-221
[11] Ebrahimi F, Chamik M and Winkler S 2004 JPEG vs. JPEG2000: An Objective Comparison of Image Encoding Quality *Proc. of SPIE Applications of Digital Image Processing XXVII* **5558** 300-308
[12] Gashnikov M V, Glumov N I and Sergeyev V V Compression Method for Real-Time Systems of Remote Sensing *Proc. 15th Int. Conf. on Pattern Recognition* **3** 232-235
[13] Gashnikov M V 2017 Minimizing the entropy of post-interpolation residuals for image compression based on hierarchical grid interpolation *Computer Optics* **41(2)** 266-275 DOI: 10.18287/2412-6179-2017-41-2-266-275
[14] Lin S 2004 *Error Control Coding: Fundamentals and Applications, second edition* (New Jersey: Prentice-Hall) p 1260
[15] Efimov V M and Kolesnikov A N 1997 Effectiveness estimation of the hierarchical and line-by-line lossless compression algorithms *Proc. of the III conf. Pattern Recognition and Image Analisys* **1** 157-161
[16] Gashnikov M V 2016 Interpolation for hyperspectral images compression *CEUR Workshop Proc.* **1638** 327-333
[17] Gashnikov M V and Glumov N I 2015 Hyperspectral images repository using a hierarchical compression *Posters proc. of 23 Int. Conf. on Computer Graphics, Visualization and Computer Vision* 1-4

Acknowledgments

This paper was funded by RFBR according to the research projects 18-01-00667, 18-07-01312.

Building extraction from satellite imagery using a digital surface model

A V Dunaeva[1,2] and F A Kornilov[1]

[1]N.N. Krasovskii Institute of Mathematics and Mechanics of the Ural Branch of the Russian Academy of Sciences, S. Kovalevskaya Street 16, Yekaterinburg, Russia, 620990
[2]Ural Federal University named after the first President of Russia B.N. Yeltsin, Mira Street 19, Yekaterinburg, Russia, 620002

Abstract. In this paper, two approaches to building extraction from satellite imagery and height data obtained from stereo images or LIDAR are compared. The first approach consists of detecting high-rise objects in a digital surface model and then improving recognition accuracy using segmentation of spectral information. The second approach uses the U-Net convolutional neural network, which showed the best results for the extraction of objects from aerospace images on a number of large datasets. Extensive experiments were carried out to evaluate the dependence of the quality of U-Net-based building extraction on the different data types (including high-resolution satellite images and digital surface model data). Building extraction quality of the trained network was also evaluated on satellite images with different spatial resolutions.

1. Introduction

At the present time, monitoring the state of the Earth's surface based on aerospace images plays an important role in many fields such as agriculture, construction, emergency analysis and others. This approach has a number of undeniable advantages: timeliness, large coverage of areas and a relatively low cost. At the same time, many tasks are handled manually by an operator, which makes it important to develop tools for automatically retrieving required information from satellite images.

The extraction and classification of terrain objects are the most important tasks in processing satellite imagery. The solutions of such problems find their application in cartography (update of topographic maps), the detection of changes in the composition of the terrain objects (i.e. analysis of urban development and illegal deforestation), navigation of aircrafts and others. However, the methods existing at the moment extract and classify objects of a terrain from space images with insufficient accuracy. First of all, this is due to the complexity and variability of the scenes under consideration, which contain a huge number of objects of different nature significantly influenced by shooting conditions. All this contributes to the importance of research aimed at overcoming these difficulties and developing software systems for processing remote sensing data.

In this paper, in order to improve the quality of the building extraction, a digital surface model (DSM) containing heights of a terrain and objects on it was applied. Height information for the DSM can be obtained by shooting a terrain from two different angles. Two such images form a stereo pair, which is passed to the input of a stereo matching algorithm, for calculating a matrix of pixel shifts (a disparity map) belonging to the same objects between two images. After that, the digital surface model of the terrain is constructed by calculating the heights of the terrain and objects on the basis of

the disparity map and the parameters of shooting. In this paper, we used a digital surface model obtained as a result of the operation of the stereo matching algorithm proposed in [1]. Since the main interest was the urban development, panchromatic images with a resolution of 0.5 m were used for stereo matching, and the resolution of the final model was 1 m per pixel. Among the works devoted to the generation of DSMs, it is also worth mentioning work [2] which describes a fast algorithm for GPUs.

The paper is organized as follows. The second section describes an approach to building extraction, based on the detection of high-rise objects of a terrain from a DSM and the refinement of their boundaries on the basis of segmentation of spectral information. The third section is devoted to the U-Net [3] convolutional neural network and details of its training. The fourth section presents the results of the comparison of considered approaches using real data.

2. Image segmentation and detection of high-rise objects

As a starting point for building extraction, we consider a segmentation algorithm, which uses heights of objects in addition to the spectral information to provide a better segmentation of satellite images.

2.1. Satellite image segmentation

Segmentation is used to pre-process images in many computer vision tasks, since it allows analyzing homogeneous (by certain criteria) regions instead of separate pixels of an image. The characteristics used to split an image can be very different: the color, the texture, etc. To date, a large number of segmentation algorithms have been developed. In [4, 5], a description and comparison of some of them is given to address the problem of building extraction. The considered algorithms demonstrate high quality segmentation, but do not use height information that can permit a better detection of object boundaries in satellite images.

To test the influence of height information on the result of satellite image segmentation, an algorithm based on the idea of grouping pixels in regions on the basis of the similarity of their brightnesses and heights was presented in [5]. The algorithm receives a four-channel image and a digital surface model. The following sequence of steps is performed.

1. Median filtering of the image.

2. Preliminary selection of a region. For each pixel in the image we form a region by recursive search for neighboring pixels of similar height and color. After that, we calculate mode values of the height and each of the four image channels in the obtained region and find in it a pixel having a height exactly equal to the calculated mode of height and the closest brightness to the modes of each channel.

3. Selection of the region. Next, from the found pixel, we form a new region by repeating recursive search, and fill the new region with the color of the computed mode values.

4. Repeat steps 2-3. Due to the noisiness of the input data, the regions obtained in the previous steps will have a small area. The repetition of the segmentation procedure will allow merging the regions into larger ones.

5. Small region removal. As a result of the brightness or height discontinuity, regions with a negligible area can remain. Such regions are combined with neighboring ones in such a way that the region to which the pixels are added has a larger area, and its height and brightness are as close to the added pixels as possible. This procedure is performed for regions in order of increasing area.

2.2. Building extraction from a DSM

After segmenting the satellite image (using the algorithm mentioned above or another one), it is required to determine whether the obtained segments belong to buildings.

To do this, it is necessary to find objects in the DSM that rise above the terrain. Such objects include buildings and forest tracts. For this purpose, the following algorithm is proposed.

1. *Search for height differences*. For each pixel p in the DSM with the height value h, we consider its one-dimensional neighborhoods $O_d^{\rightarrow}(p)$ and $O_d^{\downarrow}(p)$ of the radius d in horizontal and vertical directions. For the pixels $p_i \in O_d(p)$, we calculate $\Delta h_i = h_i - h_{i+1}, i \in [-d, d-1]$, where the index 0 corresponds to the pixel p, which is the center of the neighborhood. And if $\Delta h_0 = \max_i \Delta h_i$ or

$\Delta h_0 = \min_i \Delta h_i$, and $|\Delta h_0| \geq T$, then the pixel p is marked as having a significant height difference. The threshold T specifies the minimum height of buildings to be detected. We do not consider the pixels at which the minimum and maximum are simultaneously reached, since they are the errors in the DSM.

2. *Obtaining high regions*. From the pixels that have a significant height difference, we construct line segments, which will form the target regions. For the construction of these line segments, we search horizontally and vertically in forward and backward directions in the DSM for pixels with a height difference upward. Then we begin to draw a line segment from each selected pixel that goes:

- to the pixel with the height difference upward, the value of which is added to the value of the difference at the starting pixel. Then the construction of the line segment continues;
- to the pixel at which the height value is less or close to the height value at the beginning of the line segment;
- until the line segment reaches the fixed length.

Performing this procedure in the forward and backward directions (from left to right and vice versa, from top to bottom and vice versa) allows us to find even those regions of buildings in which one of the sides can be blurred due to the influence of noise.

Next, the pixels of the segments selected only horizontally or vertically are discarded. The remaining pixels of the segments form regions. We select only those of them that contain at least one pixel with a significant height difference. This operation allows rejecting the false intersections of the vertical and horizontal segments.

3. *Refinement of the form of the found regions*. We consider simply connected regions in the DSM, all pixels of which have the same height. If more than half of the pixels of the region were marked as high in the previous step, the remaining pixels of the region are also marked as high. Thus, the boundaries of the detected regions are smoothed and false alarms associated with slopes of hills and other sharp differences of the terrain height are rejected. The regions with small area are removed from consideration.

4. *Extraction of buildings among the found high regions*. Not only buildings and structures, but also high vegetation (trees) is related to the detected high-rise objects. The normalized difference vegetation index (NDVI) [6], which is calculated from the red and infrared channels of a satellite image, effectively distinguishes vegetation from other objects. However, in addition to vegetation, this index also extracts buildings, which will lead to skipping objects when the building extraction algorithm is running. To solve this problem, it is necessary to perform additional color filtering of the vegetation extracted using the NDVI.

The segmentation procedure splits a satellite image into homogeneous regions; however, buildings can be covered by several regions. The detection of high objects in the DSM allows us to extract an object entirely, but with low accuracy of boundary localization. We suggested the combination of the outputs of high region extraction and segmentation algorithms as follows: if more than 75% of the pixels of the segmented region are marked as high, then the whole region is considered high (the threshold was chosen empirically). This enables to increase the accuracy of object localization and reduce the number of false alarms (for regions of trees closely adjacent to buildings). The result of the proposed approach is shown in figure 1 and in table 1.

3. Building extraction using convolutional neural networks

The use of convolutional neural networks made it possible to significantly improve the performance of computer vision algorithms in solving a variety of tasks, including the task of extracting and classifying objects in satellite images. Due to the increasing interest of various organizations in this task, at least four global online competitions [7-9] on the extraction and classification of objects in aerospace images were conducted in the last two years. Among the variety of solutions, the best performance for building extraction in [7,8] was shown by convolutional neural networks – modifications of the U-Net architecture [3]. It is worth noting that a better accuracy of object extraction is achieved by training a separate model of U-Net for each class of objects. In this work, we used the implementation of U-Net described in [10].

U-Net has a deep convolutional encoder-decoder architecture. At the input, the network receives a tensor containing satellite image channels and channels with additional data about the scene. At the output, U-Net returns a single-channel mask coinciding with the size of the original image and containing the found buildings. The network is trained on batches consisting of fragments with the size of 112 ×112 pixels randomly taken from the image, to which reflection over the x or y axes is randomly applied. Its training was performed on a four-channel satellite image of an urban settlement with the size of about 5 × 5 km and 1-m resolution that contains more than 3,000 objects such as multi-storey buildings, garages, hangars and cottages. In addition, we used a digital surface model obtained from the stereo matching algorithm proposed in [1]. The examples of the input data are presented in figure 1. To assess the influence degree of the data composition on the recognition quality, four U-Net models were trained on the following data sets:

- image, NDVI, DSM;
- image, DSM;
- image, NDVI;
- image.

The results of the U-Net network trained on different data sets are shown in figure 1 and in table 1. Building extraction quality of the trained network was also evaluated using satellite images with different spatial resolutions. The results of building extraction from the images with the resolution of 2 m and 3 m per pixel can be observed in the figure 2 and table 2.

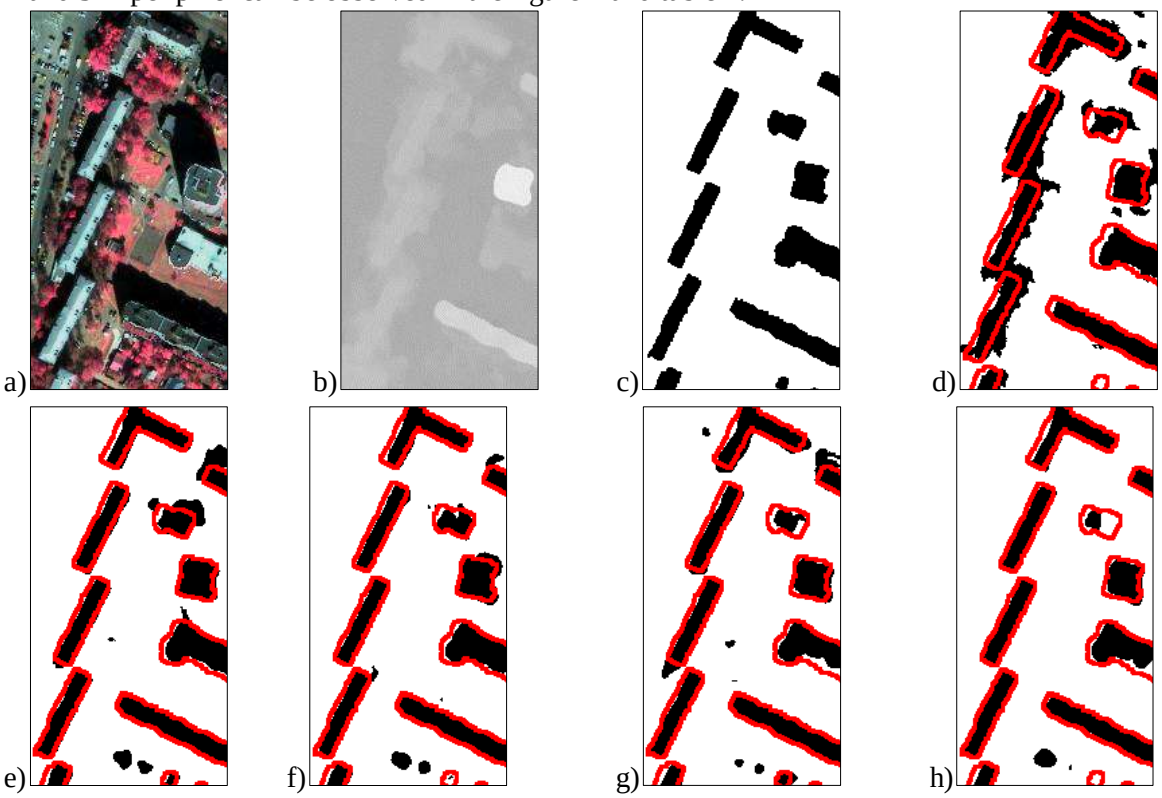


Figure 1. (a) Orthorectified four-channel satellite image (near-infrared, red, green channels) with the 1-m resolution, (b) digital surface model obtained from a stereo pair using a stereo matching algorithm, (c) the hand-drawn ground-truth building regions, (d) detection of buildings in the DSM and refinement of their boundaries using segmentation of spectral information. Building extraction by the U-Net network trained on: (e) image, DSM and NDVI, (f) image, DSM, (g) image and NDVI, (h) only on image. The extracted buildings and the ground-truth building boundaries are displayed in black and red respectively.

4. A comparison of considered approaches

The quality of the considered algorithms was evaluated using a fragment of a four-channel satellite image with the size of 900 × 500 pixels and 1-m resolution. The regions marked as buildings by the

algorithms were compared with the hand-drawn ground-truth building regions according to the following criteria:

1. Intersection over Union (also known as the Jaccard Index), which is sensitive to the accuracy of localization of the detected objects:

$$IoU = \frac{|S_{obj} \cap R|}{|S_{obj} \cup R|},$$

where S_{obj} – the regions of the extracted buildings, and R – the ground-truth building regions. Its value changes from 0 to 1, where 1 means full overlap of the found regions with the ground-truth ones.

2. The rate and amount of true positives, and false alarms in the output of the algorithms.

The results of the comparison can be observed in table 1 and in figure 1.

Of the two approaches to building extraction, the best performance was demonstrated by the U-Net network. Its advantage, as compared with the algorithm that extracts buildings only by using a DSM, is the ability to detect cottages and other low-rise buildings that are not visible in the DSM and these, at least, are one third of the objects in the initial 5 × 5 km image. Also, U-Net performs more accurate detection of the building boundaries.

According to table 1 and figure 1, U-Net shows the good result even without using additional information about the scene. However, the network was trained with height information better detect building boundaries, as was proved by the higher IoU value. The use of the NDVI lead to deterioration in building detection quality. A possible explanation for this might be that NDVI extracts not only vegetation but also buildings. The quality of the building extraction of U-Net was also evaluated on the image of another town with the size of about 5 × 5 km which can be seen in figure A1. Despite the fact that the image differs from the training set by shooting conditions and the composition and type of objects, U-Net (image, dsm) showed for it a similar rate of false alarms, correct and missed detections.

As regards building detection quality of U-Net for satellite images with different resolution, the network is also able to detect buildings in such cases. Figure 2 illustrates that the quality of building extraction depends on the building area and the network was unable to detect barely visible objects. Also, as the resolution is reduced, the quality of detection of building boundaries decreases. Perhaps, it is possible to improve the accuracy of boundary detection on images with different resolutions by adding such images to the training set of the network. It should be noted that to date several ways of improving the U-Net performance have been proposed [11-13], and it is most likely the conducted experiments will be valid for the proposed networks.

Table 1. A comparison of the building extraction algorithms. U-Net was trained on the 5 × 5 km image. The comparison was carried out on the 1 × 0.5 km image (116 buildings).

Algorithm	*IoU*	*True Positive*	*False Alarms*	*True Positive*	*False Alarms*
High-rise objects and segmentation	0.6	76	14	0.66	0.16
U-Net (image, ndvi)	0.74	95	10	0.82	0.09
U-Net (image, ndvi, dsm)	0.74	101	18	0.87	0.15
U-Net (image)	0.74	**102**	**7**	**0.88**	**0.06**
U-Net (image, dsm)	**0.76**	**102**	13	**0.88**	0.11

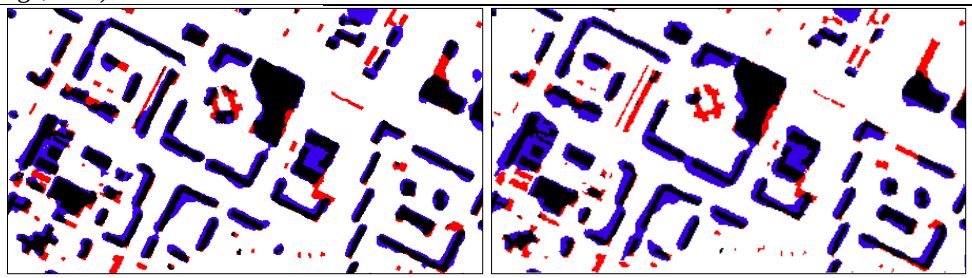


Figure 2. The results of building extraction using U-Net (image, DSM) for images with the 2-m resolution (left), and the 3-m resolution (right). True positive, false negative and false alarms regions are displayed in black, blue and red respectively.

Table 2. A comparison of the building extraction quality of U-Net using satellite images with different spatial resolutions. U-Net was trained on the satellite image with the 1-m spatial resolution (4 channels) and the digital surface model.

Satellite image resolution	*Total*	*True Positive*	*False Negative*	*False Alarms*	*True Positive*	*False Alarms*	*IoU*
1m	3141	3062	79	204	0.97	0.06	0.83
1m, only high buildings	735	728	7	–	0.99	–	–
2m	3285	1904	1381	109	0.60	0.05	0.50
2m, only high buildings	806	712	94	–	0.88	–	–
3m	3171	1282	1889	45	0.40	0.03	0.40
3m, only high buildings	754	645	109	–	0.86	–	–

5. Conclusion

This paper considered a comparison of two approaches to building extraction from satellite imagery and height data. The best performance was shown by the U-Net-based approach, which extracts a larger number of buildings with a better detection of their boundaries. Furthermore, U-Net shows the good result even without using additional information about scene. However, its application is possible only if there is a sufficiently large data set for training the network. The advantage of the approach based on extraction of buildings from a DSM is the absence of parameters that must be selected for each particular satellite image and it will work on images obtained from different satellites. In the future, it is planned to investigate the effect of the size and composition of the training set on the U-Net building extraction quality as well as the possibility of using trained models of neural networks for satellite images of different areas.

Appendix A

Figure A1. The result of building extraction using the trained U-Net model (image, dsm) for the image with the size of about 5 × 5 km and 1-m resolution. Detected buildings are displayed in green.

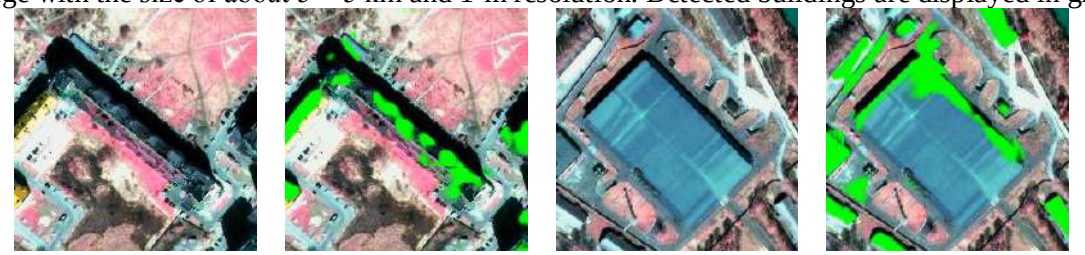


Figure A2. Examples of missed detections from figure A1. Left images in the pairs are fragments from the original image.

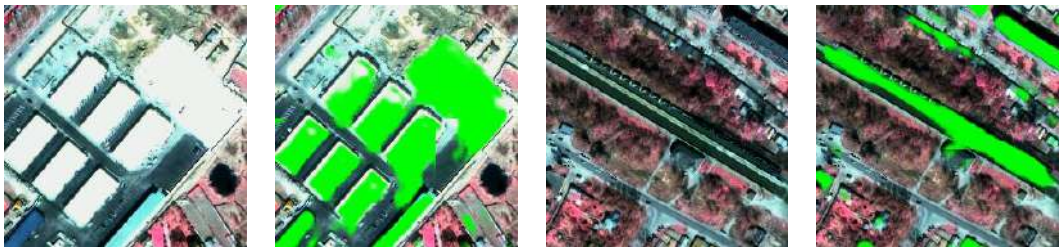


Figure A3. Examples of false alarms from figure A1. Left images in the pairs are fragments from the original image.

6. References

[1] Kostousov V B, Perevalov D S and Kornilov F A 2016 Digital terrain model generation from satellite stereoscopic data *Materials of the XXX conference of the memory of the outstanding designer of gyroscopic instruments N.N. Ostryakov* 382-388

[2] Fursov V A, Goshin Ye V and Kotov A P 2016 The hybrid CPU/GPU implementation of the computational procedure for digital terrain models generation from satellite images *Computer Optics* **40(5)** 721-728 DOI: 10.18287/2412-6179-2016-40-5-721-728

[3] Ronneberger O, Fischer P and Brox T 2015 U-Net: Convolutional networks for biomedical image segmentation *International Conference on Medical Image Computing and Computer-Assisted Intervention Lecture Notes in Computer Science* **9351** 234-241

[4] Maryanova A V 2015 The investigation of the quality of image segmentation algorithms depending on the size of objects in the image *Proceedings of the 46th International Youth School-Conference Actual Problems of Mathematics and its Applications* 129-134

[5] Dunaeva A V and Kornilov F A 2017 Building detection in remote sensing images using a digital surface model *Computational Mathematics and Information Technologies* **2(2)** 185-193

[6] Pettorelli N 2013 *The normalized difference vegetation index* (Oxford University Press) p 208

[7] *Dstl satellite imagery feature detection competition web page* (Access mode: https://www.kaggle.com/c/dstl-satellite-imagery-feature-detection) (01.05.2018)

[8] *The SpaceNet challenge web page* (Access mode: https://crowdsourcing.topcoder.com/spacenet) (01.05.2018)

[9] *Planet: understanding the Amazon from space competition web page* (Access mode: https://www.kaggle.com/c/planet-understanding-the-amazon-from-space) (01.12.2017)

[10] Iglovikov V, Mushinskiy S and Osin V 2017 Satellite Imagery Feature Detection using Deep Convolutional Neural Network: A Kaggle Competition *Preprint arXiv:1706.06169*

[11] Iglovikov V and Shvets A 2018 TernausNet: U-Net with VGG11 Encoder Pre-Trained on ImageNet for Image Segmentation *Preprint arXiv:1801.05746*

[12] Khalel A and El-Saban M 2018 Automatic Pixelwise Object Labeling for Aerial Imagery Using Stacked U-Nets *Preprint arXiv:1803.04953*

[13] Hamaguchi R, Fujita A, Nemoto K, Imaizumi T and Hikosaka S 2018 Effective use of dilated convolutions for segmenting small object instances in remote sensing imagery *IEEE Winter Conference on Applications of Computer Vision* 1442-1450

Acknowledgments

This work was carried out within the program of the Ural Branch of the Russian Academy of Sciences № 18-1-1-14.

Remote sensing investigation of inundation, elevation and land use assessment for vulnerability analysis in Moscow, Russia

K Choudhary [1], M S Boori[1, 2] and A Kupriyanov[1, 3]

[1]Samara National Research University, Moskovskoye Shosse34, Samara, Russia, 443086
[2]American Sentinel University, 2260 South Xanadu Way, Suite 310, Aurora, Colorado, USA
[3]Image Processing Systems Institute - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001

Abstract. Land use/cover change analysis assists decision makers to ensure sustainable development and to understand the dynamics of our changing environment. This research work is to understand natural and environmental vulnerability situation and its cause such as intensity, distribution and socio-economic effect in the Moscow, Russia based on remote sensing and Geographical Information System (GIS) techniques. A model was developed by following thematic layers: vegetation, LULC, geology, geomorphology and soil in ArcGIS 10.2 software using multi-spectral satellite data. With increasing scientific and political interest in regional aspects of global environmental changes, there is a strong stimulus to better understand the patterns causes and environmental consequences of LULC expansion in the elevation of Moscow state, one of the areas in the nation with fast economic growth and high population density. Satellite remote sensing images (Landsat TM, ETM and OLI) were employed to detect land cover changes. A 70 to 300 m inundation land loss scenarios for surface water and sea level rise (SLR) were developed using digital elevation models of study site topography through remote sensing and GIS techniques by ASTER GDEM and Landsat OLI data. The most severely impacted sectors are expected to be the vegetation, wetland and the natural ecosystem. Improved understanding of the extent and response of SLR will help in preparing for adaptation.

1. Introduction

Russia has a largely continental climate because of its sheer size and compact configuration. Most of its land is more than 400 km. from the sea and the centre is 3,840 km. from the sea. Russia's mountain ranges, predominantly to the south and the east, block moderating temperatures from the Indian and Pacific Oceans but European Russia and northern Siberia lack such topographic protection from the Arctic and North Atlantic Oceans. Moscow located in European Russia. It's the area of high environment sensitivity zone due to harsh climate conditions with maximum time frozen temperature below then zero [1-2]. The region is drained by numerous rivers and dotted with lakes due to heavy rainfall. Numerous studies have been performed to understand the variations in the Land surface temperature as a result of changes in the land surface properties. Since the 1960s, scientists have extracted and modelled various vegetation biophysical variables using remote sensing data and the normalized difference vegetation index is one such widely adopted index. Inverse relationship has

been reported between land surface temperature and vegetation index. Nowadays, it is recognize that climate change and sea level rise will impact seriously upon the natural environment and human society in the area [3-4]. There for sea level rise has to be one of the main impacts of climate change in Moscow. Presently remote sensing and GIS techniques are the powerful tool to investigate, predict and forecast environmental change scenario in a reliable, non-invasive, rapid and cost effective way with considerable decision making strategies. The main aim of this research work is to describe natural hazards impacts and land loss due to water level inundation from 70 to 300m in Volga river basin located in Moscow [5-6].

Vulnerability is a function of exposure, sensitivity and adaptive capacity. Where potential impacts are a function of exposure and sensitivity hence, vulnerability is a function of potential impacts and adaptive capacity. Where exposure components characterize the stressors and the entities under stress; sensitivity components characterize the first order effects of the stresses. These measures can be quantitative (e.g. precipitation variability, distance to market) or qualitative (e.g. political party affiliation, environmental preservation ethic). Other slightly different view favoured by the hazards and disasters research community is that adaptive capacity consists of two subcomponents: coping capacity and resilience. Coping capacity is the ability of people and places to endure the harm and resilience is the ability to bounce back after exposure to the harmful events. In both cases, individuals and communities can take measures to increase their abilities to cope and bounce back; again depending on the physical, social, economic, spiritual and other resources they have or have access to. Another basic issue in this analysis work is to assign weights to each factor according to its relative effects of factors considered in the vulnerability in a thematic layer. The application of subjective weightings on the one hand gives us some indication of how the relative importance of different factors might change with context and can also tell us how sensitive vulnerability rating are to perceptions of vulnerability in the expert community.

2. Study Area

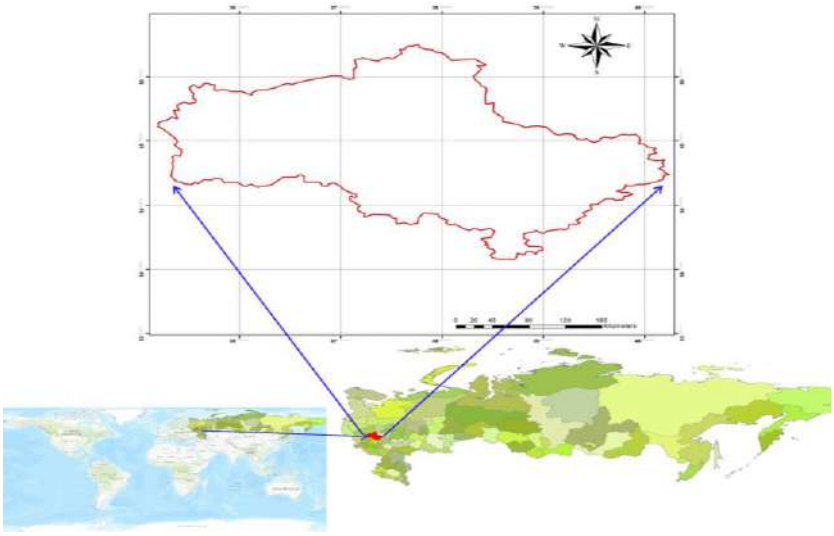


Figure 1. Location map of the study area in Moscow Region, Russia.

Moscow region an important historical, cultural, social and economic center in Russian Federation was selected for this study (fig.1). Moscow is the one of the most densely populated regions in the country and is the second most populated federal region. The Oblast has no official administrative canter, it is public authorities are located in Moscow and across other locations in the Oblast. As of the 2010 Census, its population was 7,095,120 and 7,23,1068 recorded in the 2015 Census. The latitude of the city is 55° 45’ 7” N and longitude is 37° 36’ 56" E. The region is highly industrialized, such as metallurgy, oil refining, mechanical engineering, food, energy and chemical industries [7-8].

The climate of Moscow region is humid continental, short but warm summers and long cold winters. The average temperature is 3.5 °C (38.3 °F) to 5.5 °C (41.9 °F). The coldest months are

January and February average temperature of −9 °C (16 °F) in the west and −12 °C (10 °F) in the east. The minimum temperature is −54 °C (−65 °F). Here are more than three hundred rivers in Moscow regions and most rivers belong to the basin of the Volga [9]. Which itself only crosses a small part in the north of Moscow region. They are mostly fed by melting snow and the flood fall on April-May. The water level is low in summer and increases only with heavy rain [10-11]. The river freezes over from late November until April.

3. Data and Methodology

3.1. Data

In this research work we used primary (satellite data) and secondary data such as ground truth for land use/cover classes and topographic sheets. The ground truth data were collected using Global Positioning System (GPS) for the year of 2015 in the month of June to August for image analysis and classification accuracy [12-13]. The specific satellite images used were Landsat OLI (Operational Land Imager) for landscape and Advanced Space borne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (ASTER GDEM) for elevation information.

3.2. Image pre-processing and classification

In pre-processing, first all images were georeferenced by WGS 1984 UTM projection, later on calibrated and remove there errors/dropouts. We use specific band combination and use image enhancement techniques such as histogram equalization to improve the classification accuracy [14]. At this stage, 50 points were selected as GCPs (Ground Control Points) for all images. Data sources used for the GCP selection were: digital topographic maps, GPS (Global Positioning Points) acquisitions. The data of ground truth were adapted for each single classifier produced by its spectral signatures for producing classification maps. For land use/cover classification, supervised maximum likelihood algorithm (MLC) was used in ArcGIS 10.2 software. MLC classification is based on training sites (signature) provided by the analyser based on his experience [15-16]. After training site whole image classified according to similar digital value of training site and finally classification give land use classified image of the area (fig.2). Seven main land use/cover classes have been find namely agriculture, barren land, forest, settlements, scrubland, water body and wetland in the study area (table. 1).

Table 1. Classes delineated on the basis of supervised classification.

Class name	Description
Agricultural	Cultivated areas, crop lands, grass lands, vegetables, fruits etc.
Barren land	This contains open lands mostly barren but also small vegetation.
Forest	Small trees and shrub vegetation area except for vegetation.
Scrubland	Scrub is a plant community describe by vegetation shrubs, often also including grasses and herbs.
Settlements	Includes construction activities along the coastal dunes as well as sporadic houses within the local village and some governmental buildings.
Water body	All the water within land mainly river, ponds, lakes etc.
Wetland	A wetland is a land area with standing water and low soil fertility.

A preparative requirement for the analysis of flooding impacts was the development of spatial datasets. A 1m spatial resolution digital elevation model (DEM) with error within 224 mm in elevation was constructed using ASTER GDEM images (fig.2). The GIS environment was used to classify and map the topology of land threatened by inundation [17].

The length of the sand spit at some places is more than 1km and they are highly vulnerable to river erosion basin.

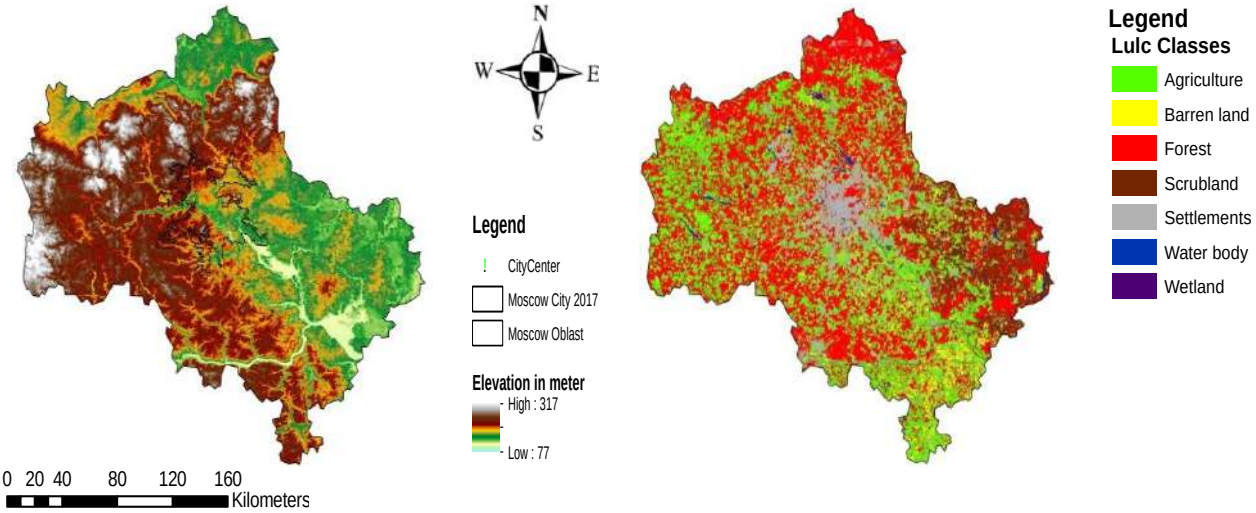


Figure 2. Elevation and land use/cover map of Moscow, Russia.

3.3. Data Analysis

All multi-spectral and temporal data were georeferenced based on topographic sheets with the help of ArcGIS 10.2 software. To improve the quality of research analysis we used different band ratio, image enhancement techniques and principal component analysis and in last supervised classification.

Table 2. Stability values of landscape units. (Ecodinamica Tricart, 1977).

Unit	Pedogenesis / morphogenesis Relation	Value
Stable	Prevails pathogenesis	1.0
Intermediate	Balance between pedogenesis and morphogenesis	2.0
Unstable	Prevails morphogenesis	3.0

Table 3. Weight table to each unite in a thematic layer.

Thematic maps/classes	Vulnerability grade levels
Land use/cover	
Agriculture	1.4
Barren land	1
Forest	1.7
Scrubland	2.2
Settlements	3
Water body	0.5
Wetland	0.8
Vegetation	
Dark coniferous forest	2.8
Grass	1.9
Grass herb	2
Oak forest	2.7
Pine	2.6
Pine leave forest	2.4
Shrub	2.3
Sphagnum bogs	1.6
Spruce	1.3
Wooded swampy fens	2.9
Geomorphology	
Plain area	2.5
Shrub land	2.3
Urban area	3
Water body	0.5
Wetland	0.8

Geology	
Flat	2.5
Gently undulating	2
Undulating	1.9
Soil	
Chernozems podzolized	1.9
Greys forest	2.1
Light-greys forest	1.8
Peats boggy	1.5
Podzols gleyic	1.4
Sod-podzolics	1.7
Sod-podzolics deepgley and gleyic	0.9

Thematic maps (fig. 3) of geology, geomorphology, soil, vegetation and land use/cover were prepared from Landsat ETM+ and OLI imageries. The weight of all landscape units based on Ecodinamica Tricart 1977 and Barbosa 1997 [18] stability concept, where stability was classified according to table 2. The weights of a landscape unit indicate the importance of any factor in relation to others. In natural vulnerability all thematic layer give same weight but in environmental vulnerability all thematic layer were given different weight based on their sensitivity in the study area.

The degree of vulnerability for all units was range from 0.0 to 3.0 (table 3) based on Barbosa and Crepani et al. (1996). The degree of vulnerability varies from 0 to 3 and is ranked as extreme, high, moderate, reasonable and low vulnerability. The weights of compensation indicate the importance of any factor in relation to others, as can be seen in the formula below for natural vulnerability map:

[(Theme 1) + (Theme 2) + (Theme 3) + (Theme 4)] / 4

For environment vulnerability we use following formula:

0.2 X [Theme 1] + 0.1 X [Theme 2] + 0.1 X [Theme 3] 0.1 [Theme 4] + 0.5 X [Theme 5]

Where: Theme 1: Geomorphology map, Theme 2: Simplified geological map, Theme 3: Soil map, Theme 4: Vegetation map and Theme 5: Land use/cover map.

The result mean was distributed in following five natural and environmental vulnerability classes:

1. Low vulnerability: less than or equal to 1.00;
2. Reasonable vulnerability: 1.1 to 1.50;
3. Moderate vulnerability: 1.51 to 2.00;
4. High vulnerability: 2.1 to 2.50;
5. Extreme vulnerability: greater than or equal to 2.51

4. Results

4.1. Land loss due to inundation

The DEM presented in figure 4 shows that low-lying land is more extensive at the north and canter of the study area. The areas lower than 1m above mean sea level (MSL), which are at risk of inundation under the minimum inundation level are vegetation, industry and urban area basically whole city.

The main results of land loss due to inundation are presented in Figure 5. The most significant changes would occur south-east side of the Moscow.

At the minimum inundation level (70m in fig.4), 0.04% (20.24 km^2) of the total area (table 4) would be flooded including: urban areas, natural vegetation and agricultural land and beaches. The area of submergence for 80m rise in water level is up to 60.12 km^2 (0.13%) and subsequently for 90m 258.29 km^2 (0.55%), 100m 895.84 km^2 (1.92%), 110m 2417.18 km^2 (5.18%), 120m 5108.59 km^2 (10.94%), 130m 8779.34 km^2 (18.80), 140m 12815.38 km^2(27.44%), 150m 16792.16 km^2 (35.95%), 180m 27976.67 km^2 (59.88%), 210m 38787.89 km^2 (82.98%), 240m 44584.13 km^2 (95.37%), 275m 46578.68 km^2 (99.62%) and 300m 46754.70 km^2 (100%) respectively (table 2). From the land use/cover map, it is clear that the maximum area is covered by agriculture which include Moscow city.

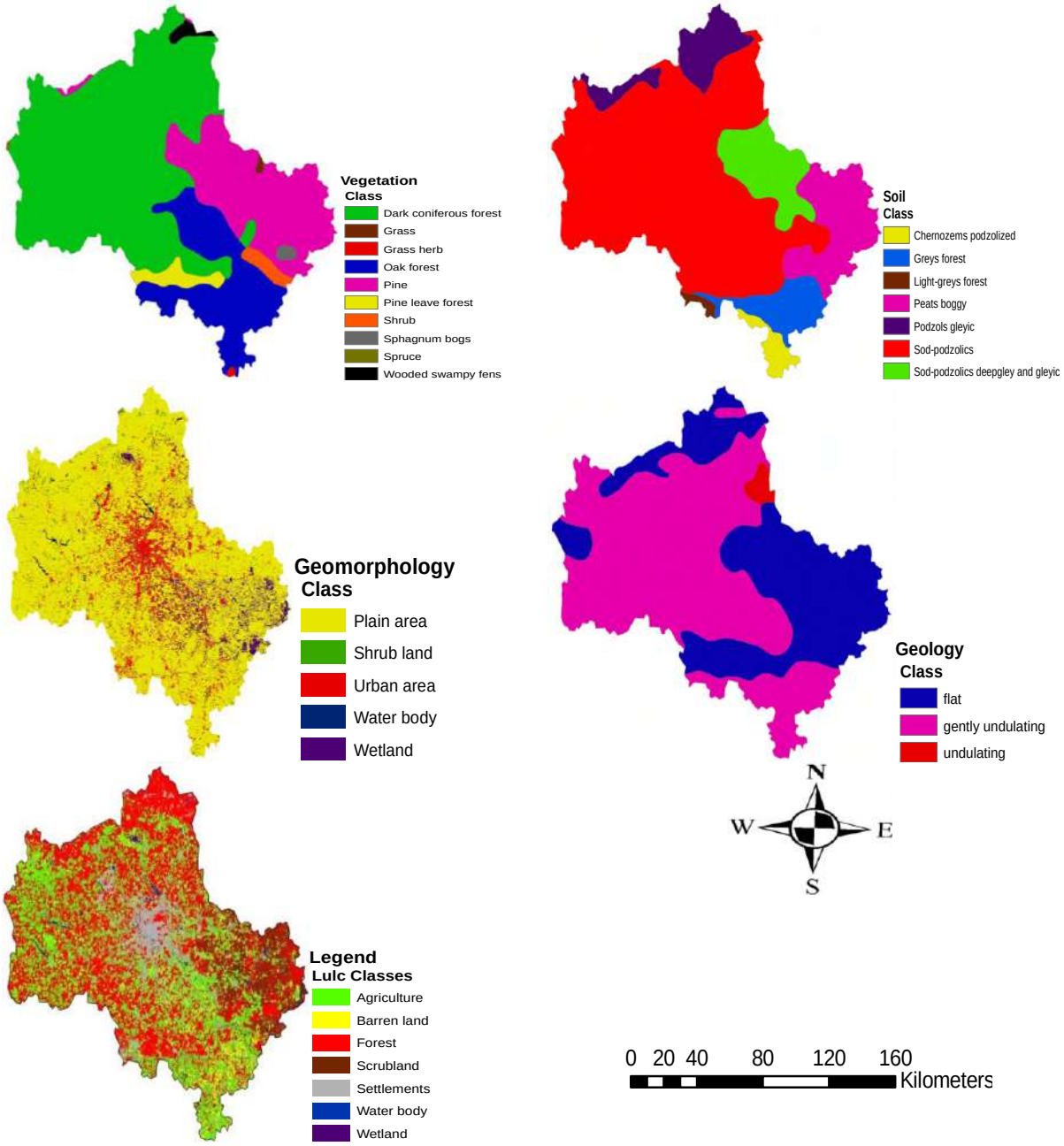


Figure 3. Simplified vegetation, soil, geomorphology, geology and land use/cover map.

At the full inundation level 300m in fig. Such a loss of land implies that the population living presently in these areas would be displaced. Even if some parts of the ecosystem of the wetland are not destroyed, because those parts could adapt to sea level rise and move landwards, the species richness is likely to decrease, due to repugnant new conditions where several plant communities and rare species would disappear. The area least vulnerable to inundation would be the southern and east part of the study area. However, parts of city and port, as well as an important river beach and natural forest would be flooded.

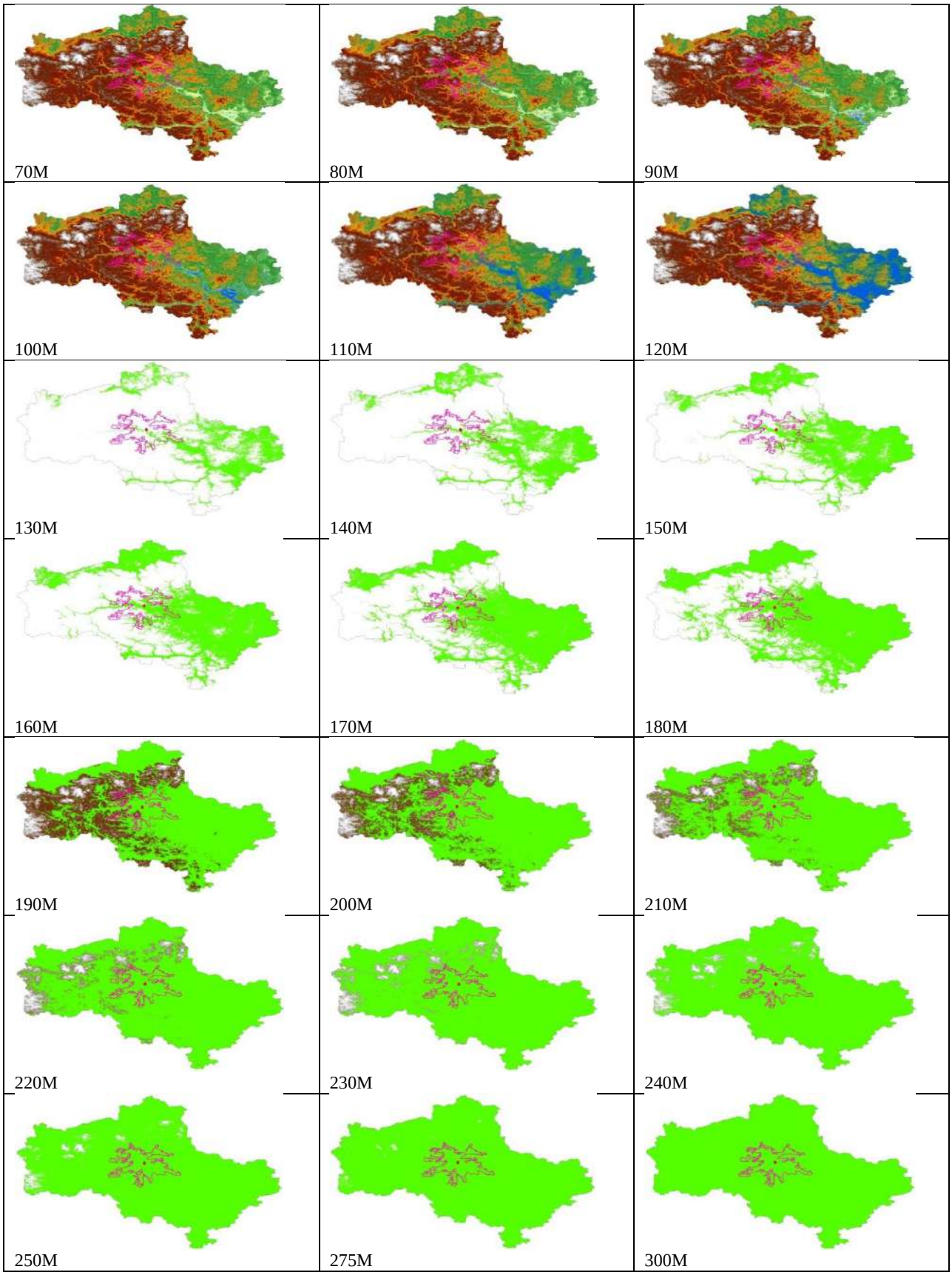


Figure 4. Land areas vulnerable to inundation in the Moscow, Russia.

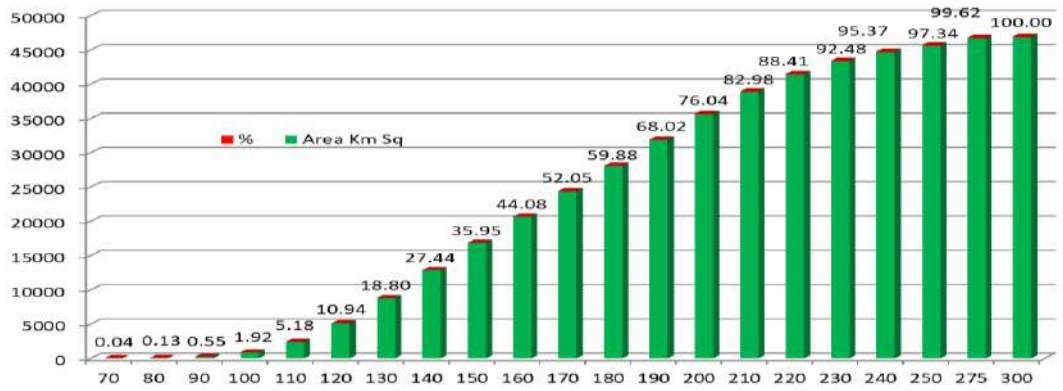


Figure 5. Inundation area graph of the Moscow, Russia.

Table 4. Potential land loss of the main sectors for 70m to 300m inundation levels scenarios (in km2 and in % of the total inundated areas).

	70M		80M		90M		100M		110M		120M		130M	
Class name	km²	%	km²	%	km²	%	km²	%	km²	%	km²	%	km²	%
Agriculture	4.92	24.32	22.86	38.02	119.46	46.25	388.84	43.41	853.78	35.32	1581.83	30.96	2557.14	29.13
Barren land	2.35	11.62	6.56	10.91	27.79	10.76	99.96	11.16	285.46	11.81	564.95	11.06	866.86	9.87
Forest	3.38	16.68	9.33	15.52	37.41	14.48	123.66	13.80	364.95	15.10	1001.52	19.60	2184.16	24.88
Scrubland	6.18	30.54	12.33	20.51	39.49	15.29	157.59	17.59	549.33	22.73	1182.53	23.15	1812.09	20.64
Settlements	1.66	8.18	4.23	7.04	16.80	6.50	71.10	7.94	229.20	9.48	542.06	10.61	1031.46	11.75
Water body	1.30	6.41	3.69	6.14	12.58	4.87	36.90	4.12	90.60	3.75	143.23	2.80	186.29	2.12
Wetland	0.45	2.24	1.12	1.86	4.76	1.84	17.79	1.99	43.85	1.81	92.48	1.81	141.34	1.61
Total	20.24	0.04	60.12	0.13	258.29	0.55	895.84	1.92	2417.18	5.18	5108.59	10.94	8779.34	18.80
	140M		150M		160M		170M		180M		190M		200M	
Class name	km²	%	km²	%	km²	%	km²	%	km²	%	km²	%	km²	%
Agriculture	3605.91	28.14	4651.49	27.70	5722.11	27.79	6903.08	28.39	8202.99	29.32	9545.05	30.03	10772.92	30.31
Barren land	1159.12	9.04	1448.88	8.63	1724.89	8.38	2000.08	8.23	2266.48	8.10	2506.25	7.88	2689.91	7.57
Forest	3667.29	28.62	5188.49	30.90	6621.71	32.16	7975.34	32.80	9352.34	33.43	10959.28	34.48	12840.17	36.13
Scrubland	2363.85	18.45	2793.90	16.64	3072.01	14.92	3225.66	13.27	3291.92	11.77	3317.75	10.44	3332.07	9.38
Settlements	1627.70	12.70	2272.80	13.53	2939.31	14.27	3611.60	14.85	4225.56	15.10	4780.97	15.04	5198.22	14.63
Water body	219.43	1.71	240.84	1.43	299.20	1.45	370.93	1.53	394.77	1.41	418.45	1.32	431.27	1.21
Wetland	172.08	1.34	195.77	1.17	213.47	1.04	227.91	0.94	242.61	0.87	258.15	0.81	273.61	0.77
Total	12815.38	27.44	16792.16	35.95	20592.70	44.08	24314.60	52.05	27976.67	59.88	31785.89	68.02	35538.18	76.04
	210M		220M		230M		240M		250M		275M		300M	
Class name	km²	%	km²	%	km²	%	km²	%	km²	%	km²	%	km²	%
Agriculture	11770.03	30.34	12506.84	30.26	13040.38	30.16	13413.83	30.09	13655.11	30.01	13881.56	29.80	13902.42	29.73
Barren land	2813.50	7.25	2886.41	6.98	2933.22	6.78	2962.56	6.64	2980.15	6.55	2992.10	6.42	2992.95	6.40
Forest	14680.51	37.85	16235.11	39.28	17450.69	40.37	18331.89	41.12	18951.63	41.64	19740.06	42.38	19889.92	42.54
Scrubland	3342.19	8.62	3350.29	8.11	3357.22	7.77	3362.99	7.54	3368.20	7.40	3376.74	7.25	3377.19	7.22
Settlements	5459.78	14.08	5620.17	13.60	5716.00	13.22	5776.87	12.96	5815.11	12.78	5847.92	12.55	5851.71	12.52
Water body	437.79	1.13	441.77	1.07	444.10	1.03	445.60	1.00	446.68	0.98	449.23	0.96	449.42	0.96
Wetland	284.08	0.73	287.53	0.70	289.38	0.67	290.39	0.65	290.77	0.64	291.07	0.62	291.09	0.62
Total	38787.89	82.98	41328.12	88.41	43230.99	92.48	44584.13	95.37	45507.66	97.34	46578.68	99.62	46754.70	100.00

4.2. Vulnerability analysis

Natural and environmental vulnerability maps are shown relationship in between landscape and vulnerability and able to tackle answers such as comparing of different types of vulnerability zones in the study area.

4.3. Natural vulnerability

Its map shows that maximum area in safe zones as 56.91% area in moderate vulnerability and 20.10% area in reasonable vulnerability zones, which represent that around 78% area of total study area is safe zone. Around 11.70% area goes in high vulnerability which is really need proper management

otherwise it will increase and harmful. The low vulnerability area is only 5.44% of the total study area, which is present in river and water body area. 5.83% area has been under extreme vulnerability, which is very less and close to water bodies. High vulnerability is due to fluctuation and extreme climate condition. Maximum vegetation area and close to river basin area under moderate vulnerability zone. Some part of wetland and vegetation under reasonable vulnerability and low vulnerability area, which represent maximum safe area in this study area. It is low vulnerability area due to less socio-economic activities and high density of vegetation (fig.6).

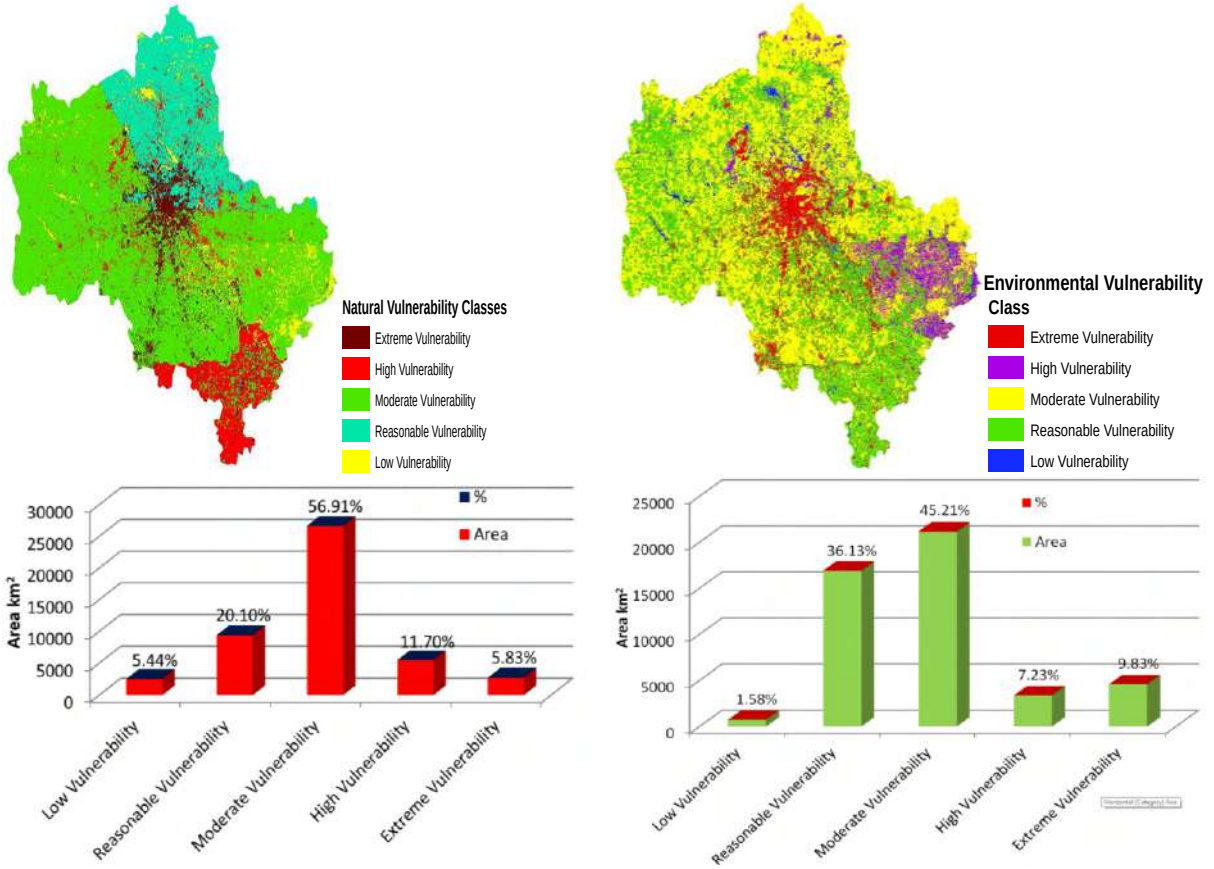


Figure 6. Natural and environmental vulnerability map.

4.4. Environmental vulnerability

Environmental vulnerability map is more sensitive than natural vulnerability. In environmental vulnerability around 46% area under moderate vulnerability zone but high and extreme vulnerability is higher than natural vulnerability. Here 7.23% area under high vulnerability and 9.83% under extreme vulnerability. Reasonable vulnerability is 36.13% and low vulnerability is 1.58%. Low vulnerability is present in river and water bodies, reasonable vulnerability present in wasteland and some parts in vegetation. Maximum study area has been under moderate vulnerability, which is present in vegetation and close to wetland and costal line. High vulnerability is present in close to river and its channels (fig.6). As study area is in north part of the Asia so maximum time of the year it is cover with ice, with harsh climatic condition. In winter only airways are the only way of approaching this area but in Summer Rivers also provide transportation facility. Here land use/cover classes and there convergent or encroachment induced by extreme cold and tough climatic condition in the study area. In extreme cold condition maximum areas convert in wasteland, where land has been unfertile. But in summer session ice has been melt and maximum land convert into wetland, forest and vegetation area etc.

5. Conclusion

Based on multi-temporal Landsat images, we determined that there was significant expansion of anthropogenic land cover in the Moscow. Analysis revealed that the area of anthropogenic land cover

was increased, resulting in a substantial reduction in natural land cover. The inundation maps can be overlaid on land use/cover maps to find out the extent of submergence of different land use/cover areas. By contrast, arable land declined by 10% due to occupation by urbanization and industrialization. It is necessary to incorporate the elevation levels for new settlements areas under the town planning acts so that human life and property are saved from natural hazards. The run-up levels can be used as guidance to determine safe locations of settlements from river basin. Vulnerability scenarios are useful for exploring uncertainties in vulnerability assessment on a regional basis, some regions show equal vulnerability to all scenarios, while other regions show different responses. This is an indicator for where we can be more or less uncertain about the future. Furthermore, it helps in indicating how society and policy can have an important role to play in future development pathways. The mapping, monitoring and modelling of land use/cover in such a vast territory as Moscow region could also contribute to the study of global environmental change.

6. References

[1] Tabak N M, Laba M and Spector S 2016 Simulating the Effects of Sea Level Rise on the Resilience and Migration of Tidal Wetlands along the Hudson River *PLoS One* **11(4)** 0152437 DOI:10.1371/journal.pone.0152437

[2] Boori M S and Ferraro R R 2015 Global Land Cover classification based on microwave polarization and gradient ratio (MPGR) *Geo-informatics for Intelligent Transportation* **71** 17-37 DOI:10.1007/978-3-319-11463-7-2

[3] Yang Z, Wang T, Voisin N and Copping A 2015 Estuarine response to river flow and sea-level rise under future climate change and human development *Estuarine, Coastal and Shelf Science* **156** 19-30

[4] McGranahan G, Balk D and Anderson B 2017 The rising tide: assessing the risks of climate change and human settlements in low elevation coastal zones *Environ. Urban.* **19(1)** 17-37

[5] Sun X, Li Y and Zhu X 2015 Integrative assessment and management implications on ecosystem services loss of coastal wetlands due to reclamation *J. Clean. Prod.*

[6] Cui L, Ge Z, Yuan L and Zhang L 2015 Vulnerability assessment of the coastal wetlands in the Yangtze Estuary, China to sea-level rise *Estuarine, Coastal and Shelf Science* **156** 42-51

[7] Sweet W V and Park J 2014 From the extreme to the mean: Acceleration and tipping points of coastal inundation from sea level rise *Earth`s future* **2(12)** 579-600

[8] Yabuki H, Park H, Kawamoto H, Suzuki R, Razuvaev V N, Bulygina O N and Ohata T 2011 Baseline Meteorological Data in Siberia (BMDS) Version 5.0, RIGC, JAMSTEC, Yokosuka, Japan *Distributed by CrDAP, Digital Media*

[9] Rotzoll K and Fletcher C H 2013 Assessment of groundwater inundation as a consequence of sea level Rise *Nature Climate Change* **3** 477-481

[10] Shalaby A and Tateishi R 2007 Remote sensing and GIS for mapping and monitoring land cover and land-use changes in the Northwestern coastal zone of Egypt *Appl. Geogr*. **27(1)** 28-41

[11] Choudhary K, Boori M S and Kupriyanov A 2017 Mapping and evaluating urban density patterns in Moscow, Russia *Computer Optics* **41(4)** 528-534 DOI: 10.18287/2412-6179-2017-41-4-528-534

[12] Tian G and Wu J 2015 Comparing urbanization patterns in Guangzhou of China and Phoenix of the USA: The Influences of Roads and Rivers *Ecol. Indic*. **52** 23-30

[13] Thakur A K, Sing S and Roy P S 2008 Orthorectification of IRS-P6 LISS IV data using Landsat ETM and SRTM datasets in the Himalayas of Chamoli district *Uttarakhand. Curr. Sci.* **95** 1459

[14] Thomlinson J R, Bolstad P V and Cohen W B 1999 Coordinating methodologies for scaling land cover classifications from site-specific to global: steps toward validating global map products *Remote Sens. Environ.* **70(1)** 16-28

[15] Choudhary K, Boori M S and Kupriyanov A 2017 Spatio-temporal analysis through remote sensing and GIS in Moscow region, Russia *CEUR Workshop Proceedings* **1901** 42-46

[16] Courchamp F, Hoffmann B D, Russell J C, Leclerc C and Bellard C 2014 Climate change, sea-level rise, and conservation: keeping island biodiversity afloat *Trends in Ecology & Evolution* **29(3)** 127-130

[17] Boori M S, Choudhary K, Kupriyanov A and Kovelskiy V 2015 Four decades urban growth and land use change in Samara Russia through remote sensing and GIS techniques *SPIE Remote Sensing and Image Formation* **9817** 01-07 DOI:10.1117/12.2227992
[18] Barbosa C C F 1997 Álgebra de mapas e suas aplicações em Sensoriamento Remote Geoprocessamento *Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais INPE* **111**

Acknowledgements
This work was partially supported by the Ministry of education and science of the Russian Federation; by the Russian Foundation for Basic Research grants (#16-41-630761; #16-29-11698, #17-01-00972).

Land cover classification and build spectral library from hyperspectral and multi-spectral satellite data: A data comparison study in Samara, Russia

M S Boori[1, 2], R Paringer[1, 3], K Choudhary[1], A Kupriyanov[1, 3] and R Banda[4]

[1]Samara National Research University, Moskovskoye Shosse 34, Samara, Russia, 443086
[2]American Sentinel University, Suite 310, Aurora, Colorado, USA
[3]Image Processing Systems Institute - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001
[4]Research Centre imarat (RCI), Defence Research & Development Organisation (DRDO) Hyderabad, India

Abstract The purpose of this research work is to compare hyperspectral and multispectral imagery to discriminating land-cover classes by *k*-nearest neighbor algorithm (KNN) supervised classification with migrating means clustering unsupervised classification (MMC) method and in last develop spectral library. We used Earth Observing-1 (EO-1) Hyperion hyperspectral data to Landsat 8 Operational Land Imager (OLI) and Advance Land Imager (ALI) multispectral data. Results indicate that KNN (95, 94, 88 overall accuracy and .91, .89, .85 kappa coefficient for Hyp, ALI, OLI respectively) shows better results than unsupervised classification (93, 90, 84 overall accuracy and .89, .87, .81 kappa coefficient for Hyp, ALI, OLI respectively). In addition, it is demonstrated that the hyperspectral satellite image provides more accurate classification results than those extracted from the multispectral satellite image. The higher classification accuracy by KNN supervised was attributed principally to the ability of this classifier to identify optimal separating classes with low generalization error, thus producing the best possible classes' separation.

1. Introduction

Remote sensing data are commonly used for land cover classification and mapping and its replaced traditional classification methods, which is expensive and time consuming. Since the early 1970s, multispectral satellite data have been widely used for land cove classification [1]. Multispectral remote sensing technologies, in a single observation, collect data from three to six spectral bands from the visible and near-infrared region of the electromagnetic spectrum [2]. This crude spectral categorization of the reflected and emitted energy from the earth is the primary limiting factor of multispectral sensors either spatially or spectrally to monitor sub-class level classification as they have very similar characteristics. Increasing the number of ''pure pixels'' through improved spatial resolution removes a large source of error in the remote sensing analysis classification. Species level mapping works well for monotypic stands, which occur in large stratifications [3]. Where species are more randomly distributed or patchy at fine scales (grain), accurate map classifications are difficult to obtain. So over the past two decades, the development of airborne and satellite hyperspectral sensor technologies has overcome the limitations of multispectral sensors [4]. Hyperspectral sensors collect several, narrow

spectral bands from the visible, near-infrared, mid-infrared and short-wave infrared portions of the electromagnetic spectrum [5]. These sensors typically collect more than 200 spectral bands, enabling the construction of an almost continuous spectral reflectance signature [6]. These bands are so sensitive to ground features that it is possible to record detailed information about earth surface. In addition, materials which have similar spectral features are possible to be discriminated [7]. However, to date, there is little research working on hyperspectral satellite data for land cover and land use mapping. As a result, accurate classification results with various land cover and land use classes are expected to be derived from a hyperspectral satellite image. Furthermore, narrow bandwidths characteristic of hyperspectral data permit an in-depth examination of earth surface features which would otherwise be 'lost' within the relatively coarse bandwidths acquired with multispectral data classification [8].

There are two broadways of classification procedures: (1) unsupervised classification and (2) supervised classification. Unsupervised classification algorithms require the analyst to assign labels and combine classes after the fact into useful information classes (e.g. forest, agricultural, water, etc). In many cases, this after the fact assignment of spectral clusters is difficult or not possible because these clusters contain assemblages of mixed land cover types. Generally speaking, unsupervised classification is useful for quickly assigning labels to uncomplicated, broad land cover classes such as water, vegetation/non-vegetation, forested/non-forested, etc). Furthermore, unsupervised classification may reduce analyst bias. Supervised classification allows the analyst to fine tune the information classes--often too much finer subcategories, such as species level classes. Training data is collected in the field with high accuracy GPS devices or expertly selected on the computer [9]. Consider for example if you wished to classify percent crop damage in corn fields. A supervised approach would be highly suited to this type of problem because you could directly measure the percent damage in the field and use these data to train the classification algorithm. Using training data on the result of an unsupervised classification would likely yield more error because the spectral classes would contain more mixed pixels than the supervised approach. Similarly, collecting in the field crop species training data is preferable to expertly selecting pixels on screen as it is often very difficult to determine which crops are growing visually [10]. Many studies have reviewed the application of hyperspectral and multispectral imagery in the classification and mapping of land use in particular water, urban, transportation and vegetation species level by detecting biochemical and structural differences. The main aim of this study is to evaluate *k*-nearest neighbor algorithm (KNN) supervised classification with migrating means clustering unsupervised classification (MMC) method on hyperspectral and multispectral imagery to discriminating land-cover classes [11]. For this purpose, a test site was selected an area located in the mainland of Samara region, Russia for which hyperspectral and multispectral imagery were made available. This research work focuses on the classification of multispectral and hyperspectral satellite imagery, in order to: (1) test the potential of hyperspectral satellite data for land cover classification till sub class levels; (2) evaluate the mapping performance of multispectral and hyperspectral satellite images and (3) finally develop spectral library.

2. Study site

We choose Samara region as a study area and its geographic coordinates are 53°12′10′′N, 50°08′27′′E (fig. 1).

Figure 1. Study area image, Samara region, Russia (source: Google Earth).

3. Methods

3.1. Selection of satellite data

In this research work we consider spatial, spectral and temporal resolution as well as cost and availability of data, when we reviewing most appropriate data. The Hyperion hyperspectral sensor (United States Geological Survey Earth Resources Observation Systems) and the multispectral OLI and ALI sensor [6] were then selected for this study. Few characteristics of all three sensors are showing in table 1.

Table 1. Characteristics of Hyperion, OLI and ALI sensors.

No.	Characteristics	Values		
		Hyperion	OLI	ALI
1	Sensor type	Push-broom	Push-broom	Push-broom
2	Wavelength range	400-2.500 nm	434-1.383 nm	433-2.350
3	Number of spectral bands	242	9	7
4	Spectral resolution	10 nm	15 – 200 nm	5 – 30 nm
5	Spatial resolution	30 m	30 m	30 m
6	Swath	7.5 km	185 km	37 km
7	Digitization	12 bits	12 bits	12 bits
8	Altitude	705 km	705 km	705 km
9	Repeat	16 day	16 day	16 day

3.2. Field work and ground trothing

Table 2. Land cover classes and their sub-classes in study area.

No.	Class level I	Class level II	Class level III
1.	Water	1.1 Inland water body	1.1.1 Deep water
			1.1.2 Shallow water
			1.1.3 Turbid water
			1.1.4 Clean water
		1.2 Lake	
		1.3 River	
2.	Vegetation	2.1 Forest	2.1.1 Conifer forest
			2.1.2 Deciduous/Broadleaved forest
			2.1.3 Mixed forest
		2.2 Agriculture	2.2.1 Heterogeneous agricultural area
			2.2.2 Permanent crops
		2.3 Mangroves	
		2.4 Grassland	
		2.5 Sparsely vegetated area	
3.	Settlements	3.1 residential	3.1.1 Old residential
			3.1.2 New residential
		3.2 Industrial	
		3.3 Park	
4.	Wetland		
5.	Bare land	5.1 Scrubland	
		5.2 Transitional woodland	
6.	Transportation	6.1 Road	6.1.1 Highway
			6.1.2 Inside road
			6.1.3 Concrete road
		6.2 Rail	
7.	Bare rocks		
8.	Sand dunes		

Fieldwork to map individual land cover classes and obtained spectral measurements of the dominant species was conducted at 60 sites in Samara region, Russia. Ground-trothing surveys should be undertaken within two weeks of acquiring satellite remote sensing imagery [7]. A random sampling method was used across the Samara region, around 7-8 samples selected in each class. The FieldSpec 3 ASD handheld spectrometer was used to obtain quantitative measurements of radiant energy easily and efficiently. We find eight meagre land cove classes and their sub-classes as shown in table 2.

3.3. Data preprocessing

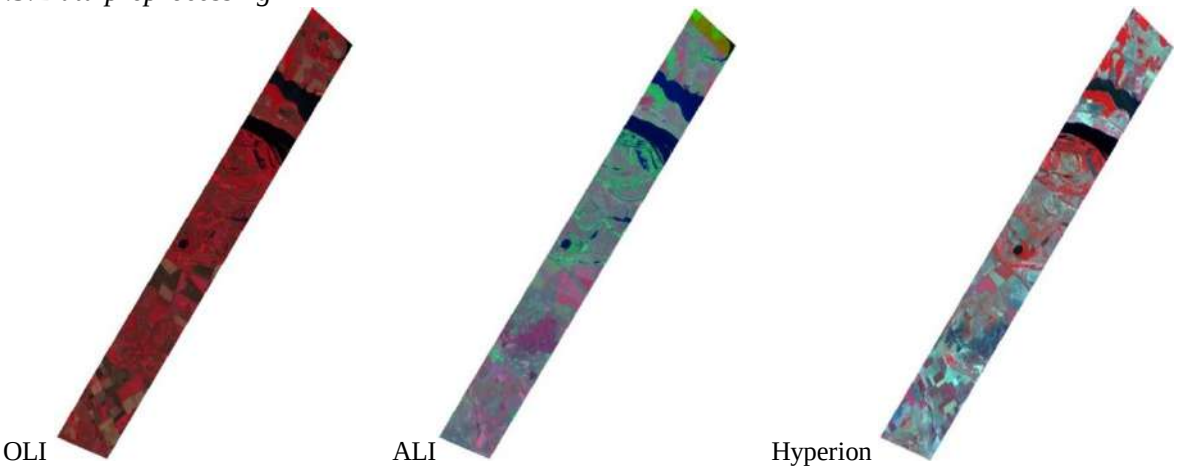


Figure 2. A sub-scene of the geometrically corrected OLI, ALI and Hyperion image over the study area in Samara region, Russia.

Digital image processing was manipulated in ArcGIS software. The scenes were selected to be geometrically corrected, calibrated and removed from their dropouts. All images were projected in UTM 39N, datum WGS 84 projection. Other image enhancement techniques like histogram equalization were also performed on each image for improving the quality of the image [8]. Some additional supporting data were also used in this study such as topographic sheets and field data. Digital topographical maps, 1:50,000 scale, were used for image georeferencing for the land use/cover map and for improving accuracy of the overall assessment. Using ArcMap, we made a composite raster data of OLI and ALI using Arc toolbox data management tools (fig. 2). Both images were composed of 9 and 7 different bands respectively, each representing a different portion of the electromagnetic spectrum. By combining all these bands, composite raster data were obtained (fig. 2). Table 3 shows details of OLI and ALI data. For pre-processing of Hyperion imagery, first georeferenced the image, subsequently were removed the non-calibrated bands of the Hyperion imagery. After this step, the resulting image was reduced to a subset of the studied region. These final 132 bands after this last pre-processing step were used in the present study (fig. 2).

Table 3. Left: Wavelength ranges of the OLI image. Right: Wavelength ranges of the ALI image.

OLI Bands	Wavelength (micrometers)	Resolution (meters)
Band 1 - Ultra Blue	0.435 - 0.451	30
Band 2 - Blue	0.452 - 0.512	30
Band 3 - Green	0.533 - 0.590	30
Band 4 - Red	0.636 - 0.673	30
Band 5 - Near Infrared (NIR)	0.851 - 0.879	30
Band 6 - Shortwave Infrared	1.566 - 1.651	30
Band 7 - Shortwave Infrared	2.107 - 2.294	30
Band 8 - Panchromatic	0.503 - 0.676	15
Band 9 - Cirrus	1.363 - 1.384	30

ALI Bands	**Wavelength (micrometers)**	**Resolution (meters)**
Pan	0.48 - 0.69	10
MS - 1'	0.433 - 0.453	30
MS - 1	0.45 - 0.515	30
MS - 2	0.525 - 0.605	30
MS - 3	0.63 - 0.69	30
MS - 4	0.775 - 0.805	30
MS - 4'	0.845 - 0.89	30
MS - 5'	1.2 - 1.3	30
MS - 5	1.55 - 1.75	30
MS - 7	2.08 - 2.35	30

3.4. Classification

In this research work we use USGS land use/cover classification system for all three images (fig. 3). For all three images, *k*-nearest neighbor algorithm (KNN) supervised classification and migrating means clustering unsupervised classification (MMC) approach was applied [9]. Training sites were collected based on field data and also take help with topography maps. Initially, training sites were chosen for all 27 sub-classes derived from all three images, than all 27 sub-classes were aggregated into following 8 meagre classes 1. Water; 2. Vegetation; 3. Settlements; 4. Wetland; 5. Bare land; 6. Transportation; 7. Bare rocks and 8. Sand dunes. For accuracy assessment 60 points were randomly collected in each image.

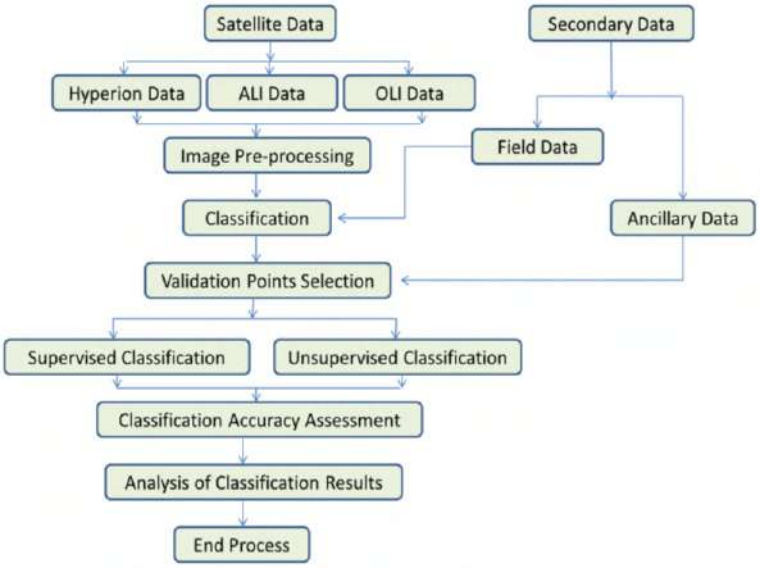


Figure 3. Flow diagram of methodological process.

3.4.1. Unsupervised classification

In unsupervised classification, image processing software classifies an image based on natural groupings of the spectral properties of the pixels, without the user specifying how to classify any portion of the image. Conceptually, unsupervised classification is similar to cluster analysis where observations (in this case, pixels) are assigned to the same class because they have similar values. The user must specify basic information such as which spectral bands to use and how many categories to use in the classification or the software may generate any number of classes based solely on natural groupings. Common clustering algorithms include K-means clustering, ISODATA clustering, and Narenda-Goldberg clustering [12].

Unsupervised classification yields an output image in which a number of classes are identified and each pixel is assigned to a class. These classes may or may not correspond well to land cover types of interest, and the user will need to assign meaningful labels to each class. Unsupervised classification often results in too many land cover classes, particularly for heterogeneous land cover types, and classes often need to be combined to create a meaningful map. In other cases, the classification may result in a map that combines multiple land cover classes of interest, and the class must be split into multiple classes in the final map. Unsupervised classification is useful when there is no preexisting field data or detailed aerial photographs for the image area and the user cannot accurately specify training areas of known cover type. Additionally, this method is often used as an initial step prior to supervise classification (called hybrid classification). Hybrid classification may be used to determine the spectral class composition of the image before conducting more detailed analyses and to determine how well the intended land cover classes can be defined from the image [13].

3.4.2. Supervised classification

In supervised classification the user or image analyst “supervises” the pixel classification process. The user specifies the various pixels values or spectral signatures that should be associated with each class.

This is done by selecting representative sample sites of known cover type called Training Sites or Areas. The computer algorithm then uses the spectral signatures from these training areas to classify the whole image. Ideally the classes should not overlap or should only minimally overlap with other classes. In ArcGIS software there are many different classification algorithms and we choose KNN supervised classification procedure as:

- ***K-nearest* neighbour algorithm** (KNN): K nearest neighbour is a simple algorithm that stores all available cases and classifies new cases based on a similarity measure (e.g., distance functions). KNN has been used in statistical estimation and pattern recognition already in the beginning of 1970's as a non-parametric technique. Pattern recognition is the scientific discipline whose goal is the classification of objects into a number of categories or classes. Depending on the application, these objects can be images or signal waveforms or any type of measurements that need to be classified. We will refer to these objects using the generic term patterns.

In supervised classification the majority of the effort if done prior to the actual classification. Once the classification is run the output is a map with classes that are labelled and correspond to information classes or land cover types. Supervised classification can be much more accurate than unsupervised classification, but depends heavily on the training sites, the skill of the individual processing the image, and the spectral distinctness of the classes. If two or more classes are very similar to each other in terms of their spectral reflectance (e.g., annual-dominated grasslands vs. perennial grasslands) misclassifications will tend to be high. Supervised classification requires close attention to development of training data. If the training data is poor or not representative the classification results will also be poor. Therefore supervised classification generally requires more times and money compared to unsupervised classification.

3.4.3. Classification accuracy assessment

Accuracy assessment of the thematic maps produced from the implementation of the supervised and unsupervised classification techniques on Hyperion, ALI and OLI imagery was also performed in ArcGIS based on the confusion matrix analysis [10]. As a result, the overall (OA), user's (UA) and producer's (PA) accuracies and the Kappa (Kc) statistic were computed. The OA provides a measure of the overall classification accuracy and is expressed as percentage (%). OA represents the probability that a randomly selected point is classified correctly on the map. Kc provides a measure of the difference between the actual agreement between reference data and the classifier used to perform the classification versus the chance of agreement between the reference data and a random classifier. PA indicates the probability that the classifier has correctly labelled an image pixel. UA expresses the probability that a pixel belongs to a given class and the classifier has labelled the pixel correctly into the same given class. In performing the accuracy assessment herein, a total of 60 sampling points for the different classes were selected (approximately 25 pixels per class) directly from the imagery following a random sampling strategy, and these points formed our validation dataset. Selection of those validation points was performed following exactly the same criteria used for the selection of training points, described earlier (Section 3.2). For consistency, the same set of validation points were used in evaluating the accuracy of the land use/cover thematic maps produced.

4. Results and discussion

4.1. Developing the spectral library

The land cover spectral library was developed by collected spectra of different sites from all three data sates and later on used as a set of reference spectra (fig. 4), to define different classes and mixed communities in Samara region, Russia. The average spectra illustrate a typical pattern, with significant divergence in the shape of the spectral curve between different land cover classes. The resulted spectral library shows all land cover class separation is possible in infrared region for all three data. In compare of all three datasets, all classes can easily separate in Hyperion data, as it have continues spectral band with very narrow bandwidth so specific bandwidth is sensitive for specific land cover class. ALI and OLI data have less capacity to separate all land cover class in compare of Hyperion

data due to less number of bands and longer bandwidth (fig. 4). In compare of ALI and OLI data sets, ALI has better results due to specific quality of sensor.

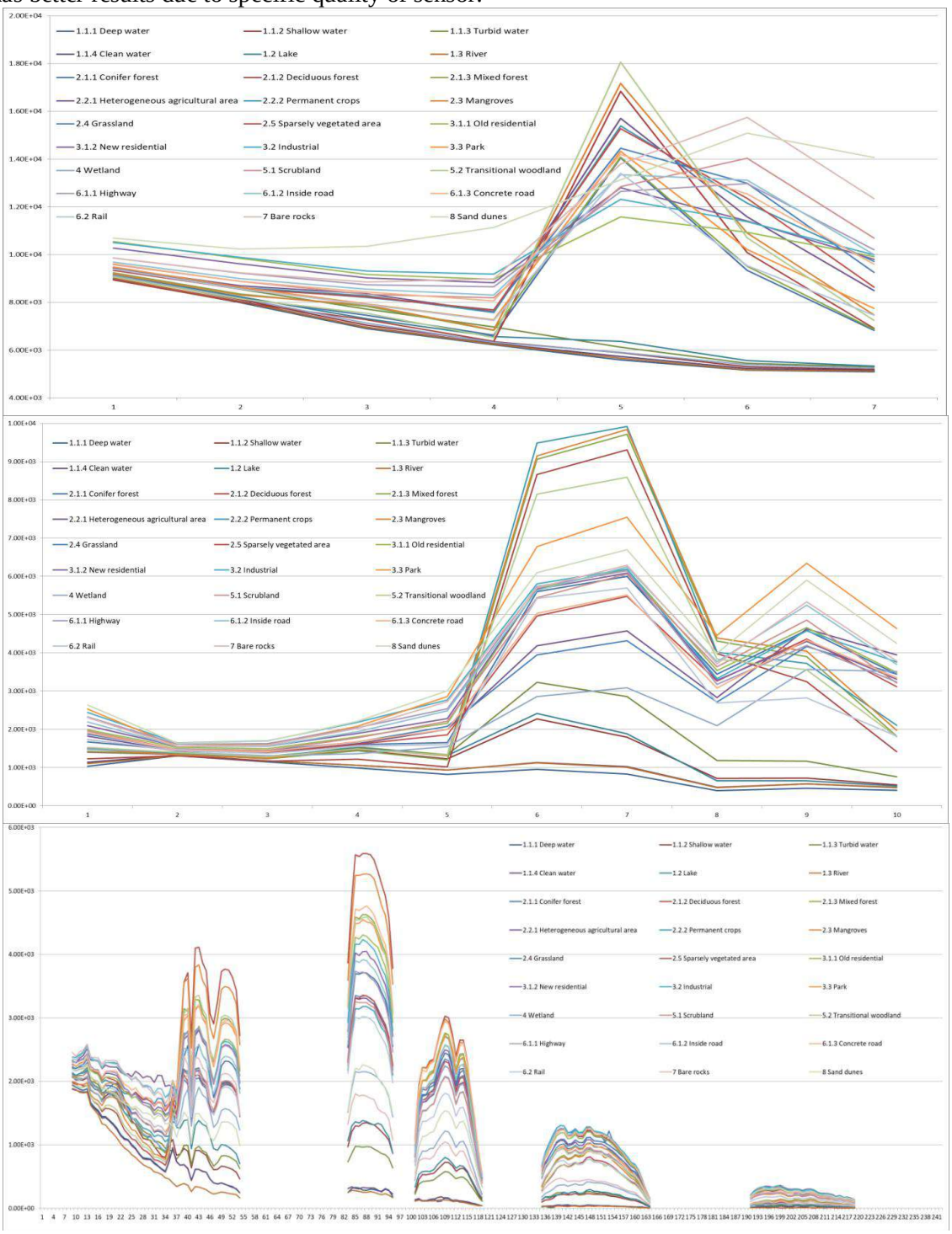


Figure 4. Representative spectra for 27 land cover classes by (A) OLI, (B) ALI and (C) Hyperion data in Samara, Russia.

Samara region land cover classes were defined into 8 major and 27 sub-classes based on species abundances and the characteristic dominant and sub-dominate land covers. For purposes of building the spectral library, a good understanding of the all land cover classes at each location in the study area was needed to utilize fully the information content of the spectra. Intra-specific and intracommunity variation were found across disturbance gradients. Phenomena included pattern,

shape-size, water content, structural changes, reduced biomass, lower ‘‘greenness’’ and chlorophyll, chlorosis and corresponding shifts across the spectral response curve. Methodological approaches to account for this variability, which can be used to assess stress, are still to be resolved. Large sets of reference spectra may be needed to fully characterize this variability. However, in this study, some land cover classes have similar spectral signature in different locations give additional benefits to sub-class level or species level mapping without a priori knowledge. However similar reflectance of mixed classes create confusing and difficult to identify class without field data or additional testing of spectral un-mixing and other spectral matching techniques.

4.2. Using spectral library for land cover classification

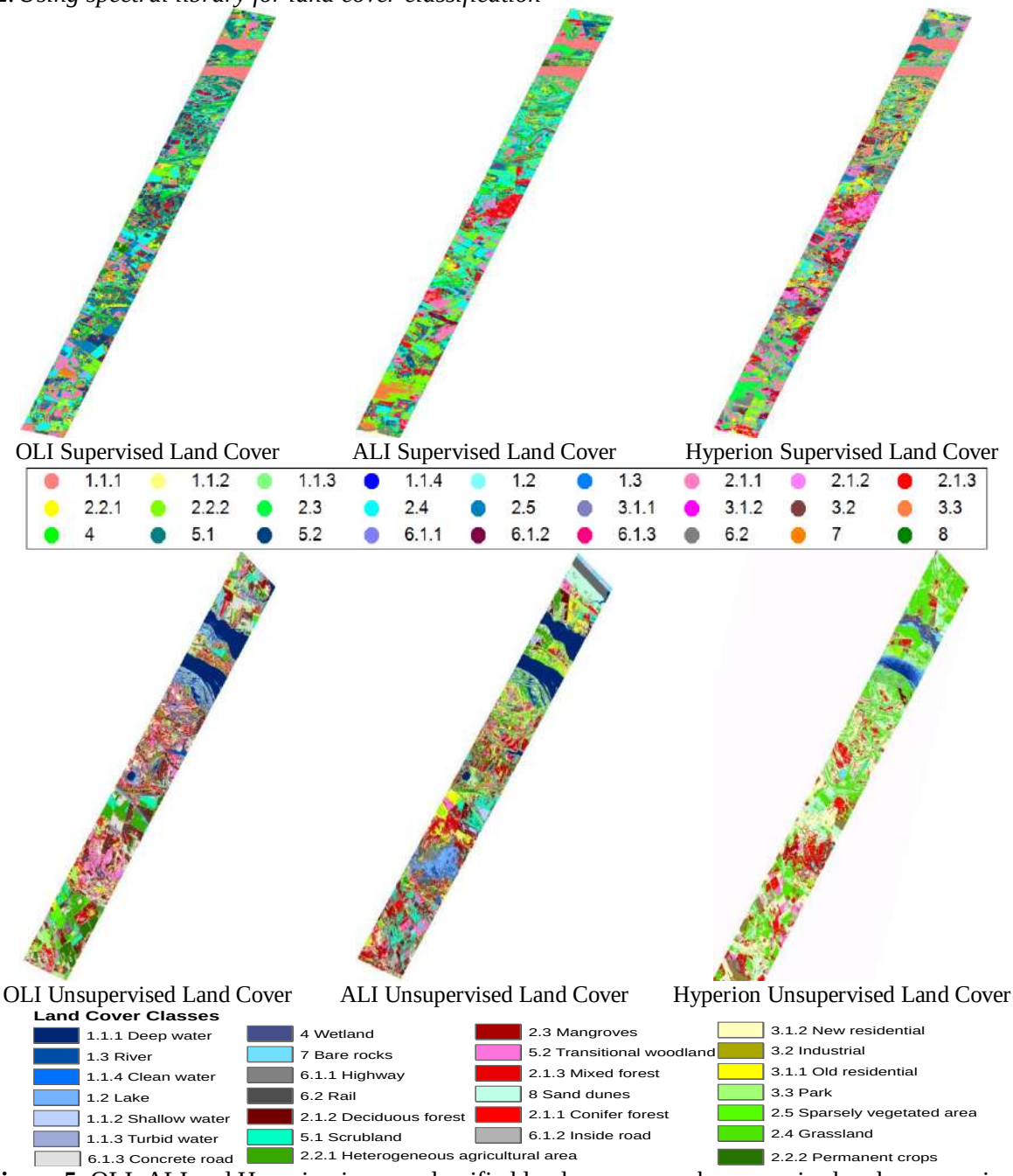


Figure 5. OLI, ALI and Hyperion images classified land cover maps by supervised and unsupervised classification methods.

Simple land use/cover classes such as forest, agriculture, settlements, water body and bare land can easily classify in high resolution data, even for their classification, we no need to use spectral library. Figure 5 show lulc images for all three data sets and in these images major land cover classes such as vegetation, water etc. can easily identify. As distinct land cover class patterns are closely related with specific bands/channels so without field data or spectral library or site situation/condition, these patterns cannot be identify, so basically, we need spectral library for sub-class level land cover classification.

A land cover map based on spectral library on hyperspectral (Hyperion) and multispectral data (OLI, ALI) produce 27 land cover classes (fig. 5). In comparison, hyperspectral data provide better results in place of multispectral data. This finding is similar to [8], who found that spectral resolution was more important for correct classification than spatial resolution, except in cases where high within pixel heterogeneity exceeded the pixel-to-pixel variance. In this research work a similar classification was produced from reference spectra extracted from the image (using GPS coordinates to identify classes) as from field-measured spectra of those land cover classes and resulted land cover map is a good representation of spectral pattern change due to continuous spectral bands in hyperspectral data.

Now we can say for wider use of hyperspectral data require improved methodologies and tools that facilitate and automate basic analyses and mapping, that can be specifically applied to land cover requirements. Both field and image methods for obtaining reference library spectra required complex processing and analysis. If a standard spectral library for land cover classes/ communities can be developed, it will aid resource managers by allowing them to utilize newer more powerful image analysis techniques while avoiding the data processing and expertise required to create the database. [4] similarly concluded that key challenges in applying these technologies on a wider scale included: building human capacity in advanced science and technology-based approaches, development of low cost and rugged IR spectroscopy instrumentation and development of decision support systems to help interpret spectroscopy data.

4.3. Classification comparison

The LULC maps produced by supervised and unsupervised classification on Hyperion, ALI and OLI data acquired over the study region are demonstrated in figure 5. The statistical results of classification accuracy assessment are shown in table 4. On the basis of accuracy assessment results, its appear that supervised classification somehow better results than unsupervised classification in overall accuracy and individual classes accuracy. Results indicate that for KNN the overall accuracy was 95, 94, 88 and kappa coefficient .91, .89, .85 for Hyp, ALI, OLI respectively, whereas for unsupervised it was 93, 90, 84 overall accuracy and .89, .87, .81 kappa coefficient for Hyp, ALI, OLI respectively. Among the two classifiers, supervised classification was the best in describing the spatial distribution and the cover density of each land cover category, as was also indicated from the statistics of the individual classes' results produced (table 4).

In all classes similar patterns were easily identify in both classification. PA and UA for the supervised classification ranged between the classes from 86% to 99%, and from 79% to 94%, whereas for unsupervised classification varied from 82% to 95% and from 75% to 92% respectively. In both classification the highest accuracy were in turbid water, permanent crops, sparsely vegetated area and bare rocks classes, followed by deep water, industrial, mixed forest, grassland, highway and sand dunes classes. In individual classes the lowest PA and UA in both classifications were shallow water, clean water, turbid water, grassland and highway classes. For all three data the highest PA and UA present in Hyperion data and lowest value present in OLI data. This was perhaps due to the similar spectral characteristics between the two classes, which was affected by the mixed pixels, caused by the low density of these vegetation types and combined with the low spatial resolution of the sensors.

So overall we can say supervised classification is better than unsupervised classification. In unsupervised classification algorithms require the analyst to assign labels and combine classes after the fact into useful information classes (e.g. forest, agricultural, water, etc). In many cases, this after the fact assignment of spectral clusters is difficult or not possible because these clusters contain assemblages of mixed land cover types. Generally speaking, unsupervised classification is useful for

quickly assigning labels to uncomplicated, broad land cover classes such as water, vegetation/non-vegetation, forested/non-forested, etc). Furthermore, unsupervised classification may reduce analyst bias. But supervised classification allows the analyst to fine tune the information classes--often too much finer subcategories, such as species level classes. Training data is collected in the field with high accuracy GPS devices or expertly selected on the computer. Consider for example if you wished to classify percent crop damage in corn fields. A supervised approach would be highly suited to this type of problem because you could directly measure the percent damage in the field and use these data to train the classification algorithm. Using training data on the result of an unsupervised classification would likely yield more error because the spectral classes would contain more mixed pixels than the supervised approach. Similarly, collecting in the field crop species training data is preferable to expertly selecting pixels on screen as it is often very difficult to determine which crops are growing visually. That`s why supervised classification is outperformed the unsupervised classification. When we compare both classification in hyperspectral and multispectral data, results show that supervised classification have highest accuracy, which authors attributed to the supervised ability to locate an optimal separating hyperplane [11].

Table 4. Summary of the results from the classification accuracy assessment conducted.

Land cover classes	Supervised Classification						Unsupervised Classification					
	Producer's accuracy (%)			User's accuracy (%)			Producer's accuracy (%)			User's accuracy (%)		
	Hyp	ALI	OLI	Hyp	ALI	OLI	Hyp	ALI	OLI	Hyp	ALI	OLI
1.1.1 Deep water	98	91	88	90	83	84	95	86	85	88	80	81
1.1.2 Shallow water	94	93	86	87	86	78	92	90	82	85	81	75
1.1.3 Turbid water	99	93	87	91	86	79	94	90	84	90	82	76
1.1.4 Clean water	95	92	87	87	86	78	91	87	83	86	83	75
1.2 Lake	95	93	87	87	85	82	90	91	82	84	81	80
1.3 River	91	93	88	85	88	80	88	90	85	81	85	79
2.1.1 Conifer forest	94	93	88	89	86	82	89	89	86	84	82	80
2.1.2 Deciduous/ Broadleaf forest	92	99	92	83	92	86	90	96	90	80	90	81
2.1.3 Mixed forest	92	97	92	84	91	86	91	94	90	81	89	82
2.2.1 Heterogeneous agricultural area	94	92	90	87	86	81	90	87	89	83	82	80
2.2.2 Permanent crops	99	92	90	94	88	85	95	88	89	92	85	81
2.3 Mangroves	96	93	91	91	88	87	92	90	90	90	83	85
2.4 Grassland	95	97	88	89	91	79	91	94	85	86	90	76
2.5 Sparsely vegetated area	99	92	88	91	84	82	96	88	84	90	81	81
3.1.1 Old residential	95	94	86	90	88	81	91	90	82	89	83	80
3.1.2 New residential	94	94	87	85	85	80	90	90	84	82	80	77
3.2 Industrial	98	94	89	93	88	85	95	91	86	91	84	81
3.3 Park	93	93	87	88	85	81	90	90	85	86	81	78
4. Wetland	94	93	88	86	88	80	91	90	84	84	86	79
5.1 Scrubland	96	92	88	89	88	81	91	89	84	85	85	78
5.2 Transitional woodland	95	92	95	87	85	85	90	90	92	83	80	82
6.1.1 Highway	94	97	87	89	91	79	89	94	84	86	90	76
6.1.2 Inside road	92	99	87	86	94	81	88	95	83	82	91	80
6.1.3 Concrete road	93	92	86	85	86	81	87	89	82	81	82	77
6.2 Rail	96	96	87	86	86	81	90	91	82	81	81	79
7. Bare rocks	99	94	88	94	86	83	94	90	85	91	83	81
8. Sand dunes	95	97	88	89	88	84	91	92	86	86	86	82
Overall accuracy	95	94	88				93	90	84			
Kappa coefficient	.91	.89	.85				.89	.87	.81			

5. Conclusions

This research work demonstrates the potential of hyperspectral and multispectral data for land cover monitoring and assessment. Currently, limitations of both data availability and cost remain, as do significant methodological and technical issues. However this research work highlights developing spectral library for land cover classes. In order to facilitate a global approach to applications of new advanced technologies for mapping and monitoring of landscape, a standardized classification system for land cover classes should be adopted to make best use of the spectral libraries and to facilitate a global remote sensing-based monitoring and assessment capacity. Additionally spectral library provide useful reference framework for landscape assessment and also support and promote new technology in terms of new space based high resolution hyperspectral instruments for earth observation. The accuracy assessment results show that supervised classification is better than unsupervised classification for all three (Hyperion, ALI and OLI) imagery. The higher classification accuracy reported by supervised classification is mainly attributed to the fact that this classifier has been designed as to be able to identify an optimal separating hyperplane for classes' separation, which the unsupervised may not be able to locate. This research found that, data analysis of hyperspectral imagery has the potential for improving classification accuracies of land cover and land use over multispectral imagery with the same resolution. If images were acquired the same day and time, then accuracies would be even more comparable. The latter, from an operational perspective, can be of particular importance particularly in the Mediterranean basin, since it can be associated to the mapping and monitoring of land degradation and desertification phenomena which are frequently pronounced in such areas.

6. References

[1] Clark M L 2017 Comparison of simulated hyperspectral HyspIRI and multispectral Landsat 8 and Sentinel-2 imagery for multi-seasonal, regional land-cover mapping *Remote Sensing of Environment* **200** 311-325

[2] Boori M S, Choudhary K, Kupriyanov A and Sugimoto A 2017 Informatics and computation method for inundation and land use study in arctic sea eastern Siberia, Russia *Proc. of SPIE* **10176** 101761D DOI:10.1117/12.2268153

[3] Lillesand T and Kiefer R 2000 *Remote Sensing and Image Interpretation* (John Wiley & Sons, Inc., New York) 363-370

[4] Akbari D, Homayouni S, Safari A and Mehrshad N 2016 Mapping urban land cover based on spatial-spectral classification of hyperspectral remote-sensing data *International Journal of Remote Sensing* **37(2)** 440-454 DOI: 10.1080/01431161.2015.1129561

[5] Boori M S, Choudhary K, Evers M and Paringer R 2017A review of food security and flood risk dynamics in Central Dry Zone area of Myanmar *Procedia Engineering* **201** 231-238 DOI: 10.1016/j.proeng.2017.09.600

[6] Boori M S, Choudhary K and Kupriyanov A 2017 Vulnerability evaluation from 1995 to 2016 in Central Dry Zone area of Myanmar *International Journal of Engineering Research in Africa* **32** 139-154 DOI: 10.4028/www.scientific.net/JERA.32.139

[7] USGS, Earth Observing-1 (EO-1) 2001 *US Geological Survey*.

[8] Ahmed F 2006 *Personal communication* (University of KwaZulu-Natal, Pietermaritzburg, South Africa)

[9] Dalponte M, Ole Ørka H, Ene L T, Gobakken T and Naesset E 2014 Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data *Remote Sensing of Environment* **140** 306-317

[10] Congalton R and Green K 1999 *Assessing the accuracy of remotely sensed data: principles and practices* (Boca Raton, FL: CRC/Lewis Press) p 137

[11] Boori M S, Choudhary K, Kupriyanov A, Sugimoto A and Paringer R 2016 Land use/cover change detection and vulnerability assessment in Indigirka river basin, Eastern Siberia, Russia *Imnage Processing, Geoinformatics and Information Security* **1638** 270-283 DOI: 10.18287/ 1613-0073-2016-1638-270-283

[12] Myasnikov V V 2015 A local order transform of digital images *Computer Optics* **39(3)** 397-405 DOI: 10.18287/0134-2452-2015-39-3-397-405
[13] Belim S V and Larionov S B 2016 An algorithm of image segmentation based on community detection in graphs *Computer Optics* **40(6)** 904-910 DOI: 10.18287/2412-6179-2016-40-6-904-910

Short-term traffic flow forecasting using a distributed spatial-temporal model

A A Agafonov[1] and A S Yumaganov[1]

[1]Samara National Research University, Moskovskoye shosse 34, Samara, Russia, 443086

Abstract. In this paper, we consider the problem of short-term traffic flow prediction. We propose a distributed model for short-term traffic flow forecasting based on the k nearest neighbors method, that takes into account spatial and temporal traffic flow distribution. To consider spatial-temporal correlations, we partition a transportation graph in clusters by an area and describe traffic flow by a feature vector defined for each cluster. The proposed model is implemented as a MapReduce based algorithm on an Apache Spark framework. The proposed traffic flow prediction model is tested using the actual average traffic speed data over a road network in Samara, Russia.

1. Introduction

Issues related to the traffic flow management are common in every major city around the world. Traffic congestion leads to economic, environmental and social problems, which emphasizes the importance of the transport planning and logistics. To solve these problems, it is important to obtain accurate and timely traffic flow information. Due to this fact, road traffic forecasting has been a subject of active research for more than 40 years.

Efforts devoted to mitigate the traffic congestion problem are usually classified in three directions: modification of the transport infrastructure, improving the operational quality of the public transport and managing traffic flows. The first and the second directions are often limited by economic or social factors, while the traffic flow management has been continuously improving due to the development of traffic data collecting and processing technologies.

Recently, much attention has been paid to the data-driven programming paradigm. This interest is explained by the development of new technologies, methods and techniques for massive data processing within the Big Data concept, the availability of multiple data sources for predicting traffic flows, and the "open data" idea, that some data should be freely available to everyone to use, without restrictions from copyright, patents or other mechanisms of control.

Short-term traffic flow forecasting considers the traffic prediction problem on the basis of current and archived information about the traffic flows state. A review of the latest achievements in the road traffic forecasting field, as well as the main unresolved technical challenges, can be found in [1]. Most research on this topic has focused on developing methods for modeling the characteristics of traffic flows (for example, density or speed). An overview of the methods of short-term traffic forecasting presented in the [2]. These methods can be classified into three categories:

1) Parametric methods [2, 3], including time series models [4], state space models, etc.

2) Non-parametric methods [5], including models of artificial neural networks [6], k-nearest neighbor (kNN) [7], support vector regression (SVR) [8].
3) Hybrid methods that combine parametric and non-parametric methods [9, 10].

However, these methods have both advantages and disadvantages when working under different conditions using different datasets, so it is hard to conclude that one method significantly superior others.

In this paper, we propose an approach based on the k nearest neighbor algorithm - one of the main non-parametric techniques for short-term traffic flow prediction. Results presented in [5, 11, 12] showed that kNN outperformed other modern comparable models, including ANN, SARIMA, random forest, and Naïve Bayes. However, if the sample data size is too large, kNN may not be suitable for real-time prediction due to the computational costs. Despite this issue, a relatively small number of works are devoted to the short-term traffic flow forecasting with a focus on processing big traffic data using the distributed computations, in particular, using the MapReduce framework [12, 13].

In this article, we consider a problem of short-term traffic flow forecasting for 10 minutes ahead. We focus on developing a distributed forecasting model based on the weighted kNN algorithm, taking into account the spatial and temporal characteristics of the transport flows in the spatially compact area of a transport network. For distributed data processing, we use MapReduce processing model implemented in the open source cluster-computing framework Apache Spark. Experimental analysis on real-world traffic data sets allows us to conclude that the proposed model has a high prediction accuracy and reasonable execution time, sufficient for real-time prediction.

The paper is organized as follows. Section 2 contains the formulation of the problem. The proposed model and its distributed implementation are described in detail, respectively, in Sections 3 and 4. In section 5, we provide experimental results of the proposed model and verify the accuracy of the proposed approach. Finally, we conclude the paper, and then present possible directions for further research.

2. Problem formulation

A road network is considered as a directed graph $G = (V, E))$, with nodes $V, N_V = |V|$ representing the road intersections and edges $E, N_E = |E|$ denoting road segments.

Let V_t^j denotes an observed traffic flow characteristic on an edge $j \in E$ at time interval t. As a traffic flow characteristic can be used travel time, average speed, density or flow.

In this work as a predicted traffic flow characteristic for the experimental study, we use the average traffic speed.

The short-term traffic flow forecasting problem can be formulated as follows: given a graph $G(V, E)$ and sequence $\{V_t^j\}, j \in E, t = 1, 2, \ldots T$ of observed traffic flow data, predict the traffic flow characteristic $\hat{V}_{t+\Delta}^j, j \in E$ at time interval $(t + \Delta)$ for a predefined prediction horizon Δ.

3. Proposed model

In this paper, we propose a short-term traffic flow forecasting model based on non-parametric regression k-nearest neighbors algorithm. To apply the kNN method to the traffic flow prediction problem, it is necessary to solve the following tasks:

1. Define a feature vector to describe traffic flow.
2. Define a suitable distance metric to determine the proximity between a feature vector describing current traffic flow characteristics and feature vectors describing historical traffic flow observations.

3. Define a prediction function to forecast a traffic flow characteristic by selected nearest neighbors.

These challenges are described in the next subsections.

3.1. Feature vector

The choice of a feature vector in the kNN method depends on the particular application of the method in practice. To solve the traffic flow prediction problem, it is reasonable to use a feature vector that takes into account spatial and temporal correlations of the traffic flow characteristics.

In the paper [12] as a feature vector authors used traffic flow of targeted road segment j, downstream road segment $j-1$ and upstream road segment $j+1$ for T time intervals:

$$(V_{t-T}^{j}, \ldots, V_{t-1}^{j}, V_{t}^{j}, V_{t-T}^{j-1}, \ldots, V_{t-1}^{j-1}, V_{t}^{j-1}, V_{t-T}^{j+1}, \ldots, V_{t-1}^{j+1}, V_{t}^{j+1}) \tag{1}$$

However, such feature vector does not consider traffic flow on adjacent segments. In addition, in some cases, the upstream / downstream road segment cannot be uniquely determined. Therefore, to describe traffic flow, it is proposed to use a feature vector that taking into account the traffic flow characteristics in the spatially-compact cluster of the transport network graph.

In this paper, we define the feature vector as follows:

1. The transportation network graph is partitioned into several spatially compact clusters $\{G_i\}$. In each cluster i the feature vector is defined as follows:

$$\{V_t^j\}^i, j \in G_i, t = t_{cur} - T, \ldots, t_{cur} \tag{2}$$

2. For the defined feature vector $\{V\}^i$ in the cluster i dimension reduction is performed using principal component analysis procedure. Result of this procedure is a new feature vector $\{X_n\}^i, n = 1, \ldots, N$.
3. Proposed feature vector for each road segment $j \in E$ is defined from the initial feature vector of the targeted road segment j and the feature vector of the cluster i such that $j \in G_i$:

$$S_j = (\{V_t^j\}, \{X_n\}^i), \quad j \in G_i; \quad t = t_{cur} - T, \ldots, t_{cur}; \quad n = 1, \ldots, N. \tag{3}$$

Graph partitioning algorithm is described in the next subsection.

3.2. Graph partitioning

Let each edge $i \in E$ corresponding to the road segment e_i with two terminal points $x_{start}^i = \left(x_{start}^0, x_{start}^1\right)^i$ and $x_{end}^i = \left(x_{end}^0, x_{end}^1\right)^i$.

Then graph partitioning by an area can be described as follows:

1. Choose the numbers of clusters M_0, M_1.
2. The cluster G_m with index $m = m_0 M_1 + m_1, \left(m_0 = \overline{0, M_0 - 1}; m_1 = \overline{0, M_1 - 1}\right)$ contains the edges $i \in E$, for which coordinates of at least one of the corresponding terminal points are inside the corresponding rectangular area Π_{m_0,m_1}:

$$G_{m_0 M_1 + m_1} \equiv \left\{ i \in E : x_{start}^i \in \Pi_{m_0,m_1} \vee x_{end}^i \in \Pi_{m_0,m_1} \right\}, \tag{4}$$

where

$$\Pi_{m_0,m_1} \equiv \left[x_{min}^0 + \frac{m_0}{M_0}\left(x_{max}^0 - x_{min}^0\right), x_{min}^0 + \frac{m_0+1}{M_0}\left(x_{max}^0 - x_{min}^0\right) \right] \times \left[x_{min}^1 + \frac{m_0}{M_0}\left(x_{max}^1 - x_{min}^1\right), x_{min}^1 + \frac{m_0+1}{M_0}\left(x_{max}^1 - x_{min}^1\right) \right],$$

$$x^s_{min} = \min_{\substack{v=\{start,end\}\\ i\in E}} x^{s,i}_v, \quad x^s_{max} = \max_{\substack{v=\{start,end\}\\ i\in E}} x^{s,i}_v, \quad s = 0, 1.$$

The number of clusters along the vertical and horizontal axis M_0, M_1 is chosen empirically. We assume, that each edge of the graph can get into only one cluster.

3.3. Proximity measure

To define the proximity between the feature vectors, it is necessary to determine a suitable distance metric. Different distance functions between feature vectors are available in the literature, including Euclidean, Mahalanobis, Hamming distance.

In this paper, we use a weighted Euclidean distance, modified to use the feature vector describing transportation network clusters. The distance is considered separately for parts of the feature vector describing traffic flows on the current segment $\{V\}$ and in the corresponding cluster $\{X\}$.

$$d(S, \bar{S}^i) = \sqrt{\sum_{t=1}^{T} \beta^{T-t+1} \left(V_t - \bar{V}^i_t\right)^2} + \alpha \sqrt{\sum_{n=1}^{N} \left(X_n - \bar{X}^i_n\right)^2}. \tag{5}$$

where $0 < \alpha \leq 1$,
T denotes the total number of time intervals in the feature vector,
N denotes the total number of elements in the feature vector describing the graph cluster,
S is the feature vector describing current traffic flow,
$\bar{S}^i$ is the feature vector describing ith historical traffic flow,
$V_t, \bar{V}^i_t$ are the feature vectors values representing respectively current and historical traffic flows on the selected road segment at time interval t,
$X_n, \bar{X}^i_n$ are the nth feature vectors values representing respectively current and historical traffic flows in the graph cluster.

3.4. Prediction function

The traditional approach for estimating the value in k-NN regression is to choose the average or the weighted average of the values of its k nearest neighbors [5].

A prediction function by the average has the following form:

$$\hat{X}_{T+1} = \frac{1}{k} \sum_{k=1}^{K} X^k_{T+1} \tag{6}$$

where $\hat{X}_{T+1}$ is the predicted traffic flow value at the next time interval $T+1$, X^k_{T+1} is the traffic flow value of the kth nearest neighbor at the time interval $T+1$, K is the total number of the neighbors.

A prediction function by the weighted average has the following form:

$$\hat{X}_{T+1} = \sum_{k=1}^{K} \frac{d_k^{-1}}{\sum_{k=1}^{K} d_k^{-1}} X^k_{T+1} \tag{7}$$

where d_k denotes the distance between the feature vector describing the current traffic data and the kth nearest neighbors.

We use the prediction function by the weighted average.

4. MapReduce implementation

The proposed model of traffic flow prediction uses a large amount of current and historical traffic flow data. To improve the efficiency of the proposed model, we implement it on the basis of MapReduce model [14] for distributed computing using Apache Spark engine [15].

MapReduce provides parallel processing of big amount of data in computing clusters. MapReduce model usually consists of three main steps: Map, Shuffle and Reduce. Figure 1 illustrates a computation flowchart of the proposed model based on MapReduce engine.

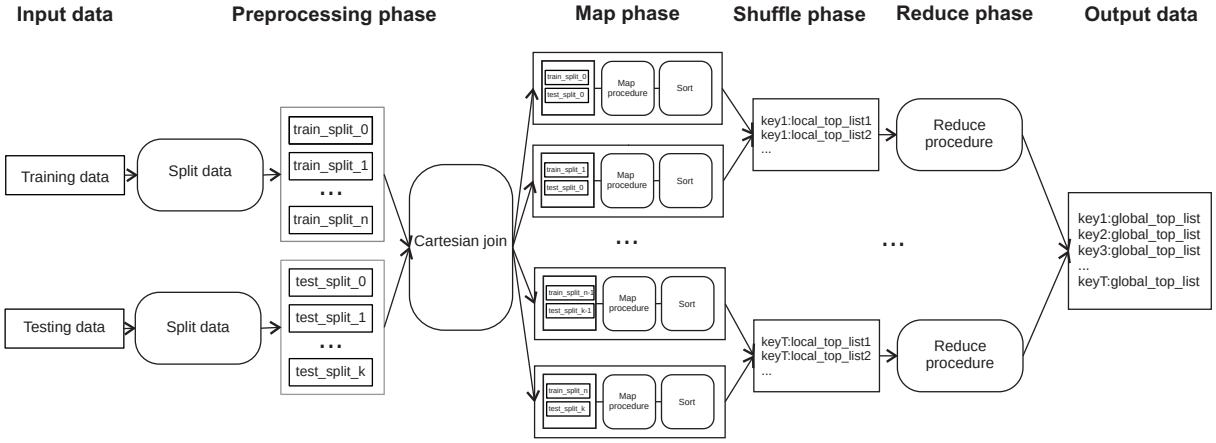


Figure 1. Data flow of MapReduce in the proposed model.

As illustrated in Figure 1, the first step is a preparation of input data for Map phase. At first, the historical and test data are divided into partitions. The optimal number of such partitions depends on the amount of processed data and the number of computing nodes. As mentioned in official Apache Spark documentation, the recommended value of partitions is 2-3 partition per CPU core in the cluster. Then, ordered pairs of historical and test data partitions are formed using the Cartesian product. Next, in the Map phase, a map function is applied to each pair of partitions. This function returns an intermediate set of key / value pairs - the test element / local list of k nearest neighbors. At the Shuffle phase, the key-value pairs are grouped and transferred to the reduce functions. At the final Reduce step for each test data element, the set of local k nearest neighbors lists is converted to the resulting (global) list of k nearest neighbors. The resulting lists of k nearest neighbors are subsequently used to find the predicted value traffic flow.

The results of evaluating the efficiency of the proposed model based on the MapReduce concept are presented in Section 5.

5. Experiments

In this work, in the experimental study, we predict average traffic speed in the city of Samara, Russia for short-term prediction horizon 10 minutes. The dataset contains records for 34 days. We compare the proposed model with the model described in [12]. This model uses feature vector in form (1), where the feature vector considers time domain and upstream / downstream road segments (denoted below as "TDUD"). Our model we denote as "Clusters" because the feature vector considers spatial-temporal correlations in graph clusters.

During testing, these models are performed on each day (test set) and the remaining days considered as a historical dataset (training set). Then the average performance across the full data set is calculated.

We conduct the experiments on an Apache Spark cluster. The traffic flow was predicted for a small area contained 698 road segments (Figure 2). Each road segment is considered as two edges with different directions. The total size of the dataset was 3.5 GB.

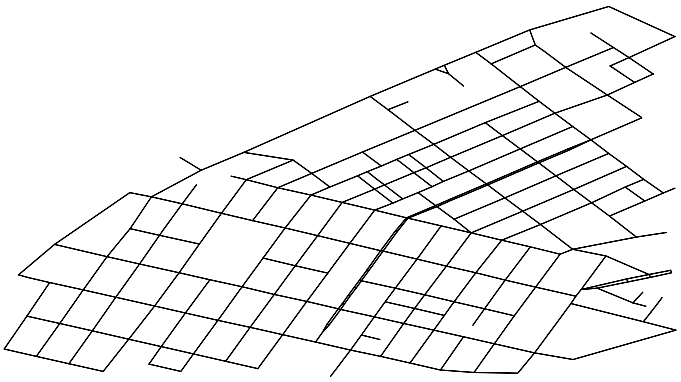


Figure 2. Samara city area.

To compare the performance of the proposed model, we use two standard metrics: mean absolute error (MAE) and mean absolute percentage error (MAPE) that can be formulated as:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}|V_t - \hat{V}_t| \tag{8}$$

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}\frac{|V_t - \hat{V}_t|}{V_t} \tag{9}$$

where V_t is the actual value of traffic flow at time interval t, $\hat{V}_t$ is the predicted value for the same time interval t, n is the total number of traffic flow observations.

Table 1. Algorithms comparison

	MAE	MAPE
TDUD	0.157	2.92
Clusters	0.149	2.82

Figure 3 and Figure 4 show the prediction result by the MAE and MAPE metrics for different days, respectively.

Based on the results above, we can conclude that considering the spatial-temporal correlation of the traffic flow in graph clusters allows improving the accuracy of the k nearest neighbor method.

6. Conclusion

The paper presents the distributed spatial-temporal model of short-term traffic forecasting based on the method of non-parametric regression k nearest neighbors. In the model, spatial and temporal characteristics of the transport flow in a compact cluster of the transport network are taken into account for the feature space description.

For distributed Big Data processing, we use MapReduce processing model implemented in the open source cluster-computing framework Apache Spark. Experimental analysis on real-world traffic data sets allows us to conclude that the proposed model has a high prediction accuracy and reasonable execution time, sufficient for real-time prediction.

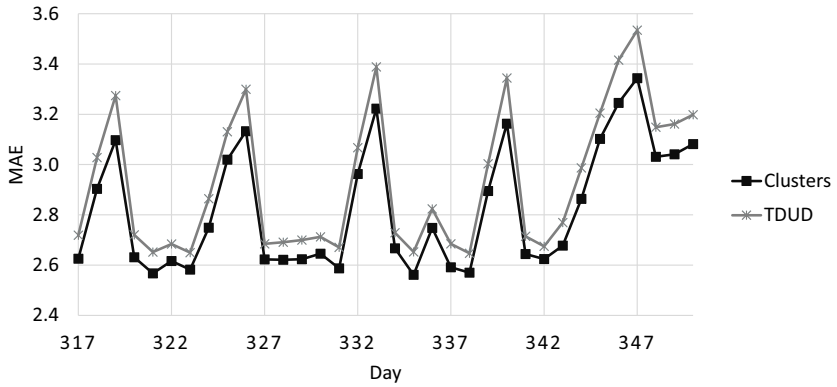


Figure 3. Mean absolute error.

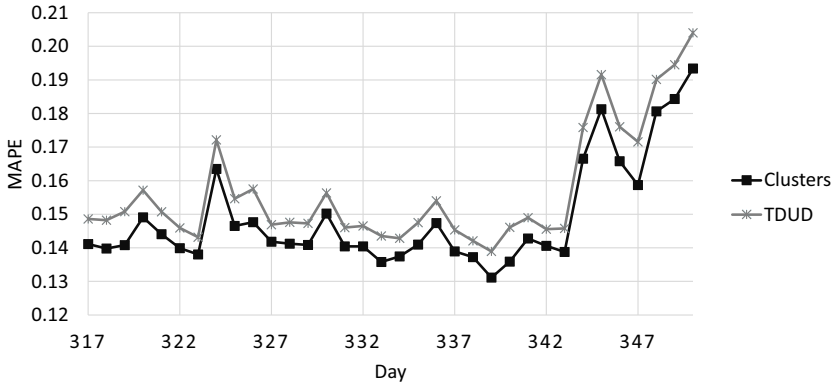


Figure 4. Mean absolute percentage error.

The possible direction of further research including dataset filtering for weekday / weekends traffic data and development of graph partitioning algorithms based on the traffic flow characteristics during a specific time period.

7. References

[1] Lana I, Del Ser J, Velez M and Vlahogianni E 2018 Road traffic forecasting: Recent advances and new challenges *IEEE Intelligent Transportation Systems Magazine* **10** 93-109
[2] Vlahogianni E, Golias J and Karlaftis M 2004 Short-term traffic forecasting: Overview of objectives and methods *Transport Reviews* **24** 533-557
[3] Karlaftis M and Vlahogianni E 2011 Statistical methods versus neural networks in transportation research: Differences, similarities and some insights *Transportation Research Part C: Emerging Technologies* **19** 387-399
[4] Shekhar S and Williams B 2007 Adaptive seasonal time series models for forecasting short-term traffic flow *Transportation Research Record* 116-125
[5] Smith B, Williams B and Keith Oswald R 2002 Comparison of parametric and nonparametric models for traffic flow forecasting *Transportation Research Part C: Emerging Technologies* **10** 303-321
[6] Yin H, Wong S, Xu J and Wong C 2002 Urban traffic flow prediction using a fuzzy-neural approach *Transportation Research Part C: Emerging Technologies* **10** 85-98

[7] Zheng Z and Su D 2014 Short-term traffic volume forecasting: A k-nearest neighbor approach enhanced by constrained linearly sewing principle component algorithm *Transportation Research Part C: Emerging Technologies* **43** 143-157
[8] Wu C H, Ho J M and Lee D 2004 Travel-time prediction with support vector regression *IEEE Transactions on Intelligent Transportation Systems* **5** 276-281
[9] Sun S and Zhang C 2007 The selective random subspace predictor for traffic flow forecasting *IEEE Transactions on Intelligent Transportation Systems* **8** 367-373
[10] Agafonov A and Myasnikov V 2015 Traffic flow forecasting algorithm based on combination of adaptive elementary predictors *Communications in Computer and Information Science* **542** 163-174
[11] Smith B and Demetsky M 1997 Traffic flow forecasting: Comparison of modeling approaches *Journal of Transportation Engineering* **123** 261-266
[12] Xia D, Wang B, Li H, Li Y and Zhang Z 2016 A distributed spatial-temporal weighted model on mapreduce for short-term traffic flow forecasting *Neurocomputing* **179** 246-261
[13] Lv Y, Duan Y, Kang W, Li Z and Wang F Y 2015 Traffic flow prediction with big data: A deep learning approach *IEEE Transactions on Intelligent Transportation Systems* **16** 865-873
[14] Dean J and Ghemawat S 2008 Mapreduce: Simplified data processing on large clusters *Communications of the ACM* **51** 107–113
[15] ApacheSpark 2018 (Access mode: `https://spark.apache.org/`)

Acknowledgments
This work was supported by the Russian Foundation for Basic Research (RFBR) grant 18-07-00605, grant 18-29-03135.

Compensation of the complex geometric distortions induced by a car digital video recording equipment

E V Myasnikov[1]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086

Abstract. In connection with the wide spread of mobile video recording equipment such as dashcam (car DVR), the restoration of distortions introduced by such equipment is of particular interest.Unfortunately, the nature of the geometric distortions introduced by such equipment is sometimescomplex, and cannot be described using the classical radial models. For this reason, in order to compensate for geometric distortions, an original method for compensating the geometric distortions was developed in this paper.The advantage of the proposed approach is, on the one hand, the possibility of a compact parametric description of the introduced distortions, and on the other hand obtaining at the final stage an improved correction field that takes into account the local inhomogeneities of distortions that occur when inexpensive systems are used.

1. Introduction

A number of models for estimating and compensating a radial distortion are described in the literature.These models include polynomial models, rational models, fisheye models and so on.In general, we consider the problem of correcting the radial distortion at the present time as an engineering problem. Nowadays this problem can be solved using calibration patterns and existing software tools (see, for example, [1, 2]).

In connection with the wide spread of mobile video recording equipment such as dashcam(car DVR) [3], the compensation of distortions introduced by such equipment is of particular interest. These systems often aim to expand the field of view of the camera in the horizontal direction, which allows recording information related to the area of interest (road, room, etc.), limiting the amount of unnecessary information outside the area of interest. This approach allows reducing both the amount of stored video information and the computational load on processors of mobile devices.

Unfortunately, the nature of the geometric distortions introduced by such equipment is sometimes complex, andcannot be described using the classical radial models. For this reason, in order to compensate for complex geometric distortions, an original method for compensating the geometric distortions was developed in this paper. The developed method is based on polynomial models and takes into account geometric distortions in general, regardless of their nature (whether they were induced by an optical system, hard- or software).

Thispaperisorganizedasfollows. The next section briefly reviews classical distortion models. The proposed approach for compensating the distortion induced by the considered mobile video recording equipment is described in Section 3. An experimental study of the proposed approach is given in Section 3. The paper ends up with the Conclusion and Refences.

2. Distortion models

The optical system of a modern camera, like many other optical systems, introduces nonlinear distortions [4]. It is accepted to distinguish the radial distortion associated with lens curvature and the tangential distortion associated with the inclination of the lens optical system with respect to the photosensitive sensor.

Radial distortion consists in shifting a point on the captured image towards or away from the optical axis with respect to its true position. If the displacement is directed towards the optical axis, the resulting distortion is called barrel-like distortion (Fig. 2a). Otherwise it is called pillow-like distortion (Fig. 2b).

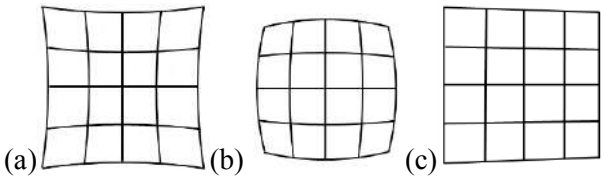


Figure 1. Image distortion: (a) radial pillow-like distortion; (b) radialbarrel-like distortion; (c) tangential distortion second picture.

The tangential distortion associated with the inclination of the sensor plane is expressed in projection distortions of the acquired image (Fig. 2c).

The presence of distortion not only leads to the curvature of straight lines, but also leads to gross errors in measurements taken with the captured image.

Elimination of distortion is usually performed in several stages:

1. The choice of a parametric distortion model,
2. Estimation of the parameters of the model,
3. Compensation of geometric distortions in accordance with the found parameters of the model.

It is essential that special calibration patterns with known parameters are used to estimate model parameters. Such patterns can contain a chessboard pattern, a uniform grid of lines, and so on (see Fig. 3). It is usually necessary to perform a series of shots of the calibration pattern to calibrate the camera using existing algorithms and software.

Figure 2. Calibration patterns often used for camera calibration: (a) uniform grid; (b) chessboard pattern [5]; (c) asymmetric circles pattern [6].

The literature describes a number of models used in the evaluation and compensation of radial distortion. Before proceeding to the description of models, let us introduce the following notations:

$x=(x,y,z)$ is the point of the object in the coordinate system associated with the object;

$u=(u,v)$ is the point on the image in the coordinate system associated with the image. The point of the object is to be projected to this point (ideal position);

$u_d=(u_d,v_d)$ is the point on the image in the coordinate system associated with the image. The point of the object is projected to this point due to geometrical distortions (real position).

Assuming that the center of distortion coincides with the origin, let us also introduce:

$$r = \sqrt{u^2 + v^2}, \quad r_d = \sqrt{u_d{}^2 + v_d{}^2},$$

where r is a distance from the origin to an ideally located point (u,v), r_d is a distance from the origin to the real point (u_d,v_d).

A quite complete review of models of radial distortion is given in [7]. The most well-known polynomial model of radial distortion (1) [8] is considered among the existing polynomial models. A special case of this model is a model with two coefficients (2) [9].

$$r_d = r(1 + k_1 r^2 + k_2 r^4 + k_3 r^6 + \ldots) \quad (1)$$

$$r_d = r(1 + k_1 r^2 + k_2 r^4) \quad (2)$$

For the latter model (2), the solution of the inverse problem is found in the form of an iterative scheme.

The model of a more general form (3) [10] and its particular case (4) [11] are also considered.

$$r_d = r(1 + k_1 r + k_2 r^2 + k_3 r^3 + \ldots), \quad (3)$$

$$r_d = r(1 + k_1 r + k_2 r^2). \quad (4)$$

For the latter model (4), the solution of the inverse problem is obtained, both with the use of approximating functions, and in the analytical form [7].

In rational models, the radial function has the form of a relation. Among the rational models, we can mention model (5) [12], and also a rational model based on the ratio of simple polynomials (6), proposed in [7].

$$r_d = \frac{\sqrt{1 + 4kr^2} - 1}{2kr} \quad (5)$$

$$r_d = r\frac{1 + k_1 r + k_2 r^2}{1 + k_3 r + k_4 r^2 + k_5 r^3} \quad (6)$$

Other models of radial distortion (for example, the fish-eye model) were proposed in the literature.

Tangential distortion is described using a model:

$$\begin{aligned} u_d &= 2p_1 uv + p_2(r^2 + 2u^2) \\ v_d &= 2p_2 uv + p_1(r^2 + 2v^2) \end{aligned} \quad (7)$$

It should be noted that methods for blind identification of radial distortion exist.

A number of works related to the evaluation of the quality of the above models are known from the literature. From the analysis of such literature, it can be concluded (in particular, in [7]) that the use of models containing more degrees of freedom (polynomial and rational models containing polynomial functions of higher orders) makes it possible to obtain more accurate results in the sense of the approximation quality. On the other hand, it was noted that using more degrees of freedom (functions of higher orders) can lead to instability when using numerical methods to estimate the coefficients.

3. Methods

As it was outlined in the Introduction, in this paper we propose a technique for correction of the complex geometric distortion induced by the considered mobile registration device.The developed method is based on the use of the video recording of the calibration pattern (uniform grid with a sufficient size). The developed method consists in the estimation of the field of correction of geometric distortions using a frame with the subsequent aggregation over all evaluated frames. Actually, the estimation of the field of correction of geometric distortions consists of the following steps:

- Filtering the horizontal lines of the calibration grid using the Gabor filter bank;
- Search for the spanning lines of the calibration grid in the horizontal direction;
- Approximation of horizontal lines of the calibration grid using polynomials;
- Construction of the field of correction based on horizontal lines by interpolation of the received polynomials;
- Filtering the vertical lines with the subsequent search for the spanning lines, and approximation of vertical lines of the calibration grid using polynomials;
- Evaluation of the correction field for the frame by interpolation of the received polynomials.

An example of the distorted calibration grid is shown in Figure 3 (a). The image is slightly cut from below to exclude the influence of auxiliary inscriptions. As it can be observed, the distortion has quite a complex nature. Horizontal lines are substantially convexwith relation to the center of the

image(barrel-like), while vertical lines change from slightly concave (pillow-like) to almost straight. In addition, local non-uniformities of the lines are noticeable.

At the first step of the method, for the input grayscale image $I(x,y)$, we obtain filtered images $I_1^H(x,y)$ and $I_1^V(x,y)$ by using Gabor filter banks, which are tuned to distinguish horizontal lines (with orientation equal to 45, 67.5, 90, 112.5, and 135 degrees), and vertical lines (with orientation equal to -10, 0, 10 degrees). The examples of the filtered images are shown in Figure 3 (b,c).

At the second step of the method, we perform the search of the local maxima of filtered images, which may correspond to the spanning lines of the calibration grid:

$$I_2^H(x,y) = \begin{cases} I_1^H(x,y), & if\ \left(I_1^H(x,y) = \max\limits_{i=-h..h} \left\{ I_1^H(x,y+i) \right\} \right) and\ \left(I_1^H(x,y) > t \right) \\ 0, & otherwise \end{cases},$$

$$I_2^V(x,y) = \begin{cases} I_1^V(x,y), & if\ \left(I_1^V(x,y) = \max\limits_{i=-w..w} \left\{ I_1^V(x+i,y) \right\} \right) and\ \left(I_1^V(x,y) > t \right) \\ 0, & otherwise \end{cases}$$

Then the tracing of horizontal and vertical lines is performed, and the sets $L_i^H = \left\{ \left(x_{i,j}^H, y_{i,j}^H \right) \right\}_{j=1..n_i^h}, i = 1..N_H$ and $L_i^V = \left\{ \left(x_{i,j}^V, y_{i,j}^V \right) \right\}_{j=1..n_i^v}, i = 1..N_V$ of the pixels corresponding to the vertical and horizontal lines are formed. Here N_H is the number of horizontal lines, N_V is the number of vertical lines, n_i^h is the number of pixels in the horizontal line L_i^H, and n_i^v is the number of pixels in the vertical line L_i^V.

Approximation of the horizontal L_i^H and vertical L_i^V lines is performed by the polynomials of degree k:

$$P(x) = p_0 + p_1 x + \ldots + p_K x^K$$

so that the coefficients $p_{i,k}^H$ and $p_{i,k}^V$, k=1..$K$of the polynomials are optimal in a least squares sense:

$$\sum_{(x,y) \in L_i^H} \left(P_i^H(x) - y \right)^2 \xrightarrow[p_{i,k}^H]{} \min, \quad i = 1..N_H$$

$$\sum_{(x,y) \in L_i^V} \left(P_i^V(y) - x \right)^2 \xrightarrow[p_{i,k}^V]{} \min, \quad i = 1..N_V$$

The examples of the approximation by polynomials are shown in Figure 3 (d, e).

After construction of the distortion model, the correction of an input frame can be performed using the interpolation of the coefficients of the polynomials of the constructed model. For example, in the case of linear interpolation, the interpolated polynomial $P()$ has the coefficients

$$p_k = \lambda p_{i,k} + (1-\lambda) p_{i+1,k}, \quad k=1..K.$$

Here polynomial $P_i()$ correspond to the bottom (left) edge of the considered cell, and polynomial $P_{i+1}()$ correspond to the top (right), $\lambda=(S-d)/S$, where S is a linear size of the grid cell, and d is the distance to the bottom (left) edge of the cell.

It is worth noting that horizontal and vertical lines do not necessary have to be approximated by the polynomials of the same order.

Here we suppose that the optical axis is approximately orthogonal to the plane of the grid pattern and the orientation of the camera and pattern is approximately known.

The advantage of the proposed approach is, on the one hand, the possibility of a compact parametric description of the introduced distortions in the form of the coefficients of polynomials, and on the other hand obtaining at the final stage an improved correction field that takes into account, among other things, the local in homogeneities of distortions that occur when inexpensive systems are used.

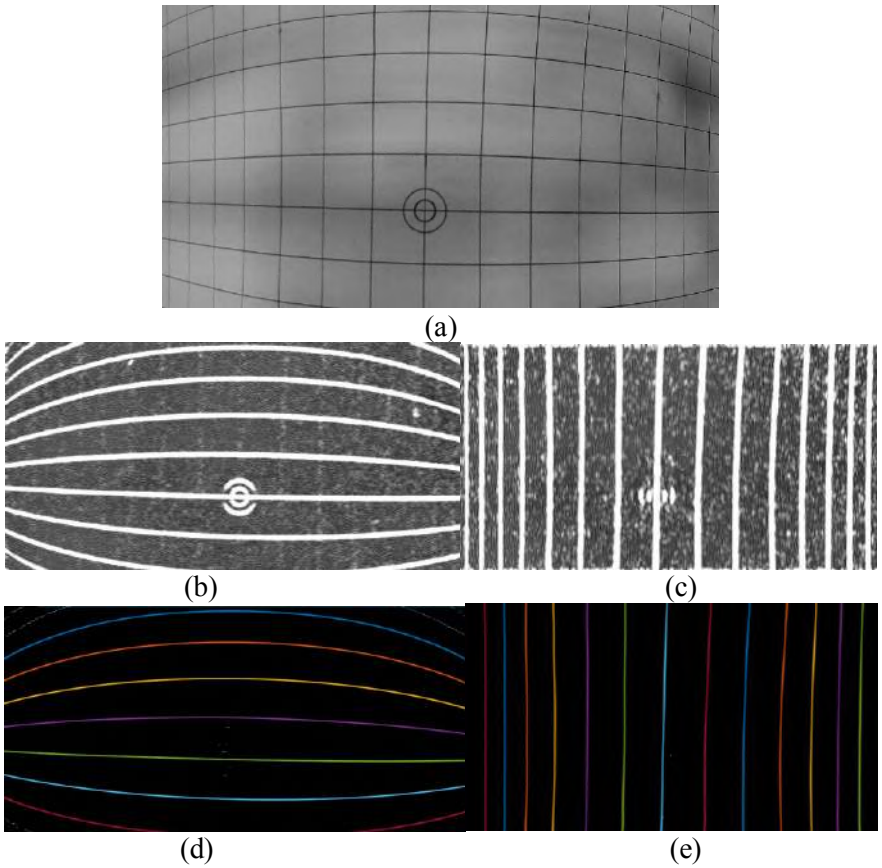


Figure 3. An example of the distorted calibration grid (a); the results of the Gabor filtering (b, c); the approximation of lines by polynomials (d, e).

4. Experiments

In this section we provide experimental results of the developed technique. The experimental setup contained the car digital video recorded system equipped with a wide-angle camera. The characteristics of the equipment are as the following: 1920×1080 pixelsresolution, 30 frames per second, 170 degreesview angle (diagonal, as stated by the manufacturer), MOV / H.264codec.

To evaluate the developed technique, we used 20 frames of the video, which were taken with 10 frames interval (total approximately 7 seconds of the observation of the grid pattern). For each frame we performed the following operations:

- Undistortion of the image according to the technique described in the previous section;
- Filtering the horizontal and vertical lines of the calibration grid separately using the Gabor filter bank;
- Searching for the spanning lines of the calibration grid using the correlation approach separately in the horizontal and vertical directions;
- Search for the crosses of the vertical and horizontal lines;
- Evaluation of the projective transformation that compensates for the mutual mismatch between the recording device and the calibration pattern;
-Transformation of all the points of the vertical and horizontal lines using the projective transformation found at the previous step;
- Evaluation of the mean squared error over all the projected points with respect to their ideal positions.

The dependence of the average reconstruction error evaluated using the above scheme on the degree of the polynomials is shown in Figure 4. The estimated error in restoring geometric distortions is approximately 0.7 pixels. The best value was obtained for the polynomials of the 10 degree. The reported results were obtained for the model estimated using only one frame without aggregation.

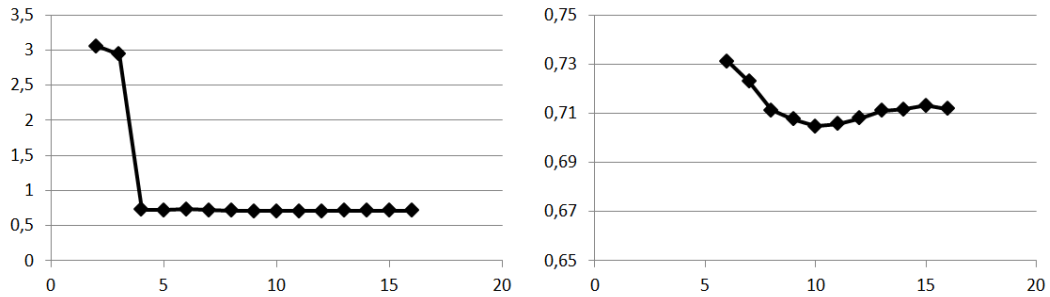


Figure 4. The dependence of the reconstruction error on the degree of the polynomials.

5. Conclusion

A new technique is proposed in this paper for the correction of the complex geometric distortion introduced by the mobile video registration device.

For the model constructed using only one frame, the estimated reconstruction error is approximately 0.7 pixels. The advantage of the proposed approach is, on the one hand, the possibility of a compact parametric description of the induced distortions in the form of the coefficients of polynomials, and on the other hand obtaining at the final stage an improved correction field that takes into account, among other things, the local inhomogeneities of distortions that occur when inexpensive optical systems are used.

The future directions of research include the experiments with different mobile video recording equipment, increasing the efficacy of the developed method, and the use of dynamic targets for the evaluation of geometric distortions.

It is assumed that the results of research can be used potentially to improve the recognition of vehicles and road signs [13], as well as to improve the quality of road assessment and traffic planning [14].

6. References

[1] *Camera calibration with OpenCV* (Access mode: https://docs.opencv.org/2.4/doc/tutorials/calib3d/camera_calibration/camera_calibration.html)
[2] *Camera Calibrator* (Access mode: https://www.mathworks.com/help/vision/ref/cameracalibrator-app.html)
[3] *Dashcam* (Access mode: https://en.wikipedia.org/wiki/Dashcam)
[4] Myasnikov E 2017 Technique for processing imagery of plans and drawings obtained using digital camera *Procedia Engineering* **201** 353-362
[5] *Open CV chessboard calibration pattern* (Access mode: http://docs.opencv.org/2.4.1/_downloads/pattern.png)
[6] *Open CV asymmetric circles grid* (Access mode: http://docs.opencv.org/2.4.1/_downloads/acircles_pattern.png)
[7] Ma L, Chen Y and Moore K L 2004 Rational radial distortion models of camera lenses with analytical solution for distortion correction *International Journal of Information Acquisition* **1(02)** 135-147
[8] Slama C C 1980 *Manual of Photogrammetry* (American Society of Photogrammetry, fourth edition)
[9] Zhang Z 1999 Flexible camera calibration by viewing a plane from unknown orientation *IEEE Int. Conf. on Computer Vision* 666-673
[10] Hartley R and Zisserman A 2000 *Multiple View Geometry* (Cambridge University Press)
[11] Ma L, Chen Y and Moore K L 2003 Flexible camera calibration using a new analytical radial undistortion formula with application to mobile robot localization *IEEE International Symposium on Intelligent Control*
[12] Brauer-Burchardt C and Voss K 2001 A new algorithm to correct fish-eye- and strong wide-angle lens-distortion from single images *Int. Conf. on Image Processing* 225-228

[13] Yakimov P Y 2015 Tracking traffic signs in video sequences based on a vehicle velocity *Computer Optics* **39(5)** 795-800 DOI: 10.18287/0134-2452-2015-39-5-795-800
[14] Agafonov A A and Myasnikov V V 2016 Method for the reliable shortest path search in timedependent stochastic networks and its application to GIS-based traffic control *Computer Optics* **40(2)** 275-283 DOI: 10.18287/2412-6179-2016-40-2-275-283

Acknowledgments
The reported study was funded by RFBR according to the research project №17-29-03190.

Hybrid face recognition solution for security

Y Donon[1]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086

Abstract. This article introduces a design that aims threw the combination of open source and closed source technologies, to make a, simple to implement,low-cost and high-performing face recognition solution. The solution provides identification, emotions and facial features recognition as well asdangerous objects spotting.This article exposes the concept of the solution, explains its importance on the market and provides details of a proof of concept prototype.

1. Introduction

The market of face and image recognition technologies is booming and forecasted a brilliant future. Although it is seen more and more in specialised magazine or promoted by giants of information technologies, many smaller actors are left behind as they perceive the technology as inaccessible or too expensive.

Numerous researchesaboutthose systems have been made in the recent times and during those years of research, computers science has evolved beyond measure.But what really have changed since a few years, are the cameras. What makes this ground of research more prolific than ever todays is that we all havephones in our pockets which sensors have an average of 14 megapixels, that we can buy full HD webcams for less than a hundred dollars. 15 years ago, a digital camera's resolution would be a fifth of what a webcam hasnow and be ten times its price. [1]

Although face recognitions attempts have been around for more than 50 years now, it still appears as a new technology to most of people. If we had indeed technologies able to perform those tasks back in the sixties, pictures would have to be taken according to very precise specifications. Attempts were multiplied; it became a trend in the nineties, some artefacts from that time, such as the ORL Database if Faces from Cambridge are even still in use today. In the beginning of the two-thousands, an international contest has even been thrown on the subject of face recognition. [2] Yet with all of that, it is only now and in the upcoming years that we really can and will perceive ground-breaking advance in those technologies. [3]

Nowadays, we have the tools, we have the necessary sensors for an efficient recognition and new actors on this market are emerging every day.Those solutions represent a trend on the security market of course; it allows to recognise not only people, but also specific objects and track them if necessary. The industry alsostarts to use emotion recognition systems to understand better their customer. [4] In this paper, I will introduce a solution to exploit this new market and make it accessible to everyone threw a low-cost, high performing face recognition solution for security. A design that is easy to deploy without high computing capabilities. The goal of this paper is for everyone to understand the stakes of this market, how accessible it is now and how it can be used in our everyday life.

2. Market and projections

2.1. Hybrid

As the market is still emerging but have been around for a long time, both open-source and closed source solutions exists. Closed source solutions are efficient to spot faces,can differentiate them, making an authentication possible. Those solutions, however, falls short when it comes to analyse a picture's details, such as emotions, facial details or objects. Closed source image recognition providers, however, are usually specialized and therefore extremely good when it comes to identifying those details. [4]

The design presented in this paper tries to take profit of this reality. Combining open source technologies and closed source ones, taking to both worlds what they are good at, allows making a first analysis on a local computer, even one having low computation capabilities and, over the internet, using solutions provided by the majors of image recognition, to analyse pictures in-depth,beyond the capacities of open source solutions.

2.2. Projection

As mentioned, the face recognition market is still emerging. It is expected to be worth between 7.5 and 10 billion dollars by 2022, 2 to 3 time more than it was worth in 2016. The year before that, the main client of those systems was US Homeland Security.By now the use of such solution for security has already spread in several countries and is used by such actors as the British police. Since its beginning this technology has been viewed as a major asset in security systems. [4]

Open source solutions are forecasted to improve their algorithms in 2D and thermal face recognition, while it is believed that online services will keep the specialized market (complex emotions, facial features details, 3D modelling, etc...) , although open source alternatives exists and will also improve, but not with the same precision rate. [5] The main uses between 2017 and 2022 are forecasted to be emotion recognition, tracking and monitoring, access control and law enforcement. [4] Therefore the design suggested in this paper fits the needs of the market to have an affordable solution, using the full capacities offered by the different actors of face recognitions solutions. It also is appropriate as thiscurrent is forecasted to be stable over at least the upcoming four years.

Making profitable for SMEs (Small and medium-sized enterprises), which are 98% of economic environment, a multi-billions digital economy market threw the design presented is a breakthrough for face recognition as it makes it an accessible tool.

3. Results

3.1. Functioning

In this design, if the picture is of sufficient quality for an optimal analysis[6], the system queries first a micro database of a handful of the most recent faces, loaded on the computer's RAM (1). This reduces the load on the disk's database and accelerates the program, as between two frames, it is usually the same faces that show up.

If the face hasn't been recognized on the first database, a query is sent to the second one, which can store up to a thousand of faces, depending of the capacities of the computer (2). This database is typically conceived to store the faces of all the employees of a company and manage access controls.

If no match is found in the second database (the confidency of the comparison between the shown face and the existing ones is too low), the system querries online services, that can analyse the picture, confirm that the idividual in unknown via an online database (3), and differenciate its emotions, facial features as well as alaysing its environement, detecting immediate threats, such as weapons.

Finnaly, the result of this detection will be added to the RAM-loaded database to avoid detecting and alaysing again the same face (4). Each query to an online service having of course a cost.

3.2. Results

The performance reached by the test program fitted all of our expectations, if sometimes the description suffers small imprecisions it offers a real-time identification on video with 5 frames per

second, spotting simultaneously several object[7], more than enough for a security camera, giving even an impression of relative fluidity in the capture. With identified facial features such as hair colour or emotions, a very precise recognition differentiating identical twins without any hesitation and beingable to detect some specific object such as weapons, we can say that, on a technical point of view, the performance test of the design is a complete success.

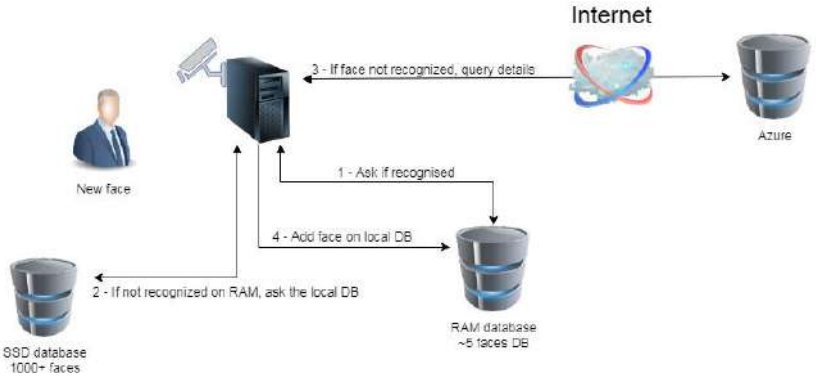


Figure 1. Design's business process.

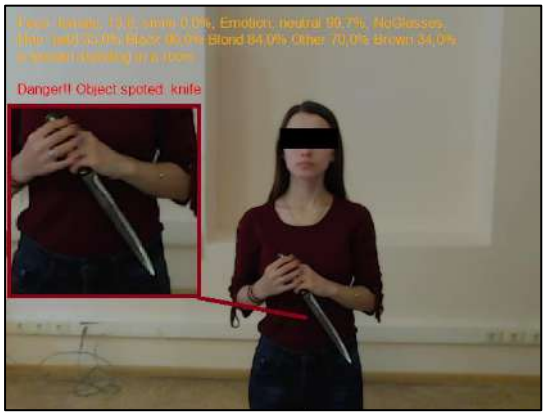


Figure 2. This image illustrates the analysis of a person caught on camera. Some of the information, such as the approximation of the age are not correct, however, the capture allows a clear identification (the program select the frame having the "best quality of face").

Figure 3. Assets that although the quality on an image might be poor, the program is able to extrapolate correct information.

Table 1. Features analysis for figure 2 and 3.

Feature	Values figure 2	Precision (appreciation if data unavailable) figure 2	Values figure 3	Precision (appreciation if data unavailable) figure 3
Facial	13.8	72%	20	97%
Smile	0.0%	Correct	0.0%	Correct
Emotion	Neutral 99.7%	Correct	Neutral 98.7%	Correct
Glasses	No glasses	Correct	Reading glasses	Correct
Hair	Bald 33%	15%	Bald 33%	15%
Hair	Black 99%	85%	Black 100%	Error
Hair	Blond 84%	10%	Blond 53%	Correct
Hair	-	-	Brown 42%	Correct
Hair	Other 70%	-	Other 38%	-
Description	A woman standing in a room	Correct	A woman in a blue shirt	Correct
Object	Knife	Correct	gun	Correct

Although some obvious progress are to be made on the hair colour detection, the features calculated are generally close to reality and most importantly allowsa human identification of a person, even without a the subsequent picture.

3.3. Technical specifications

A software has been developed as a design proof of concept. It has been developed in C#, using an OpenCV wrapper for this platform, OpenCvSharp, Fisherfaces recognition algorithm and Microsoft's Face and Vision API.

The use of Fisherfaces has been motivated over other methods for its search of discrimination criteria, which is more reliable to exclude possible faces match, enhancing the security offered by the solution. We widely favour a false negative, which leads to a control on the server that the person truly isn't identified in our database, than a false positive, which would allow an intruder to get through the system. [8-11]

The use of the Microsoft cognitive systems has been decided as it fitted the technical needs of the environment, offered a good transparency and as they send back details from the analysis of the image such as face coordinates, allowing further extrapolation. The other considered providers whichbilling systems were adapted to this design were Google Cloud Platform and IBM Watson.

The goal being to make the market as accessible as possible, it was important to reduce every source of costs. The system has been tested on several Microsoft Windows platforms (Win 7 and superior versions), it function and manage real-time recognition on computers having 4Gb of RAM, a dual core processor and a SSD of 64 GB,inferior configurations haven't been tested.

3.4. Costs

The design described here is of course flexible, meaning any online service could be used alternatively to Microsoft's. The calculation of costs for such an access control system was made considering an arbitrary a company size of a hundred workers (big company on the SME environment). Considering each of the employees comes into the company's building twice a day every working day, it is 4000 controls a month. If those faces are all stored locally, they should be recognized and therefore not generate any cost. If every day fifty unknown person comes into the building it will make about a thousand controls that are not perceived by the local recognition system. Those numbers all falls under the "free calls pool" of Microsoft Azure subscription, even considering that some queries of the analysis must be done in several steps, generating as many calls. However, this represents a laboratory reality which always differs from the "field". For the same amount of people, used in a production environment, the price of the online analysis has been calculated to be about 10 to 15 dollars a month, taking in account all the frequent errors of the software. [12]

As a onetime cost, it is necessary to get a small computer and a webcam to run the software. Multiple devices have been assessed on that purpose, all in a price range of 250 to 350 dollars for the computer and as for the camera between 50 and 80 dollars. For a total cost of 300-430 dollars a door. Counting the cost of electricity to power the system, the total cost of the installation is estimated to 1500 dollars for a period of 5 years (total cost of ownership).

4. Reliability

The test program realized to proof this design is able to distinguish similar faces such as twins easily. The confidence criterion has been configured severely, to make sure the local recognition system wouldn't give any false positive. This confidence has been set according to previous researches. [13] Tests have been repeated several times on thousands of frames without any mistake from the software.

To assess the efficiency of the software, some further tests and comparison have been conducted. The computer has been presented pictures from five pairs of twins identified in the database and two pairs of pictures of the same person on different pictures and has to differentiate them. Humans, on the other hand, have been presented a similar set of pictures and were simply asked, having two seconds for each picture to tell which subjects were twins and which were not. [14]

The precision of the software couldn't be assessed with accuracy, as so far, the program hasn't been cheated on successfully. Whenever the confidence of the local face analyser is too low, faces are

sent online for analysis. Since the program is at its final stage of development, the success rate has been of a hundred percent. Therefore, the upcoming paragraph, assessing the reliability of such systems, is based on external information and other systems.

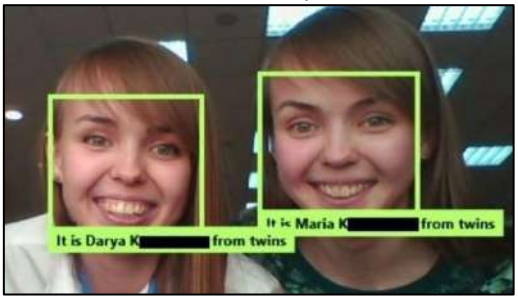


Figure 4. Twins differentiation.

Figure 5. The common test sample between the computer and humans for twins' assessment.

Table 2. Software control.

Feature	Value
Couple tested	7
Total frames	2500+
Accuracy	100%

Table 3. Human control.

Feature	Value
Human subjects	18
Couple tested	6
Total frames	72
Accuracy	61%

Unveiling its last iPhone, Apple claimed its face recognition system has a reliability of one in a million, meaning that once in a million times two faces would be confused and recognised as being the same, this is the closest comparison possible to do to the online services used. [14] To correlate this number we can take the code of a credit card in Russia, 4 digits or 10'000 possibilities, fingerprints , reputable unreliable once in 50'000 samples or an average home key (6 tumblers, 7 heights), which makes about 120'000 possibilities. Weather the reliability of the system is comparable to Apple's claim about its owns is discussable, but, nevertheless, the tests in laboratory are in favour of assessing a very high index of reliability forcomparable face recognition systems.

5. Conclusion

As underlined in this presentation, face recognition is a fast-developing market at the moment, much is already done but much is left to be built and this design has a place in the development of the market. Every major actor involved in security should now consider getting themselves an access to this kind of technology, especially now it is more accessible than ever and as the market trend makes it very profitable.

In the future, the detection will be improved by assessing the liveness of faces. Checking that we are not given a picture of a face but that it is a genuine face we have in front of the camera. This can be made by different methods, but the most adapted to a system of those dimensions is the analysis of the micro-behaviour of the eyes. [16-19] The identification system on RAM will also be compared in efficiency to a YOLO system (You Only Look Once) in order to asset their respective efficiency and choose the most appropriate technology to keep a target acquired and analyse it only once.

This kind of system could also be used on security cameras to get frames with a higher resolution and filter them threw an artificial intelligence able to understand which frames are relevant by an analysis of the pictures. Allowing selecting only relevant frames for storage, gives the possibility to significantly augment the quality of the camera's captors without being confronted to the problem of the storage space saturation. Emotion recognition and specifically this design can be adapted to the numerous of other uses such as home automations, alarms, research of wanted persons and many others that haven't been mentioned in this article. It is up for everyone, on this new market, to develop their own ideas.

Of course, this paper wasn't about a purely technical breakthrough, however I hope that the reader understands better now the face recognition market, how to use it efficiently and make it profitable, in particular with the design offered.This kind of design will make the difference between an emerging market and a fully grown and accessible one, bringing a new technology to the consumer. In other words, I want everyone to understand how face recognition systems are now in the reach of their hands.

6. References

[1] *Digital Photography review* (Access mode: https://www.dpreview.com/articles/5778663183/ten-unique-cameras-from-the-dawn-of-consumer-digital-photography) (20.8.2013)

[2] Philips P J, Flynn P J, Scruggs T, Bowyer K W, Chang J, Hoffman K, Marques J, Min J, Worek W 2005 Overview of the face recognition grand challenge, Computer Vision and Pattern Recognition *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* DOI: 10.1109/CVPR.2005.268

[3] Zhao W, Chellappa R, Philips P J, Rosenfeld A 2003 Face recognition: A literature survey *ACM Computing Surveys* **35** 399-458

[4] Gates K A 2011 *Our biometric future: facial recognition technology and the culture of surveillance* (New York University press) p 263

[5] Rybintsev A V, Konushin V S and Konushin A S 2015 Consecutive gender and age classification from facial images based on ranked local binary patterns *Computer Optics* **39(5)** 762-769 DOI: 10.18287/0134-2452-2015-39-5-762-769

[6] Nikitin M Yu, Konushin V S and Konushin A S 2017 Neural network model for video-based face recognition with frames quality assessment *Computer Optics* **41(5)** 732-742 DOI: 10.18287/2412-6179-2017-41-5-732-742

[7] Protsenko V I, Kazanskiy N L and Serafimovich P G 2015 Real-time analysis of parameters of multiple object detection systems *Computer Optics* **39(4)** 582-591 DOI: 10.18287/0134-2452-2015-39-4-582-591

[8] Jaiswal S, Bhadauria S S and Jadon R S 2011 Comparison between face recognition algorithm Eigenfaces, Fisherfaces and Elastic Bunch Graph Matching *Journal of Global Research in Computer Science* **2(7)** 187-193

[9] Yang M-H 2002 Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition Using Kernel Methods *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition* 215-220

[10] Turk M A and Pentland A O 2002 *Face recognition using Eigenface* (The Media Laboratory MIT)
[11] Kalinovskii I A and Spitsyn V G 2017 Review and testing of frontal face detectors *Computer Optics* **40(1)** 99-111 DOI: 10.18287/2412-6179-2016-40-1-99-111
[12] *Microsoft's Computer Vision API Version 2.0 documentation, Microsoft* (Access mode: https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/home) (22.8.2018)
[13] Vizilter Yu V, Gorbatsevich V S, Vorotnikov A V and Kostromov N A 2017 Real-time face identification via CNN and boosted hashing forest *Computer Optics* **41(2)** 254-265 DOI: 10.18287/2412-6179-2017-41-2-254-265
[14] *How secure is Face ID?* (Access mode: https://www.macworld.co.uk/feature/iphone/how-secure-is-face-id-3663992/) (01.11.2018)
[15] Z Caplova, Obertov Z, Gibelli D M, Mazzarelli D, Fracasso T, Vanezis P, Sforza C and Cattaneo C 2017 The Reliability of Facial Recognition of Deceased Persons on Photographs *Journal of Forensic Sciences* **62** 1286-1291
[16] Pan G, Wu Z and Sun L 2008 Liveness detection for face recognition, recent advances in face recognition *IntechOpen* 9 DOI: 10.5772/6397
[17] Blanz V and Vetter T 2003 Face recognition based on fitting a 3D morphable model *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25(9)**
[18] Kosinski M and Wang Y 2018 Deep neural networks are more accurate than humans at detecting sexual orientation from facial images *Journal of Personality and Social Psychology* **114(2)** 246-257
[19] Pan G, Sun L and Wu Z 2017 Eyeblink-based Anti-Spoofing in Face Recognition from a Generic Webcamera *IEEE 11th International Conference on Computer Vision* DOI: 10.1109/ICCV.2007.4409068

Reduction of the computational complexity of stochastic gradient algorithms of image parameters estimation for a priori optimization of the local sample volume

A G Tashlinskii[1], M G Tsaryov[1] and D G Kraus[1]

[1]Ulyanovsk State Technical University, Severny Venets str. 32, Ulyanovsk, Russia, 432027

Abstract. At stochastic gradient estimation of image parameters the estimates convergence character and computational expenses essentially depend on image samples local sample size used for obtaining the stochastic gradient. In the paper the possibility of a priori optimization of the volume of a local sample to minimize computational costs at geometrical images deformations estimation is considered. The minimum of the given computational costs for the conventional unit of expectation of the improvement of the evaluation is chosen as an optimization criterion. The block diagram of one of the algorithms and the examples of calculation results are presented.

1. Introduction

Recently, systems in which the initial information is a dynamic data array represented in the form of images have become widespread because such a presentation has visibility, compactness and information capacity.

The methods of image parameters estimation are based, as a rule, on four approaches. This is a comparison of image fragments, spatial-temporal filtering, morphological analysis [1], and analysis of the optical flow. For large-scale images, processing with stochastic gradient algorithms (SGA) [2-5], based on the analysis of the optical flow, is effective. In this case, the vector of estimates $\hat{\vec{\alpha}}$ of the investigated parameters $\vec{\alpha}$ is formed iteratively:

$$\hat{\vec{\alpha}}_t = \hat{\vec{\alpha}}_{t-1} - \mathbf{\Lambda}_t \vec{\beta}_t \left(\mathrm{Q} \right), \tag{1}$$

where $\mathbf{\Lambda}_t$ – gain matrix, $\vec{\beta}_t$ – stochastic gradient of the objective function Q of the estimation quality, $t = \overline{1,T}$ – iteration number; $\hat{\vec{\alpha}}_0$ – initial approximation of parameters. The algorithms are recurrent, have good accuracy characteristics and high speed of execution, do not require a preliminary evaluation of the images parameters, applicable to images with smoothly varying heterogeneity. The parameters estimated by the algorithm converge to the optimal values under rather weak conditions [6] and are resistant to impulse noise [7].

The study of the temporal dynamics of images requires analysis of the parameters of sequences of image frames, in particular, interframe geometric deformations of, for example, images $\mathbf{Z}^{(1)} = \left\{ z_{\vec{j}}^{(1)} \right\}$ and $\mathbf{Z}^{(2)} = \left\{ z_{\vec{j}}^{(2)} \right\}$, where $\vec{j} \in \Omega$, Ω – samples grid. As the objective function is chosen the mean square of the interframe difference in the estimation of interframe deformations under conditions when

it is possible to neglect the brightness distortions, or the sampling coefficient of interframe correlation at interframe brightness distortions close to linear [8].

The key problem is to increase the speed of the SGA. Various approaches are being explored to solve this problem. In particular, in [9] a procedure for stochastic gradient optimization of the second order in a linear time was proposed, in [10] to accelerate the optimization process using algorithms based on stochastic gradient descent the Nesterov moment method is applied, in [11] a convolutional neural network is used to estimate the geometric mismatch parameters between two images in accordance with a given geometric model (the affine model and the thin-plate spline transformation are studied), in [12] the acceleration of the stochastic gradient descent is achieved by taking into account the probability of smoothness of separate areas of the image. Also one of the approach is to reduce the volume μ of a two-dimensional local sample $Z_t = \left\{ z_{\vec{j}t}^{(2)}, \tilde{z}_{\vec{j}t}^{(1)} \right\}$. It is used at each iteration of the estimation to find the stochastic gradient $\vec{\beta}(\mathrm{Q})$ of the objective function, where $z_{\vec{j}t}^{(2)} \in \mathbf{Z}^{(2)}$, $\tilde{z}_{\vec{j}t}^{(1)} = \tilde{\mathbf{Z}}^{(1)}$; $\tilde{\mathbf{Z}}^{(1)}$ – oversampled image $\mathbf{Z}^{(1)}$ using the current estimates $\hat{\vec{\alpha}}_{t-1}$ of the interframe deformations parameters. But in so doing, the possibilities of a priori and a posteriori optimization of the local sample volume according to various optimality criteria have been poorly investigated. In this paper we consider the possibility of a priori optimization of the local sample volume by the criterion of minimum computational costs when estimating one parameter.

2. Optimizing the local sample volume

Suppose that, in accordance with the given error in the estimation of the parameter α, the mismatch $\varepsilon = \alpha_{\text{ev}} - \hat{\alpha}$ of the parameter estimate $\hat{\alpha}$ and its exact value α_{ev} should change from $\varepsilon_{\max}$ to $\varepsilon_{\min}$. Consider the possibility of minimizing the computing costs of the SGA by optimizing the volume of the local sample for each iteration of the estimation for the given conditions. We use the following optimality criterion.

At each t-th iteration of the stochastic gradient estimation, we will search for such a volume μ_t of the local sample that provides the minimum of the computational cost per the conditional unit of the mathematical expectation of an parameter estimate $\Delta\alpha_t$ improvement

$$\mu_t = k\big|_{\min \frac{g(k)}{\Delta\alpha_t}},\ k = 1,\ 2,\ \ldots, \tag{2}$$

where $g(k)$ – the computational cost of implementation by the algorithm of the t-th iteration for a local sample volume equal to k; $g(k)/\Delta\alpha_t$ characterizes the computational costs, normalized to the conditional unit of the mathematical expectation $\Delta\alpha_t$ of an estimate improvement (an expression for the calculation $\Delta\alpha_t$ using relay type of stochastic gradient estimate sets out later in this paper).

Because at the estimation iterations the parameter mismatch consistently changes from $|\varepsilon_{\max}|$ to $|\varepsilon_{\min}|$, then for the T iterations the proposed criterion will provide the minimum total computational costs

$$G = \sum_{t=1}^{T} g(\mu_t), \tag{3}$$

where T – the number of iterations required to perform the condition $|\varepsilon_T| < |\varepsilon_{\min}|$; ε_T – the mismatch of the parameter estimate $\hat{\alpha}$ and its exact value at the T-th iteration.

A detailed analysis of computational costs requires consideration not only the features and structure of the calculated ratio, but also many other influencing factors. These include the sampling time and conditions of images, the class of computing device, the time spent on the operations of addition, multiplication, division, access to memory, move data, and other auxiliary operations. Many of these factors depend on the specific image recording devices and the computers used. Therefore, in this

paper, the computational cost components will not be concretized. However, we will assume that the computational costs $g(\mu_t)$ of performing the SGA t-th iteration contain two components:

$$g(\mu_t) = g_{Zt} + g_o, \tag{4}$$

where g_{Zt} – computational cost for the formation of a local sample; g_o – other computing costs.

In this case, the cost of forming a local sample will be considered proportional to the volume μ of the local sample: $g_{Zt} = \mu \Delta g$, where Δg – computational costs for the formation of a local sample of a unit volume. Then

$$g(\mu_t) = \Delta g\left(c^{-1} + \mu_t\right), \tag{5}$$

where $c = \Delta g / g_o$ – the coefficient characterizing the proportion of the g_o when the volume of the local sample increases by one.

For relay SGA, the mathematical expectation of the improvement in the estimation $\Delta\alpha_t$ of the parameter under study by the t-th iteration can be found [13] by using the drift probability of the estimates [14]

$$\Delta\alpha_t = \mathrm{M}\left[\varepsilon_{t+1} - \varepsilon_t\right] = \left(\varepsilon_t - \Lambda_t\right)\rho^- + \varepsilon_t\rho^o + \left(\varepsilon_t + \Lambda_t\right)\rho^+ = \varepsilon_t + \Lambda_t\left(\rho^+ - \rho^-\right), \tag{6}$$

where ρ^+ – the probability that, for a given mismatch ε, the estimate $\hat{\alpha}$ will change toward the exact value of the parameter ($\operatorname{sign}\varepsilon_t = \operatorname{sign}\Delta\alpha_{t+1}$); ρ^- – the probability that, for a given mismatch ε, the estimate $\hat{\alpha}$ will change away from the exact value of the parameter, that is $\operatorname{sign}(\varepsilon_t) = -\operatorname{sign}\Delta\alpha_{t+1}$; ρ^o – the probability that the estimate will not change ($\Delta\alpha_t = 0$). Obviously, these probabilities constitute exhaustive events: $\rho^+ + \rho^o + \rho^- = 1$. We also note that, in the strict sense, the drift probability ρ^+ characterizes not the probability of improving the estimate, but the probability of changing the estimate in the "right" direction.

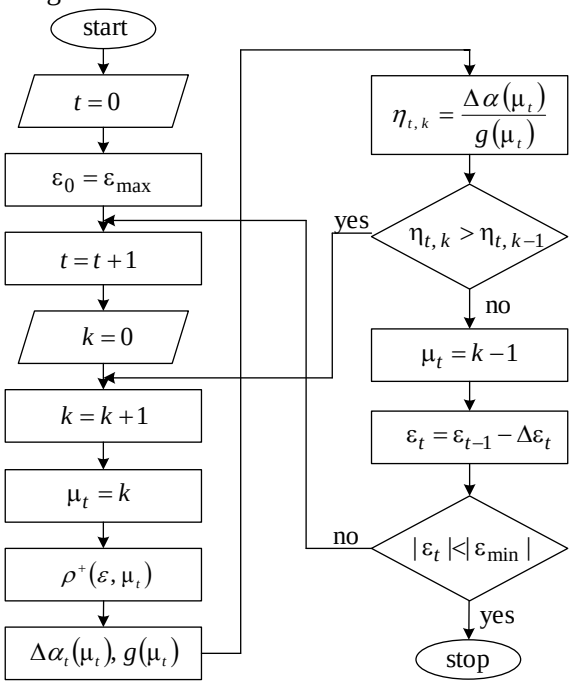


Figure 1. Block-diagram of the algorithm of the local sample volume optimization.

A block diagram of one of the possible algorithms for finding the optimal volume of a local sample are presented on figure 1. Here, for simplicity, it is assumed that $\rho^o = 0$, then $\rho^- = 1 - \rho^+$ and $\Delta\alpha_t = \Lambda_t\left(2\rho^+ - 1\right)$. To sequentially calculate the volume μ_t of the local sample at the t-th iteration,

$t=\overline{1,T}$, in the range of the deviation of the estimate from $|\varepsilon_{\max}|$ to $|\varepsilon_{\min}|$, the initial conditions are given $t=0$ and $\varepsilon_o=\varepsilon_{\max}$. Next, the volume μ_1 of the local sample is calculated, at which the minimum of the reduced computational costs (the minimum of the ratio $g(k)/\Delta\alpha_1$) is reached at the first iteration. Then, the local sample volume is estimated at the next iteration, i.e. is computed μ_2 and so on, until condition $|\varepsilon_t|<|\varepsilon_{\min}|$ is fulfilled.

3. Examples of calculating of the local sample volume

An example of computation results of optimal value of the local sample volume as a function of the mismatch is shown in figure 2. An interframe parallel image shift was evaluated. The value of the parameter c is chosen equal to 5%. Curve 1 corresponds to the noiseless images, and curve 2 corresponds to the signal-to-noise ratio $\sigma_x^2/\sigma_\theta^2=10$. It is assumed that the noise model of the researched images $\mathbf{Z}^{(1)}$ и $\mathbf{Z}^{(2)}$ is additive: $z_{\bar{j}}=x_{\bar{j}}+\theta_{\bar{j}}$, where $x_{\bar{j}}$ – image with dispersion σ_x^2, $\theta_{\bar{j}}$ – independent Gaussian noise with zero mathematical expectation and variance σ_θ^2.

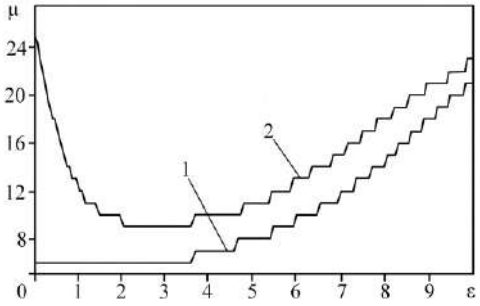


Figure 2. Relation the mismatch with the optimal of the local sample volume.

For the conditions corresponding to curves 1 and 2, the table 1 shows the results of the experiment. They show the loss in computing costs when using a constant volume of local sample ($\mu=const$) in comparison with the case of using the optimal volume of local sample. When specifying in SGA $\mu=const$, the value of μ corresponded to the average value μ_{avg} of the optimal sample size, and $\mu_{\text{avg}}-2$, $\mu_{\text{avg}}-1$, $\mu_{\text{avg}}+1$ и $\mu_{\text{avg}}+2$.

Table 1. Gain in computing costs, %

	$\mu_{\text{avg}}-2$	$\mu_{\text{avg}}-1$	μ_{avg}	$\mu_{\text{avg}}+1$	$\mu_{\text{avg}}+2$
Curve 1	4.1	3.9	3.9	4.5	4.9
Curve 2	2.1	1.7	1.8	2.4	2.9

4. Conclusion

The approach to increasing the rate of algorithms for stochastic gradient estimation of image parameters is considered on the example of inter-frame deformations estimation. The purpose is achieved by optimizing the size of the two-dimensional local sample used at each iteration of the estimation to determine the stochastic gradient. A priori optimization based on the criterion of minimum computational costs for the case of estimating one parameter is used. At the same time at each iteration of the estimation, the local sample size providing a minimum of computational costs for the conventional unit of the mathematical expectation of the parameter estimate improvement is determined. It is shown that, as the number of iterations increases, the mismatch modulus of the estimate and the exact value of the parameter decreases, the proposed approach provides a minimum of total computational costs. To determine the mathematical expectation of an improvement of the

studied parameter estimate, the probabilities of drift estimates (the probability of changing the estimates towards the exact value of the parameter and from it) are used.

The carried out modeling for one of the possible algorithms realizing the proposed approach confirmed the set purpose. So, for the given example of results, the gain in computational costs in comparison with the situation of using the constant volume of the local sample amounted to no less than 3.9% in the absence of noise, and not less than 1.8% for a signal-to-noise ratio equal to 10 (in variance). Thus, the proposed approach for algorithms of stochastic gradient estimation of image parameters makes it possible to determine the optimal size of the local sample for each estimation iteration, which ensures the minimization of computational costs.

5. References

[1] Rubi A Yu, Lebedev M A, Vizilter Yu V and Vygolov O V 2016 Morphological image filtering based on guided contrasting *Computer Optics* **40(1)** 73-79
[2] Su H R and Lai S H 2015 Non-rigid registration of images with geometric and photometric deformation by using local affine Fourier-moment matching *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition* pp 2874-2882
[3] Moritz P, Nishihara R and Jordan M I 2016 A linearly-convergent stochastic L-BFGS algorithm *Proc. of the 19th Int. Conf. on Artificial Intelligence and Statistics, AISTATS* pp 249-258
[4] Borisova I V, Legkiy V N and Kravets S A 2017 Application of the gradient orientation for systems of automatic target detection *Computer Optics* **41(6)** 931-937
[5] Taslinskii A G 2008 Optimization of goal function pseudogradient in the problem of interframe geometrical deformations estimation *Pattern Recognition Techniques, Technology and Applications* (Vienna: I-Tech) pp 249-280
[6] Tsypkin Ya Z 1995 *Information theory of identification* (Moscow: Fizmatlit) p 336 (in Russian)
[7] Tashlinskii A G 2007 Pseudogradient estimation of digital images interframe geometrical deformations *Vision Systems: Segmentation & Pattern Recognition* (Vienna: I Tech Education and Publishing) pp 465-494
[8] Tashlinskii A G 2008 The specifics of pseudogradient estimation of geometric deformations in image sequences *Pattern Recognition and Image Analysis* **18(4)** 701-706
[9] Agarwal N, Bullins B and Hazan E 2017 Second order stochastic optimization for machine learning in linear time *Journal of Machine Learning Research* **18(116)** 1-40
[10] Allen-Zhu Zeyuan 2017 Katyusha: the first direct acceleration of stochastic gradient methods *Proc. of the 49th Annual ACM SIGACT Symposium on Theory of Computing* pp 1200-1205
[11] Rocco I, Arandjelovic R and Sivic J 2016 Convolutional neural network architecture for geometric matching *Proc. CVPR* **2** 6148-6157
[12] Allen-Zhu Zeyuan, Richt´arik Peter, Qu Zheng and Yuan Yang 2016 Even faster accelerated coordinate descent using non-uniform sampling *Proc. of the 33rd Int. Conf. on ICML* **48(4)** 1110-1119
[13] Tashlinskii A G and Tikhonov V O 2001 The method for analyzing the error of pseudo-gradient measurement of multidimensional processes parameters *Izvestiya vuzov: Radioelektronika* **44(9)** 75-80 (in Russian)
[14] Tashlinskii A G and Voronov I V 2014 The probability of demolition of estimates of parameters of interframe geometric deformations of images under pseudo-gradient measurement *Izvestiya of the Samara Scientific Center of the RAS* **16 N6(2)** 612-615 (in Russian)

Acknowledgments

The reported study was funded by RFBR and Government of Ulyanovsk Region according to the research projects 16-01-00276 and 18-41-730006.

The technology of correction of dynamic distortions on mobile devices

E F Fatkhutdinova[1] and V A Fursov[1,2]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086
[2]Image Processing Systems Institute - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001

Abstract. We propose the technology of image correction in mobile devices based on the use of a parametric FIR filter. The filter is constructed using a frequency response specified in the form of a parametric family of centrally symmetric frequency characteristics. A model of a one-dimensional frequency response is used in the form of segments of three functions: parabola, constant, and exponential function. Within the framework of the proposed filter model, a mobile application has been created that implements two tuning schemes: identification of the filter parameters from the reference image and adjustment without a reference. Besides, it is possible to vary the frequency response parameter characterizing the mid-range interval, which provides additional possibilities for adjusting the quality of the recovery. An example of processing test images is given.

1. Introduction

Statistics of recent years show that the volume of traffic of mobile devices on the Internet is growing rapidly. A mobile application is a program that works on tablet PCs and smartphones. The number of smartphone owners around the world is growing every day. Music, working with photos, social networks - the most popular types of applications among smartphone owners [1]. One of the most popular functions of mobile devices is image registration[2-3]. Mass use of this function is associated with the ability to quickly register images in unexpected and unique situations. In this case, often you have to deal with problems due to defocusing and/or blurring (when the object being recorded is rapidly moving relative to the camera) [4]. Distortions such as blurring and defocusing are usually called dynamic. The task of correcting dynamic distortions in images recorded by mobile devices using the digital image processing algorithms implemented directly in the mobile device itself is extremely relevant.

Known classical approaches to the construction of filters for image processing are inverse filtering and Wiener filtering. When constructing an inverse filter, one often has to face the fact that the inverse operator does not exist or the corresponding transfer function has poles close to zero [5]. At the same time, noise is emphasized on the restored image. In the Wiener filter, this problem is overcome by taking into account in the filter transfer function the frequency characteristics of noise [6]. However, in practice these characteristics are often absent, and the synthesis of the optimal Wiener filter becomes a serious

problem. Unfortunately, other known methods of constructing filters, in one form or another, also face these problems [5], [6]. Therefore, often filters for the correction of dynamic distortions are built by parametric identification in the class of filters with finite impulse response (FIR filters). In [7], such a technology was considered, based on the use of a parametric model of the frequency response in the form of segments of quadratic and exponential functions. In this work, two variants of the technology were considered: identifying the parameters of the impulse response by use cases and setting the filter by indirect characteristics without a reference image.

In [8], the model of the frequency response from segments of quadratic and exponential functions was generalized, in particular, the region of medium frequencies was expanded due to the adding of a frequency response constant interval. On model examples it was shown that this modification improves the quality of image reconstruction in comparison with the previous realization of the quadratic exponential FIR filter. However, the implementation of a filter with an extended frequency response region requires, albeit insignificantly, an increase in computational resources. In this paper, we present the results of work on developing a mobile application that implements both of the above options.

The work is structured as follows. Section 1 describes the methods and algorithms for filter adjustment, both for the reference image and for the absence of a reference. In Section 2, the libraries used for constructing the mobile application, pseudo- code and user interface are explained. Section 3 provides examples of processing test images in a mobile device.

2. Method and algorithms

An FIR filter with a radially symmetric real frequency response [5, 6] is constructed with a supporting region D in the form $N \times N$ - a square with the center at the point $k_1 = 0$, $k_2 = 0$. Assuming that the central sample of the impulse response $h(0,0)$ of the reference region D is at point n_1, n_2 readings of the restored image $y(n_1, n_2)$ [5] can be represented as:

$$y(n_1, n_2) = \sum_{k_1=-\frac{N-1}{2}}^{\frac{N-1}{2}} \sum_{k_2=-\frac{N-1}{2}}^{\frac{N-1}{2}} h[r(k_1, k_2)] x(n_1 - k_1, n_2 - k_2), \quad (1)$$

where $r(k_1, k_2) = \sqrt{k_1^2 + k_2^2}$, and $h[r(k_1, k_2)]$ – counts of the one-dimensional impulse response defined on a set of circles with radii $r = r(k_1, k_2),\ k_1, k_2 \in D$.

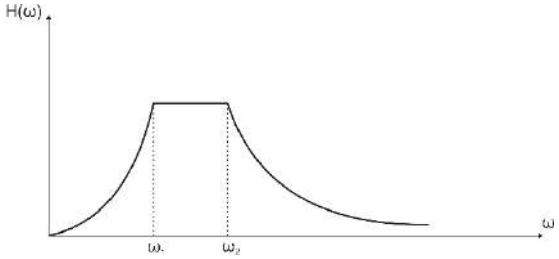


Figure 1. A typical graph of the spectrum of the GSE filter.

The one-dimensional frequency response function is used for all values $0 \le \omega < \infty$ in the form of three consecutive segments: parabola, constant and exponential function:

$$S(\omega) = \begin{cases} a\omega^2, in\ 0 \le \omega < \omega_1, \\ A = const = a\omega_1^2, in\ \omega_1 \le \omega \le \omega_2, \\ e^{-c\omega}, in\ \omega \ge \omega_2, \end{cases} \quad (2)$$

$$S(\omega_2) = a\omega_1^2 = e^{-c\omega_2}.$$

Filter, corresponding to the described frequency response, hereinafter referred to as a generalized square-law exponential filter (Generalized Square-Exponential) or briefly a GSE-filter. The graph of this function is shown in figure 1.

The impulse response corresponding to this spectral characteristic, by virtue of the radial symmetry property, is obtained as a function of the spatial parameter r inverse Fourier transform:

$$h(r)=\frac{1}{\pi}\mathrm{Re}\int_0^{\infty}S(\omega)e^{j\omega r}d\omega=\frac{1}{\pi}\mathrm{Re}\left\{\int_0^{\omega_1}a\omega^2e^{j\omega r}d\omega+\int_{\omega_1}^{\omega_2}Ae^{j\omega r}d\omega+\int_{\omega_2}^{\infty}e^{-c\omega}e^{j\omega r}d\omega\right\}= \quad (3)$$

$$=\frac{e^{-c\alpha\omega_1}}{\pi}\left\{\frac{\sin(\omega_1 r)}{r}\sin(\omega_1 r)+\frac{2\cos(\omega_1 r)}{\omega_1 r^2}-\frac{2\sin(\omega_1 r)}{\omega_1^2 r^3}+\frac{\sin(\alpha\omega_1 r)-\sin(\omega_1 r)}{r}+\frac{c\cos(\alpha\omega_1 r)-r\sin(\alpha\omega_1 r)}{c^2+r^2}\right\}$$

(we excluded a and ω_2, taking into account (2) substitution $a=e^{-c\omega_2}/\omega_1^2$ and take $\omega_2=\alpha\omega_1$).

It is easy to see that when $\alpha=1$ impulse response (3) coincides with the impulse response of the quadratic-exponential filter, which we considered in the previous paper [7]. In this paper we will build a mobile application in which the user can set this parameter (or the range of this parameter change) taking into account subjective requirements for the quality of recovery and available computing resources.

The counts of the two-dimensional impulse response, as in [8], are determined by discretizing the continuous function (3) for all directions corresponding to all samples of the reference region. For each sample (points k_1, k_2) of the support region in the ratio (3), the argument $r=r(k_1,k_2)=\sqrt{k_1^2+k_2^2}$. Since at *r=0* in (3) there is uncertainty, the value of the central reading is calculated as the sum of all samples except for the central one:

$$h(0,0)=\sum h(k_1,k_2),\ \forall k_1,k_2\in D, k_1,k_2\neq 0. \quad (4)$$

Then, all the samples in the reference area are normalized so that the requirement of maintaining the average brightness level of the processed image is fulfilled:

$$\sum h(k_1,k_2)=1,\ \forall k_1,k_2\in D. \quad (5)$$

The filter adjustment algorithm is constructed in the following sequence of steps:

1. Assigning the initial estimates of the parameters $\hat{\omega}_1, \hat{c}, \hat{\alpha}$ and criterion $Q_k(\hat{\omega}_1,\hat{c},\hat{\alpha})$ (with *k = 0*).

2. Calculation of pulse response samples for all points of the reference region $D(n_1,n_2)$ with the use of ratios (3), (4) and the normalization of all samples, satisfying (5).

3. Processing of the distorted image and calculation of the quality criterion $Q_k(\hat{\omega}_1,\hat{c},\hat{\alpha})$.

4. If the $Q_k(\hat{\omega}_1,\hat{c},\hat{\alpha})>Q_{k-1}(\hat{\omega}_1,\hat{c},\hat{\alpha})$, estimates $\hat{\omega}_1, \hat{c}, \hat{\alpha}$ save, otherwise, according to some rule, a new variant of estimates is formed and the transition to step 2 is performed. If all estimates from the range of admissible values are "viewed" is an output.

Based on the described technology, it is possible to adjust the filter parameters both from the reference image and without the reference. When setting up a benchmark as a criterion $Q_k(\hat{\omega}_1,\hat{c},\hat{\alpha})$ proximity of the restored image to the standard is used indicator:

$$PSNR=10\log_{10}\left(\frac{MAX^2}{MSE}\right), \quad (6)$$

where *MAX* – is the maximum value received by the image pixel, and *MSE* – is the root-mean-square error (MSE) characterizing the proximity of the restored image to the reference image.

In the filter adjustment technology without a reference image, an MSE is calculated between the recovered and the original distorted images. In this case, in contrast to the setting according to the etalon, the improvement in the quality of the restored image is accompanied by a decrease in the PSNR (6). As the PSNR value is reduced, the quality of the resulting image can be either higher or lower, an additional condition is used: increasing the variance of the processed image. This requirement provides an increase in the average contrast, which is usually observed with an increase in the sharpness of the image. In addition, a restriction on the minimum allowable value of PSNR is introduced, since at its small values, significant distortions, characterized by high contrast, are possible.

Thus, checking the fulfillment of the conditions on the k-th iteration $Q_k(\hat{\omega}_1,\hat{c},\hat{\alpha})>Q_{k-1}(\hat{\omega}_1,\hat{c},\hat{\alpha})$ on step 4 of the described technology, it reduces to verifying that the following conditions are met:

$$PSNR\left(\hat{\omega}_k, \hat{c}_k, \hat{\alpha}_k\right) < PSNR\left(\hat{\omega}_{k-1}, \hat{c}_{k-1}, \hat{\alpha}_{k-1}\right)$$
$$D\left(\hat{\omega}_k, \hat{c}_k, \hat{\alpha}_k\right) > D\left(\hat{\omega}_k, \hat{c}_k, \hat{\alpha}_k\right), \quad (7)$$
$$PSNR\left(\hat{\omega}_k, \hat{c}_k, \hat{\alpha}_k\right) > PSNR_{доп}.$$

Here $PSNR(\hat{\omega}_k, \hat{c}_k, \hat{\alpha}_k)$ value indicator (6) calculated on the k - th iteration of the image reconstructed with filter parameters $\hat{\omega}_k, \hat{c}_k, \hat{\alpha}_k$ when compared with the reference (when tuning to the etalon) or the original distorted (when configured without reference); $D(\hat{\omega}_k, \hat{c}_k, \hat{\alpha}_k)$ – the dispersion of the reconstructed image, and $PSNR_{\text{доп}}$ – the minimum permissible value of the exponent (6).

3. Development of mobile application

The development of any mobile image processing application is divided into 2 parts: the development of a filter and the development of the main part. For the first part - the implementation of the image processing algorithm - the OpenCV library was used, a special version for the Android platform. OpenCV is a crossplatform open source library that is free for commercial and academic use. OpenCV contains algorithms for interpreting images, determining optical distortions, determining the similarity, analyzing the movement of an object, determining the shape of an object, and much more, written in C / C ++ and using parallelism and multi-core for the tasks performed.

To use the OpenCV library on the Android platform there is also a ready-made Android set NDK - tool that allows to realize part of the Android app compiled programming languages such as C and C ++ and contains a library to manage activities and access to various physical components of the device. Android NDK is integrated with tools from a set of components for software development (Android SDK), as well as with the integrated development environment of Android Studio [9-11].

Thus, there are 2 ways to use the OpenCV library on the Android platform:

1. OpenCV Java API + Android SDK: functionality is written in the Java language;
2. OpenCV native interface + Android NDK: functionality is written in C ++.

We will use the first way.

The main library modules used are:

- module Core - contains basic operations, including arithmetic;
- module Imgproc - responsible for image processing;
- module Utils - contains helper methods, e.g., conversion Bitmap image format into a format Mat OpenCV and back and Mat download from a resource identifier for the resource.

Since we are working with a limited amount of memory, it is advisable to load an image with a lower resolution into the memory. The version with the reduced resolution should correspond to the size of the component of the user interface that displays it. A high-resolution image takes up a large amount of memory and does not provide a visible advantage when displayed. The choice of the option can be made by the user taking into account the characteristics of a particular mobile device. Figure 2 shows the program's pseudo code, which implements the algorithm described in Section 2, written with regard to the features of programming on mobile devices.

To work with bitmaps, we use the type Bitmap, which allows you to maintain the responsiveness of the user interface and reduce the amount of memory [12,13]. Next, for image processing, the Bitmap type is converted to the Mat type. Mat is a class for storing images that can be interpreted in C ++.

Using the BitmapFactory class, you can learn the resolution and type of graphics data before creating a Bitmap object and before allocating memory to that object. When you receive information about the resolution and size of the image, you can decide whether to load into memory the full-sized version of the image or its reduced version. We list some factors that need to be taken into account.

Start
for w ∈ (w_In, w_End)/c ∈ (c_In, c_End)/α ∈ (α _In, α_End) with Step_w/Step_c/Step_ α do
start ***function1***
for i ∈ (1, N) with step = 1 do
for j ∈ (1, N) with step = 1 do

$$r = \left|\sqrt{\left(i - \frac{N+1}{2}\right)^2 + \left(j - \frac{N+1}{2}\right)^2}\right|$$

if $r > 0{,}5$ then

$$h(i,j) = \frac{\exp(-c\alpha\omega_1)}{\pi}\left(\frac{\sin\omega_1 r}{r} + \sin\omega_1 r + \frac{2\cos\omega_1 r}{r^2\omega_1} - \frac{2\sin\omega_1 r}{r^3{\omega_1}^2} + \left(\frac{\sin\alpha\omega_1 r - \sin\omega_1 r}{r}\right) + \left(\frac{c\cos\alpha\omega_1 r - r\sin\alpha\omega_1 r}{c^2 + r^2}\right)\right)$$

end
end
end
end
start ***function2***
for i ∈ (1, N) with step = 1 do
for j ∈ (1, N) with step = 1 do

$$XR(k,l) = XR(k,l) + X\left(k - \frac{N+1}{2} + i, l - \frac{N+1}{2} + j\right)h(i,j)$$

end
end
end
Input: min value of PSNR
New value of PSNR = psnr(XR, X, 1)
if calculated value of PSNR < previous value of PSNR and
dispersion of processed image(XR) > dispersion of distorted image (X) then
keep values of w/c/α
keep value of PSNR
end
Call ***function1*** и ***function2*** with new parameters w, c, α
Output: Processed image.
End

Figure 2. The pseudocode of the program.

1. Estimated memory consumption when downloading the full version.
2. The amount of memory that can be spent on an image, considering the total memory consumption of the application.
3. Resolution of the component that will display the downloaded data.
4. The size and density of the screen points on the current device.

Methods of the BitmapFactory class should not be executed in the main thread of the user interface, so as not to reduce system performance. It is impossible to predict how long it will take to download data and processing them, it depends on various factors (speed of reading from the disk, the size of pattern to be displayed, processing power, and e.g.). If such a task will block the main UI thread, the system will mark the application as hovering, and you can get out of it, and even if removed.

If you want to perform operations that take some time, they must necessarily be performed in separate threads, called "background" or "worker" threads. To implement this task, we use the class AsyncTask,

which offers a simple and convenient mechanism for moving laborious operations into the background thread.

The method of the AsyncTask class allows you to perform asynchronous operation in the user interface. About nor can perform minor image loading operations, file operations, database operations etc. in the workflow, and then publish the results in the main thread of the user interface without having to independently process the threads and/or handlers.

Figure 3 shows an application mockup built using an online tool for prototyping applications marvelapp.com.

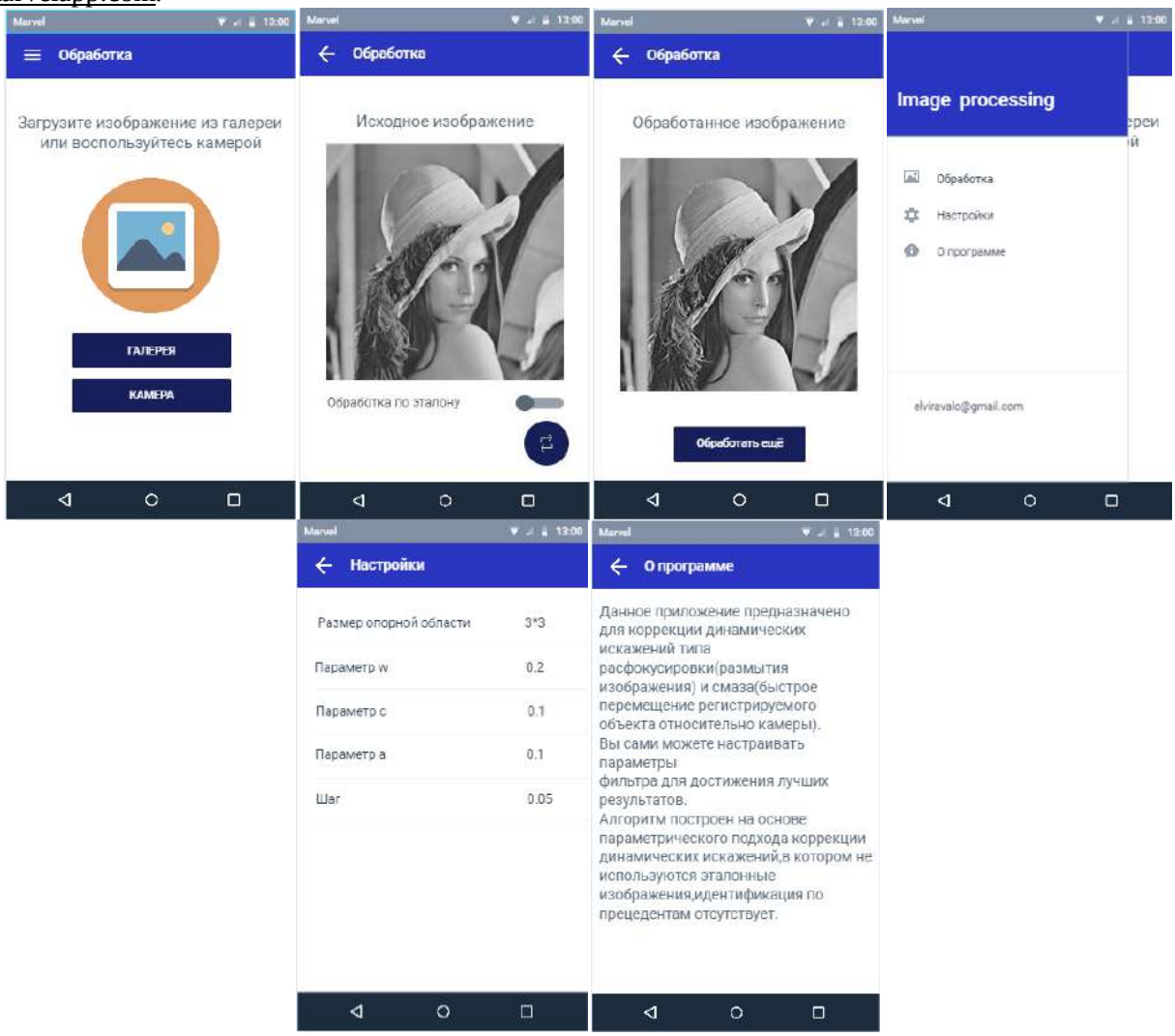


Figure 3. Applicaion layers.

When you enter the application, the user enters the main screen, where he can choose how to load the image: choose the finished one from the "gallery" or take a photo using the camera. And for newer versions of Android, starting at 6, you need to allow access to the application to the camera. In the next step, the user can see the image he uploaded in the user interface component, make a decision, process whether it is an image by the standard, and click on the button that will start the process. After processing, the user goes to the screen where he can see the result. The processed image is saved in a special folder of the application, accessible through the built-in "gallery" application.

To change filter parameters for better results during processing, the user can go to the "settings" section of the application main menu.

To get information about the application, the user needs to go to the "About the Program" section.

The diagram of the main application classes that implement the described functionality is shown in Figure 4.

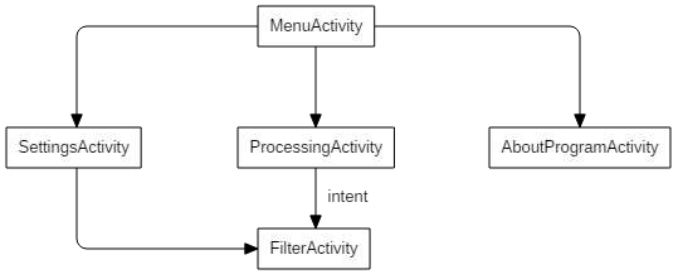


Figure 4. Class diagram.

4. Experimental results

Figure 4 shows the result of the filter with the configured parameters ω, c и a. Figure 6 (a) shows an example of a distorted halftone image by modeling a low-pass Gaussian filter with a blur: $\sigma = 3$. Figure 6 (b) shows the reconstructed image using a filter configured without a reference image, with a reference area dimension of 7×7. The result achieved during the restoration: PSNR = 25.13.

To quantify the achievable quality of recovery, we used test images (undistorted and distorted). We emphasize that the original undistorted image was not used to adjust the filter and restore it. We carried out a "blind" configuration.

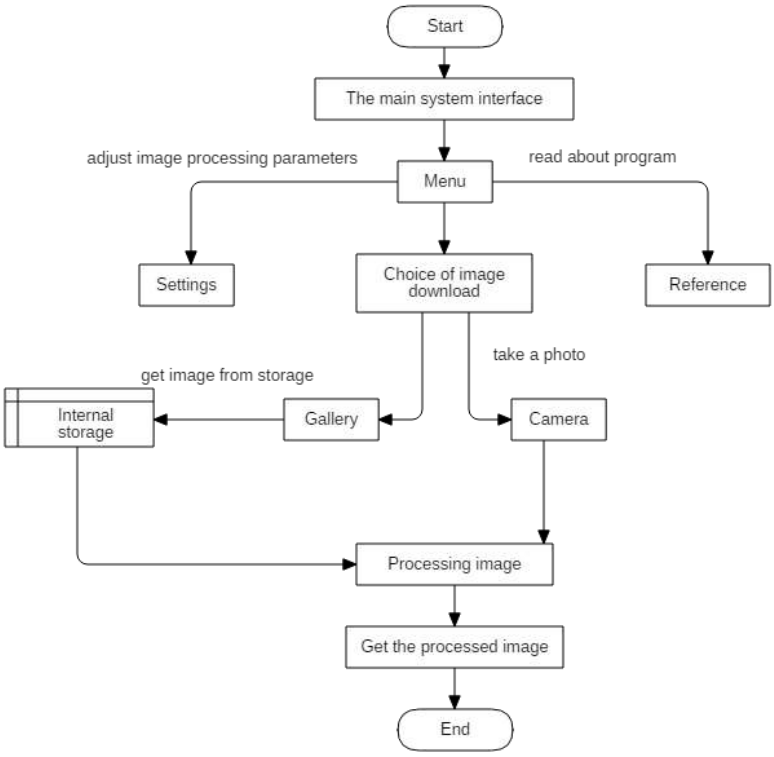


Figure 5. The scheme of the application.

Figure 5 shows a detailed scheme of the application.

(a)  (b) 

Figure 6. Images of "Lena": (a) - distorted with a blur σ = 3, (b) – processed.

This algorithm for image restoration without using a reference image can also be used to restore color images. In this case, the processing is performed using the same algorithms for each component.

5. Conclusion

The method of correction of dynamic distortions is based on the "blind" identification of the parameters of the restoring filter, which does not require the reference image. A program is developed on the Android platform, designed for image processing in mobile devices. An example of processing test image obtained by modeling the lowpass Gaussian filter is given. The result is achieved when reconstructing the image with a blur: σ = 3 by PSNR = 25, 13. The results obtained may be of interest to users of tablets and smartphones. Further plans of the authors are in order to make it accessible to the wide range of users.

6. References

[1] Mail.Ru Group, Mobile Internet in Russia (Access mode: https://corp.mail.ru/media/files/40314-researchmob i l e mail.pdf)
[2] Greisukh G I, Ezhov E G, Kazin S V and Stepanov S A 2017 Diffractive elements for imaging optics of mobile communication devices *Computer Optics* **41(4)** 581-584 DOI: 10.18287/2412-6179-2017-41-4-581-584
[3] Greisukh G I, Ezhov E G, Kazin S V and Stepanov S A 2017 Single-layer kinoforms for cameras and video cameras of mobile communication devices *Computer Optics* **41(2)** 218-226 DOI: 10.18287/0134-2452-2017-41-2- 218-226
[4] Nikonorov A V, Petrov M V, Bibikov S A, Kutikova V V, Morozov A A and Kazanskiy N L 2017 Image restoration in diffractive optical systems using deep learning and deconvolution *Computer Optics* **41(6)** 875-887 DOI: 10.18287/2412-6179-2017-41-6-875-887
[5] Soifer V A 2010 *Computer Image Processing, Part II: Methods and algorithms* (VDM Verlag) p 584
[6] Pratt U 1982 Digital image processing (Moscow: Mir) **2** p 480
[7] Fursov V A and Yakimov P Y 2017 Internet technology for correcting dynamic distortions on images in mobile devices *Proceedings of the XIX All-Russian Scientific Conference* (Moscow: IPM them. M V Keldysh) 436-445 DOI: 10.20948/abrau-2017-09
[8] Fursov V A 2018 Construction of quadratic-exponential FIR filters with an extended average frequency response region *Computer Optics* **42(2)** 297-305 DOI: 10.18287 / 2412-6179-2018-42-2-297-305
[9] *Class AsyncTask* (Access mode: http://developer.alexanderklimov.ru/android/theory/AsyncTask.php)
[10] *Android developers* (Access mode: https: // developer.android.com/index.html)

[11] Paramonov I V 2013 *Development of mobile applications for the Android platform: a tutorial* (Yaroslavl: YarSU) p 88
[12] Hardy B, Phillips B, Stuart K and Marsicano C 2016 *Android. Programming for professionals* (St. Petersburg: Peter) p 640
[13] Howse J 2013 *Android Application Programming with O penCV* (Packt Publishing Ltd.) 1-131

Acknowledgements
The work was supported by the Ministry of Education and Science of the Russian Federation and the Russian Foundation for Basic Research, projects No. 17-29-03112, No. 16-07-00729-a.

Fuzzy model for support investment decisions under risk

E V Orlova[1]

[1]Ufa State Aviation Technical University, K. Marks st. 12, Ufa, Russia, 450000

Abstract. One of the most important problem related to understanding the investor's behavior is to study the ways he selects, analyzes and interprets the available information and then uses it to make investment decisions. It is necessary to find out how an investor forms a certain opinion and comes to his own behavior strategy. The investor's behavioral model is variable since the financial market is volatile. The behavior of the investor is determined by a combination of rational (objective) and irrational (subjective) factors. To describe the influence of a combination of factors on the investor's behavior, a model describing this influence is needed. A formal description of interaction is complicated since a number of factors are of a qualitative nature, and the factors are also interrelated. The paper identifies the main causes and factors of irrationality in investor behavior which is the basis for its analyze and control. The fuzzy model, which allows to link a lot of behavioral factors with the utility (efficiency) of the solution is developed. Simulation results can be used for the investor's utility functions designing that is required to decisions making justification.

1. Introduction

It was shown [1-7] that under conditions of uncertainty and risk investors are being influenced a lot of irrational factors. The majority of financial theories [8-10] are based on the assumption of rational investor behavior. A number of researches show that many economic systems cannot be described with a sufficient degree of reliability by classical financial theories. Actually, the investor behavior is not strong rational and the majority of classical theories erroneously describe real systems [7, 9, 11].
Investigations of investor behavior in conditions of uncertainty and risk allow revealing a number of laws that have influencing to one of fundamental factors - "propensity to risk". The understanding of the factor "propensity to risk", studying the characteristics that influencing it, will enable to design mechanisms which will allow to consider the investor behavior.

For the decision of this problem it is necessary, firstly, to reveal principal causes of irrational behavior to analyze and supervise similar situations and as a result to avoid negative consequences of the "incorrect behavior". Secondly, it is necessary to predict in time and to estimate objectively other agents actions, using any deviations from a "rational" behavior.

In the paper the problem of irrational factors detecting has been decided. A set of investor's subjective factors has been revealed and their impact to investor behavior with using developed artificial intelligent model has been revealed.

2. Rational decision-making theory: advantages and disadvantages

The scientific basis for making decisions in different economic systems is economic theory. Economic science abounds in a variety of economic laws and patterns – demand (consumer choice theory) and supply (firm theory), diminishing returns (economies of scale), diminishing marginal utility. Knowing these universal relationships between events does not always make it possible to identify the causal

mechanism connecting the economic agent's actions and the functioning of economic institutions with the predicted results by these theoretical laws [14, 17].

One of the provisions of the neoclassical theory is the rational behavior of economic agents. Its main content reduced to the fact that the agent has all the information he needs, has enough time to analyze it, and the decision he makes as a result maximizes some of his utility function, which is considered known. At the same time, utility is understood as a measure of the psychological and consumer value of different sets of goods. If one set of benefits is more preferable, then its utility will be greater. The utility function reflects the comparative consumer value of sets of benefits based on a preference relation. In this form, the rationality of behavior also enters into the hypothesis of an efficient market, on which many analytical techniques are based for practical application.

Studies of the economic agent's behavior have revealed a number of features that deny rationality in the sense of the above description. For example, the law of demand can be interpreted as a causal law explaining the economic behavior by the causes, preferences and beliefs of rational agents, as a result of which they (agents) form the causal mechanism according to which the demand for each commodity changes back to a change in the price of the goods. At the same time, the universality of the law of demand has exceptions in the form of Griffin's paradox, in which a negative effect of consumption on price change is observed. To implement practical goals, it is necessary to use individual demand curves that take into account the preferences and expectations of future prices, the consumer's money income and the prices of other goods. That is, the neoclassical theory of consumer choice explains the pattern of the negative slope of demand curves, while Griffin's goods, which have a positive slope in the demand curve, are rare.

The neoclassical firm theory uses the notion of a positive slope of supply curves using the price or issue factors as strategically significant and has the basic premise of maximizing profit under specified technology and demand constraints or maximizing the utility function including profit, free time, liquidity, prestige , or about the maximization of the total sales volume with the minimum acceptable production efficiency. However, in this description the firm represents entrepreneurs as economic agents seeking to maximize some index of preferences, including monetary and non-monetary benefits. This approach is similar to the approach to consumer behavior in the theory of demand, the utility of entrepreneurs is reduced to the observed monetary profit, while the risk factors and uncertainties, time, costs of obtaining information in the formation of the utility function are not taken into account.

The normative nature of the rational decision-making theory is to explain how to behave in order to be rational. The utility theory forms a number of formal conditions that the individual's preferences and choices should satisfy. To determine the essence of rational preferences and choice, it is necessary to give the subject a set of prescriptions in the form of these conditions (the prerequisites of utility theory) about how to prefer and make a choice rationally. This is the essence of the difference between the normative theory of decision-making and the positive theory, which, in turn, describes, predicts and explains the behavior of individual economic agents.

To take into account the uncertainty in decision-making, Neumann and Morgenstern introduced the notion of expected utility [12], taking into account the probabilities of each possible outcome in the evaluation of utility. However, common in all these areas is the development of the idea of rationality and a stable orderly structure of preferences in conditions of complete, accessible and qualitative information about the probabilities of the expected outcomes.

Thus, the initial prerequisite for studying the behavior of an economic agent in the process of consumption of economic goods in the normative theory of decision-making is the rationality of the individual, it is assumed that the consumer is a kind of personality deprived of personal qualities that makes rational economic decisions: weighs the quality of the product and its price, searches for options for the best ratio of these two parameters, and then makes a decision. But the rationality of an economic agent is limited, depending on its limitations on the perception and processing of information, attitudes and prejudices and depends heavily on the degree of involvement in the consumer choice process. This approach requires improvements and additional assumptions, for which it is necessary to attract excellent approaches and models of behavior of economic agents, on the basis

of which it will be possible to identify the factors dominating in the behavior and to generate models with acceptable predictive ability.

Summarizing, we can say that the sources of uncertainty in the decision-making process are:

- Objective characteristics – lack of information, incompleteness of information resources, probabilistic nature of simulated processes, multivariate goals, criteria and alternatives;
- Subjective characteristics – low efficiency or incorrectness of the applied methods and technologies of data analysis, comparison and choice of alternatives due to subjective evaluation and interpretation of information, the disinclination of the person making the decision, responsibly perform all stages of decision-making.

As a result of the analysis carried out on the study of psychological factors, the violation of the rationality of the choice of decisions and their impact on the effectiveness of decisions, these factors are grouped in accordance with the structure of the decision-making process, as well as the preconditions for the manifestation of these factors. It should be noted that there are a lot of ways to classify these factors - by the degree of perception of information for decision-making, by the specifics of the subject area, etc. However, it seems more constructive for modeling and for more adequate assessment of the influence of subjective factors on the quality, timeliness and effectiveness of decisions taken Structuring the behavioral factors of the decision-maker in the enlarged stages of the decision-making process: 1 – identification of the target; 2 – collection of information and the formation of alternatives and criteria; 3 – formation of the attitude of preferences of alternatives; 4 – choice of alternative.

3. Factors of investor irrational behavior: the identification

One of the major problem associated with understanding of investor behavior is to study the way in which he selects, analyzes and interprets the available information and then uses it to form some principles and beliefs. In other words, one must determine how the investor understands and comes to a definite decision.

As a result of the analysis of psychological factors [13-16], the violation from the rationality, these factors are grouped in accordance with the structure of the decision-making process, as well as the preconditions for the manifestation of these factors. It should be noted that there are a lot of ways to classify these factors - by the degree of perception of information for decision-making, by the specifics of the subject area, etc. However, it seems more constructive to model and more adequately assess the influence of subjective factors on the quality, timeliness and efficiency of decisions taken Structuring of the behavioral factors of the decision-maker in the enlarged stages of the decision-making process: 1 – purpose identification; 2 – information collection and alternatives and criteria forming; 3 – alternatives preference relation forming; 4 - alternative choosing.

The first group includes factors of deviation from rational choice associated with the staging process - a problem situation description and the purpose forming. The definition of a goal requires significant, not always motivated costs and resources from the decision-maker, the impossibility, the difficulty of understanding and expressing an adequately informative goal that is the basis for decision-making, inconsistency:

- The factor of information representation consists in the difference in problem situation perception in case of its description in different formulations ("negative" or "positive"), and, consequently, in different preferences of alternatives corresponding to these two statements, which contradicts the rationality of the choice in making decisions;
- The factor of re-investment. When implementing investment projects, there is a periodic situation when the investor, investing a certain amount of financial, temporary resources, decides to continue financing to maintain its primary investments, even if the prospects for the project have deteriorated significantly. The probability of irrational investment of the project is directly proportional to the amount of invested money;
- Biased assessment of assets. The tendency of the subjects to attach greater importance to their own assets and to assign to them a higher cost, in comparison with the valuation of another's property. This effect is quite clearly manifested in the example of selling your own business, the

value of which is estimated based on the effort and money spent on creating a business, without taking into account the economic indicators and the value of similar assets.

The second group includes factors and personality traits caused by difficulties in the formation of objective constraints in the choice of the analysis of possible options and the allocation among them is feasible. Personality traits that contribute to these deviations from rationality are dreaminess, impossibility or unwillingness to distinguish desired and real.

- Information flow factor. The decision-maker is often influenced by a large number of heterogeneous information, often unreliable and the opinions of other people, often incompetent in the problem at hand;
- The factor of conservatism. Delayed change by the decision-maker of established beliefs and principles under the influence of a new information flow;
- Use of incomplete and inaccurate information. In certain situations, limited information is perceived as exhaustive and sufficient. This leads to its misinterpretation and, as a result, to irrational decision-making;
- The tendency to subjectively perceive the situation. A biased evaluation of existing information forms a subjective, often erroneous, opinion and, as a result, erroneous decisions;
- incorrect use of instrumental methods of information evaluation. In the conditions of existence of necessary and sufficient information, individuals can use incorrect methods to assess it;
- The factor determinism. It manifests itself in a tendency to see patterns in situations in which there is in fact an accident. The desire to predict certain events inclines the decision-maker to describe situations with deterministic characteristics. A similar situation occurs when several similar events create a belief in the occurrence of a phenomenon;
- Propensity to simplification. If the complexity and uncertainty of the situation increases, the subject loses rationality and begins to use simplifications. In connection with the difficulties of processing a large amount of complex information, some of this information can be lost, which leads to a simplification of the task. However, unrecorded information often has a high degree of significance. This explains why fairly simple approaches work well on the market, and complex decision-making systems, even if applied, are not always adequate;
- The factor of simple access. The economic agent attaches too much importance to information to which there is simple access, so the frequent repeated use of such information can be perceived as an irrefutable truth;
- Subjective assessment of probability. There is a difference between the actual probability of an event and the way an individual evaluates this probability;
- The factor of "slow" changes. Greater weight is given to general, not absolute changes. Individuals may not take into account the non-standard behavior of the system, if it occurs gradually, at certain intervals in time;
- The factor of the greatest significance of recent events. The most recent events tend to be given much more weight. The entity may feel that the business is no longer working after a sequence of unprofitable trades, although in fact it continues to function within the computed profit-and-loss relationships;
- Low propensity to change target goals. The essence of this effect is that in the subconscious of the individual there can be a conflict between his beliefs (assumptions) and the real reality. To avoid this, the subconscious tries to resolve this contradiction by "adjusting" historical facts to the existing beliefs. In other words, the human subconscious often "writes off" emerging problems on the "minor" shortcomings of the applied method, instead of pointing out the need for its modernization;
- Difficulties in choosing criteria and alternatives. Often such difficulties arise in individuals who are not prone to self-restraint or who are painfully related to external restrictions. In this case, virtual alternatives are formed, which can lead to undesirable consequences.

The third group integrates factors of personality traits leading to irrationality, and is associated with problems of evaluating the preferences of alternatives on a variety of criteria. On the subjective formation of transitive relations, the following factors influence the preferences of alternatives:

- Accounting for differences, not similarities. In order to simplify the choice between different alternatives, individuals do not notice (ignore) the common features of phenomena, focusing on their differences. This can lead to different preferences for the same alternatives in the same situations;
- Nonlinearity of preferences. This effect is manifested in situations of comparison of possible amounts of profits or losses, at which the significance of the difference in absolute value between them is smoothed as the values increase;
- Giving more importance to growth than to absolute change. The individual perceives not so much the absolute value of his wealth as his change, and the losses always seem more significant than the equivalent income. Individuals are more likely to take a greater risk to avoid losses than to obtain additional profits.

The fourth group of irrational behavior factors is associated with the decision-making stage and is connected with such personality traits as caution, indecisiveness, radicalism, impulsiveness:

- Rejection of losses. The negative emotions experienced in connection with losses are much stronger than the positive emotions associated with making a profit. The investor gives twice the value of losses than profits. In reality, an investment project with the same probability of profit and loss will not be of interest to the investor, even if the profit is one and a half times more than the loss;
- The factor of risk - competence. Most decision-makers tend to take more risks in areas in which they are more competent, no matter how this competence and professionalism can influence the likelihood of a decision;
- Subjective control factor. The propensity of the subject to a greater risk in situations in which, in the opinion of this subject, there are real opportunities to influence the results of events. A similar situation develops if the subject needs to take a set of decisions that do not directly affect the future result;
- Propensity to take risk depending on previous financial results. The degree of risk aversion largely depends on previous results (for example, investment decisions). If they were positive, then the risk aversion may temporarily decrease, and vice versa, after a succession of failures it only aggravates, leads to the appearance of "fear of mistakes";
- Increased risk for net profit. The individual is inclined to take a much greater risk when investing the funds of previous financial transactions. The propensity to risk the profit received from investing is increasing;
- The factor of choice of alternatives is manifested in the fact that when all alternatives are presented simultaneously, the subject chooses certain of them on some principle, and if the alternative is sequentially provided, he concentrates his choice on one of them.

The diagram of the cause-effect relationship of the factors of the investor's irrational behavior reflects the most significant factors and is presented in fig. 1.

Thus, there is a significant set of personal, psychological characteristics of economic agents that cause significant deviations from rationality in the choice of solutions. The individual, in whose behavior the listed features are clearly expressed, does not take into account the whole set of available information about the problem, in fact making the boundaries of the objective function, decision criteria, limited many alternatives, or unreasonably exposing them as the best part of them.

The revealed effects and patterns of behavior of economic agents in conditions of uncertainty and risk allow to explain many facts of economic agents irrational behavior. Especially brightly, irrational behavior manifests itself under uncertainty and risk in entrepreneurial, investment, innovation, and financial activities. Risk propensity is one of the most significant factors in the implementation of investment activities of innovative projects. The studying of factors influencing propensity to risk, determining the degree of their significance, provides certain possibilities for using this knowledge to influence the behavior of the economic agent (decision-maker).

In order to model the agents' behavior and take into account the consequences of irrational decisions arising from the occurrence of these deviations (errors), it is necessary that the modeling scheme include the following independent models: the model for controlled object, the model for

decision-maker, the model for generate alternatives. In microeconomic modeling, the enterprise acts as an object if a decision is made to implement an investment project for the modernization of fixed assets, the release of new products, optimization of the structure of output, etc. The model of the object reflects the influence of the decisions made on the controlled object. The model of the subject (decision maker) reflects the existing psychological, cognitive features of the subject and the structure of the factors of the decisions made. The model of alternatives presents a description of attributes and alternatives and reflects the evaluation of options from the perspective of goals and selection criteria and the structure of their set. An example of the application of decision-maker model is shown below.

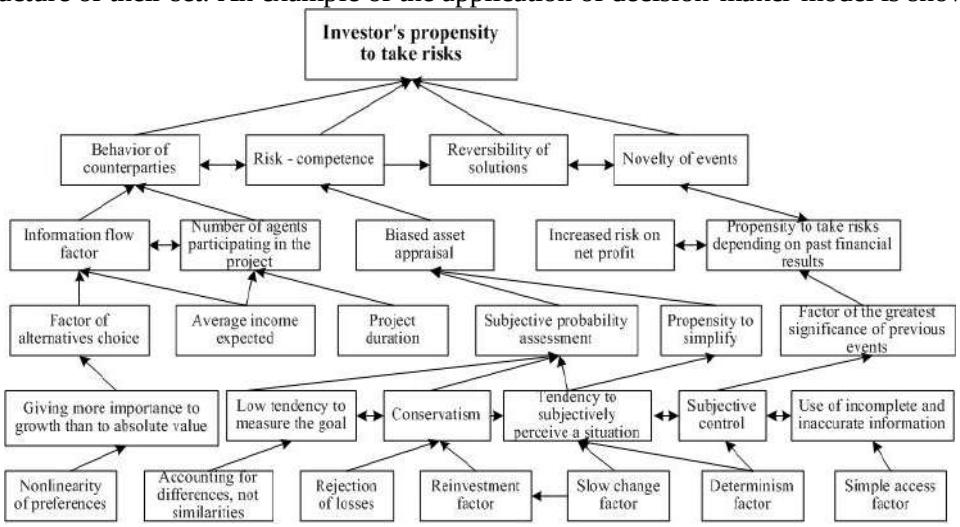


Figure 1. Cause-effect relationship of subjective factors of the investor's propensity to take risk.

4. The model for decision making

The propensity to take some decision differs for each decision maker. It coordinates by objective external and internal constraints which include:

- Objective environment factors;
- Internal research object factors;
- Demand, pressure of economic partners, competitors, consumers;
- Legislative regulatory restrictions that determine the liability (administrative, financial) for the management functions and regulatory between participants in the economic process;
- Social and moral obligations as determined goal that is implementation of the event will be successful.

The decision to choose is influenced by the relationship to decision maker to the described factors. One of the major unsolved problems of decision-making process modeling is the problem of subjectivity, which is not described by classical mathematical methods. To solve these problems a utility functions are used. In analyzing a decision in finance on the investor behavior the rational choice theory is usually used. It is based on the approach generated optimization model. Rational behavior is behavior which provides the best decision in terms of a particular purpose. As shown above a number of decisions based not only on rational considerations, but also in social traditions, subconscious reactions, moral installations scattered facts of personal experience in some field, and are the result of irrational behavior. Under high degree of uncertainty economic agents are not able to analyze the whole complex of factors and goals, and often apply special fragmentary discourse.

Classical analytical approach does not involve consideration of subjectivity in the decisions making, not investigate reasons and methods of mutual influence in the construction of economic evaluations.

To solve the problem of decisions selection about investment strategy we take into account the different risk propensity of potential investors and use tools based on the utility function of Neumann-Morgenstern [12]. To construct a utility function it is necessary to determine the nature of the behavioral study of alternatives influences on the functions type, as well as consider the impact of subjective factors to the risk propensity.

Practical application of utility theory in assessing of investment projects attractiveness identifies the following advantages of utility curve:

- as the expression of individual investor (expert) preferences utility curve once being built can make investment decisions in the future, taking into account their preferences, but without any consultation with expert;
- in general utility function can be used to delegate the decision making. It is logical to use a utility function of top management as to ensure its position in the decision it seeks to take into account the conflicting interests of all contractors. The utility function may change over time, reflecting the financial terms. Thus, utility theory can be formalized approach to risk and thus research to justify decisions taken under uncertainty and risk.

Because of the complex cause-effect relationship of subjective factors, irrational behavior, and risk propensity to determine the impact of factors it is appropriate to use artificial intelligence (AI) methods [16, 18-20]. With the help of fuzzy set theory the AI model was developed which allows to analyze and evaluate the impact of subjective factors for investors' risk propensity.

The implementation of fuzzy model is carried out in Matlab environment using the module Fuzzy Logic Toolbox [13, 14]. The simulation results are used in construction the utility functions of investment projects. The input data serves 25 parameters – subjective factors. Initial data processed by the system and generates an output variable is the risk propensity. All variables in the model are normalized in the interval from 0 to 1. The membership function is constructed for the resulting parameters "risk propensity" and "counterparty behavior". A rules base consist of 100 rules and has a form of "if-to" records and reflects different combinations of interconnection all input factors (irrational factors) and output (result) indicator "risk propensity". As shown above subjective factors of the investor's risk propensity have a complex relationship. Therefore, the input variables are in turn the resulting parameters for the factors of the lower hierarchical level.

Table 1. Simulation results.

Experiment No.	Risk – Competence	Reversibility of solutions	Novelty of events	Behavior of counterparties	Number of counterparties	Biased asset appraisal	Profit	Risk propensity
1	0.055	0.095	0.985	0.145	0.975	0.115	0.155	**0.147**
2	0.185	0.095	0.965	0.145	0.955	0.115	0.155	**0.202**
3	0.205	0.175	0.925	0.195	0.935	0.145	0.225	**0.341**
4	0.205	0.185	0.925	0.225	0.915	0.145	0.225	**0.359**
5	0.245	0.205	0.855	0.245	0.875	0.225	0.245	**0.403**
6	0.285	0.245	0.835	0.285	0.835	0.265	0.275	**0.434**
7	0.325	0.305	0.775	0.315	0.785	0.315	0.315	**0.464**
8	0.355	0.325	0.775	0.345	0.785	0.335	0.335	**0.476**
9	0.385	0.355	0.765	0.385	0.765	0.375	0.365	**0.494**
10	0.425	0.365	0.625	0.395	0.625	0.415	0.385	**0.532**
11	0.425	0.395	0.475	0.415	0.485	0.455	0.435	**0.585**
12	0.455	0.435	0.415	0.445	0.455	0.475	0.465	**0.605**
13	0.505	0.475	0.395	0.495	0.435	0.515	0.495	**0.645**
14	0.625	0.615	0.375	0.615	0.405	0.635	0.635	**0.655**
15	0.735	0.745	0.345	0.775	0.385	0.745	0.705	**0.685**
16	0.755	0.795	0.315	0.805	0.345	0.775	0.725	**0.717**
17	0.795	0.795	0.295	0.805	0.325	0.775	0.765	**0.723**
18	0.835	0.815	0.285	0.835	0.285	0.825	0.815	**0.792**
19	0.885	0.845	0.255	0.865	0.255	0.875	0.855	**0.881**
20	0.915	0.885	0.235	0.925	0.235	0.915	0.895	**0.961**
21	0.915	0.915	0.195	0.955	0.195	0.925	0.925	**0.965**
22	0.935	0.925	0.135	0.975	0.165	0.955	0.955	**0.985**
23	0.975	0.965	0.135	0.985	0.135	0.975	0.995	**0.995**

The simulation result is represented as the vector of values of variation of subjective factors (input variables) and assessing of parameter "risk propensity" (output variable), table 1.

As shown in fig. 2 the curve that determines the dependence of the variable “profit” and investors' “risk propensity”, has three different areas.

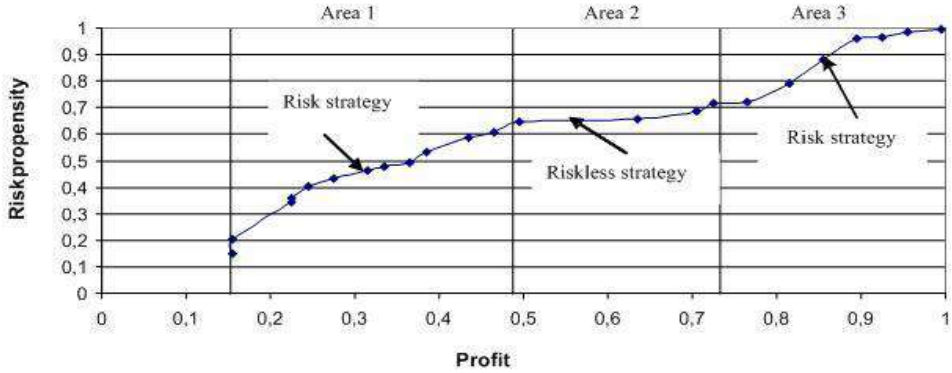


Figure 2. Utility function.

Area 1: Strategy of risky investment. Under the profit increasing the risk propensity has the tendency of significantly growth. Changing the low profits significantly affect the decision.

Area 2: Strategy of risk-free investment. Under the profit increasing the parameter “risk propensity” is practically unchanged. The mean change in income has weak effect on the decision.

Area 3: Strategy of risky investment. Under profit increasing the risk propensity has the tendency of significantly growth. The profit growth significantly affects the decisions.

Thus, based on AI model the impact of subjective factors on the risk propensity were determined. The results can be used in construction of utility functions of investment projects, taking into account the influence of subjective factors on the tendency of decision-makers to take risks.

Due to the application of AI model the problem of impact of subjective factors on the investor's risk propensity was solved. With the help of fuzzy modeling of the expert system, a fuzzy model was developed that makes it possible to analyze and evaluate the influence of subjective factors of the investor's propensity to take risks for the final decision adoption. The obtained results can be applied for constructing the utility functions of innovative projects that take into account the influence of subjective factors of the propensity of the decision maker to risk.

5. Results and conclusions

The transformation of methodological approaches in the decision-making theory in the study of the economic agents behavior is considered. The impossibility of an isolated using of rational choice theory as an adequate description of economic activity due to the objective existence of a multitude of causes of the irrationality of the decision-maker is shown. The main factors of the individual's irrational behavior, structured according to the stages of the decision-making process, based on the economic agent's personality-stable psychological features are selected and systematized. The AI approach for taking these effects in the decision making is proposed.

Given the complex causal relationship of irrational factors and to determine the degree of their influence on the investor's risk propensity, the fuzzy model is proposed. This model is used to construct the utility function, which is an objective basis for selecting behavioral alternatives. The practical use of the utility function can be found in the problems of economic decisions selecting. This will improve the quality and effectiveness of management decisions.

6. References

[1] Hausman D 2012 *Philosophy of the Economy* (Moscow: Publishing House “Anthology”) (in Russian)

[2] Hands U 2012 Regulatory Rational Choice Theory: Past, Present and *Future Problems of Economics* **10**

[3] Friedman M 2003 *Freedom* (Moscow) (in Russian)

[4] Van Horne and James C 2000 *Finance Management* (Moscow) (in Russian)
[5] Rusavin G I 2005 Theory of Rational Choice and its Limitations *Issues of Philosophy* **6** (in Russian)
[6] Rudic N B 2009 *Behavioral Finance in Economy* (Moscow) (in Russian)
[7] Blaug M 2004 *The methodology of economics or how economists explain* (Moscow) (in Russian)
[8] Kahneman D, Tversky A 2003 Rational choice, values and frames *Journal of Psychology* **4**
[9] Piskunov L P 2012 Transformation of Theoretical Approaches to the Study of Consumer Behavior *Science and Business: the Ways of Development* **9**
[10] Glukhov B 2010 Household Finances: Structure and Management *Problems of the Theory and Practice of Management* **1** (in Russian)
[11] Kleiner G B 2003 Modeling Methodology of Decision-making by Economic Agents *Economics and Mathematical Methods* **2** (in Russian)
[12] Neumann D, Morgenstern O 1970 *Theory of Games and Economic Behavior* (Moscow)
[13] Orlova E V 2012 Modeling of Utility Function With Using Irrational Factors *St. Petersburg State Polytechnical University Journal. Economics* **3** 24-30 (in Russian)
[14] Orlova E V 2014 Economic Behaviour: Synthesis of the Rational and the Irrational *Problems of Theory and Practice in Management* **3** 127-136 (in Russian)
[15] Orlova E V 2016 Modeling and Decision Support for the Firms' Pricing Policy under a Chaotic Dynamic of Market Prices *Proceedings of the Workshop on Computer Modeling in Decision Making* **1726** 81-88
[16] Orlova E V 2016 Simulation Model for the Firms' Financial Resource Management *Proceedings of the 28th International Business Information Management Association Conference on Vision 2020: Innovation Management, Development Sustainability, and Competitive Economic Growth* 1317-1321
[17] Orlova E V 2017 Control over Chaotic Price Dynamics in a Price Competition Model *Automation and Remote Control* **78(1)** 16-28
[18] Soldatova O P, Lyozin I A, Lyozina I V, Kupriyanov A V and Kirsch D V 2015 Application of fuzzy neural networks to determine the type of crystal lattices observed on nanoscale images *Computer Optics* **39(5)** 787-795 DOI: 10.18287 / 0134-2452-2015-39-5-787-794
[19] Orlov E P and Sizova I M 2016 Analytical representation of statistical characteristics of random processes with arbitrary spectra *Computer Optics* **40(4)** 560-571 DOI: 10.18287 / 2412-6179-2016-40-4-560-571
[20] Amosov O S, Ivanov Y S and Zhiganov S V 2017 Human localization in the frame of the video stream using an algorithm based on growing neural gas and fuzzy inference *Computer Optics* **41(1)** 46-58 DOI: 10.18287 / 2412-6179-2017-41-1-46-58

Minimisation of robust estimates of the sums of parametrised functions

Z M Shibzukhov[1,2], M A Kazakov[1] and D P Dimitrichenko[1]

[1]Institute of Applied Mathematics and Automation KBSC RAS, Shortanova str. 89A, Nalchik, Russia, 360000
[2]Moscow Pedagogical State University, Malaya Pirogovskaya str. 1, Moscow, Russia, 119991

Abstract. A robust approach to the design of machine learning algorithms, based on minimising finite sums of the parametrised functions is considered. This method implies using robust finite-sum differentiable aggregating functions that are resistant to outliers.

1. Introduction

The majority of machine learning problems can be reduced to problem of minimising finite sums of parametrised functions:

$$Q(\mathrm{w}) = \sum_{k=1}^{N} v_k \, \ell_k(\mathrm{w}),$$

where $\ell_1(\mathbf{w}), \ldots, \ell_N(\mathbf{w})$ – are non-negative basis functions, $\mathbf{w}$ – is the vector of unknown parameters, $\mathbf{w} \in \mathbf{W} \subseteq \mathbb{R}^m$, $v_1, \ldots, v_N \geq 0$ – is nonnegative weights. Most often $v_k = \mathrm{const}$, for example, 1 (arithmetic sum) or $1/N$ (arithmetic mean). Q target function is minimised by the optimal set of parameters $\mathbf{w}^*$:

$$Q(\mathbf{w}^*) = \min_{\mathbf{w} \in \mathbf{W}} Q(\mathbf{w}).$$

Most of algorithms for neural networks (NN) learning are based on this principle. In particular back propagation (BP) algorithm is based on minimisation of arithmetic mean squared errors.

However, if the distribution of the basic functions values contains outliers, the minimization of $Q(\mathbf{w})$, as a rule, leads to a distortion of $\mathbf{w}^*$. This is due to the fact that the arithmetic sum and the arithmetic mean are not resistant to the outliers.

Of course, the problem of outliers could be solved by choosing the values of the weights $v_1, \ldots, v_N$, which, on the one hand, would suppress the values of outliers, and on the other hand, leave the rest left unchanged. However, the selection of such weights is difficult task and is essentially equivalent by the complexity to identifying the outliers in the empirical distribution $\{\ell_1(\mathbf{w}^*), \ldots, \ell_N(\mathbf{w}^*)\}$.

One of the effective way of dealing with this problem is to use robust aggregation functions to calculate the sum or average. Thus, we get definitions for the function Q:

$$Q(\mathbf{w}) = \mathrm{med}_{k=1,\ldots N} \ell_k(\mathbf{w})$$

for robust estimation of a mean

$$Q(\mathbf{w}) = \sum_{k=1}^{N-p} \ell_{(k)}(\mathbf{w})$$

and for robust estimation of a sum. Here $z_{(1)}, \dots, z_{(N)}$ is the sequence of numbers obtained by arranging the initial sequence $z_1, \dots, z_N$ in ascending order. For example, to build a robust regression with $\ell_k(\mathbf{w}) = (f(\mathbf{x}_k, \mathbf{w}) - y_k)^2$ there have been proposed LMedS and LTS (Least Trimmed Squares) [1,2].

Last sum could be rewritten as trimmed arithmetical mean:

$$Q(\mathbf{w}) = \frac{1}{N-p} \sum_{k=1}^{N-p} \ell_{(k)}(\mathbf{w})$$

Minimizing the above estimates on data with outliers (up to 50%) allows finding adequate estimates for $\mathbf{w}^*$. However, minimization algorithms for LTS and LMS include a combinatorial component form of $\mathbf{w}^*$ search through subsets, since their gradients are singular. It makes application of gradient based algorithms almost impossible. This also reduces the scalability of such algorithms and their application in training neural networks and in problems with big data.

Another way of robust estimation of $\mathbf{w}^*$ is using a winsorized sum

$$Q(\mathbf{w}) = \sum_{k=1}^{N} \max\{\ell_k(\mathbf{w}), \overline{\ell}(\mathbf{w})\}$$

or an average

$$Q(\mathbf{w}) = \frac{1}{N} \sum_{k=1}^{N} \max\{\ell_k(\mathbf{w}), \overline{\ell}(\mathbf{w})\},$$

where $\overline{\ell}(\mathbf{w})$ is threshold value for the empirical distribution $\{\ell_1(\mathbf{w}), \dots, \ell_N(\mathbf{w})\}$.

In this paper, we consider general approach where for estimation of average empirical losses it will be used M-averaging aggregation functions (M-averages). This approach generalises M-regression method [5] and provides universal technique for solving the problem of the empirical risk minimisation in presence of outliers. It allows to use differentiable M-averages that could be treated as a sort of approximations of median and quantiles. In such cases a general gradient based procedure could be constructed for NN robust training.

2. Minimisation of M-averages from parametrised functions

For the median case, the problem can be solved using M-averages [5,7-10], which are differentiable and, in a sense, are approximate median:

$$\mathsf{M}_\rho\{z_1, \dots, z_N\} = \operatorname*{argmin}_u \sum_{k=1}^{N} \rho(z_k - u),$$

where ρ – is the nonnegative strictly convex function, $\rho(0) = 0$.

Here are some examples of M-averages:

- Collection of symmetrical averages:

$$\mathsf{M}^\gamma\{z_1, \dots, z_N\} = \operatorname*{argmin}_u \sum_{k=1}^{N} |z_k - u|^{1+\gamma},$$

- where $0 \le \gamma \le 1$ (M^0 is median, M^1 is arithmetical mean).
- Collection of non-symmetrical averages:

$$\mathsf{M}^\gamma_\alpha\{z_1, \dots, z_N\} = \operatorname*{argmin}_u \sum_{k=1}^{N} |z_k - u|^{1+\gamma}_\alpha,$$

- where $|u|^{1+\gamma}_\alpha = (\alpha - [u > 0])u|u|^\gamma$, $0 \le \gamma \le 1$ (M^0_α is α-quantile, M^1_α is α-expectile).

Here is the sufficient condition: if ρ – is twice differentiable, then $\mathsf{M}_\rho\{z_1, \dots, z_N\}$ has all partial derivatives:

$$\frac{\partial \mathsf{M}_\rho}{\partial z_k} = \frac{\rho''(z_k - \overline{z})}{\rho''(z_1 - \overline{z}) + \dots + \rho''(z_N - \overline{z})}.$$

Besides, $\frac{\partial \mathsf{M}_\rho}{\partial z_k} \geq 0$ and

$$\sum_{k=1}^{N} \frac{\partial \mathsf{M}_\rho}{\partial z_k} = 1.$$

In order to find out in which cases the function M-average M_ρ can be stable with respect to outliers we consider the following inequality:

$$|\mathsf{M}_\rho\{z_1,\ldots,z_N+\Delta\} - \mathsf{M}_\rho\{z_1,\ldots,z_N\}| = \frac{\rho''(\tilde{z}-u_{\tilde{z}})\Delta}{\sum_{k=1}^{N-1}\rho''(z_k-u_{\tilde{z}})+\rho''(\tilde{z}-u_{\tilde{z}})} < \rho''(\tilde{z}-u_{\tilde{z}})\Delta,$$

where $\rho(r)$ is convex, $\rho''(r)$ is continouous function, $\Delta > 0$ is value of distortion, $\tilde{z} \in [z_N, z_N + \Delta], u_{\tilde{z}} = \mathsf{M}_\rho\{z_1,\ldots,z_{N-1},\tilde{z}\}$. Let M_ρ be some M-averaging function. We define empirical risk based on M-averaging function M_ρ, as follows:

$$\mathcal{Q}_\rho(\mathbf{w}) = \mathsf{M}_\rho\{\ell_1(\mathbf{w}),\ldots,\ell_N(\mathbf{w})\}.$$

The classical empirical risk is a special case when M_ρ is arithmetical mean. The best set of the parameters for $\mathbf{w}^*$ have to minimize the function with respect to the minimization principle:

$$\mathcal{Q}_\rho(\mathbf{w}^*) = \min_{\mathbf{w}} \mathsf{M}_\rho\{\ell_1(\mathbf{w}),\ldots,\ell_N(\mathbf{w})\}.$$

Since the median and quantile are not continuously differentiable, the gradient procedures for minimisation of the risk functional are not practical. However, instead of median we can use continuously differentiable parametric family of M-average functions based on the dissimilarity function $\rho_\varepsilon(z-u)$ that satisfy the following requirements:

1. $\lim_{\varepsilon\to 0}\rho_\varepsilon(z-u) = |z-u|$;
2. $\lim_{\varepsilon\to 0}\rho'_\varepsilon(z-u) = \operatorname{sign}(z-u)$;
3. $\lim_{\varepsilon\to 0}\rho''_\varepsilon(z-u) = \delta(z-u)$ (Dirac's δ-function).

We demonstrate, for example, that for the role of "approximate" median the following functions can be used:

- $\rho_\varepsilon(r) = \sqrt{\varepsilon^2 + r^2} - \varepsilon$;
- $\rho_\varepsilon(r) = |r| - \varepsilon\ln(\varepsilon + |r|) - \varepsilon\ln\varepsilon$.

Such M-averages M_{ρ_α} are continuously differentiable and robust with sufficiently small ε. This implies that they are resistant to outliers (in some cases up to 50%).
To approximate the α-quantile, one can use the function

$$\rho_{\varepsilon,\alpha}(r) = \begin{cases} (1-\alpha)\rho_\varepsilon(r), & \text{if } r < 0 \\ \alpha\rho_\varepsilon(+0) + (1-\alpha)\rho_\varepsilon(-0), & \text{if } r = 0 \\ \alpha\rho_\varepsilon(r), & \text{if } r > 0. \end{cases} \quad (1)$$

The $\mathbf{w}^*$ search algorithm is an IR-ERM (Iteratively Re-weighted Empirical Risk Minimization) [3]:

procedure IR-ERM($\mathbf{w}_0$)
  $t \leftarrow 0$
  repeat
    $z_1 = \ell_1(\mathbf{w}_t), \ldots, z_N = \ell_N(\mathbf{w}_t)$
    $\overline{z}_t \leftarrow \mathsf{M}\{z_1,\ldots,z_N\}$
    for $k = 1,\ldots,N$ **do**
      $$v_k = \frac{\rho''(z_k - \overline{z}_t)}{\rho''(z_1-\overline{z}_t)+\cdots+\rho''(z_N-\overline{z}_t)}$$
    end
    $$\mathbf{w}_{t+1} \leftarrow \operatorname{argmin}_{\mathbf{w}} \sum_{k=1}^{N} v_k\,\ell_k(\mathbf{w})$$
  $t \leftarrow t+1$
  until $\{\overline{z}_t\}$ and $\{\mathbf{w}_t\}$ stabilize
end

At the heart of IR-ERM is the process of iterative re-weighting, as well as in the IRLS [4]. The IR-ERM algorithm differs from the IRLS in the way of recalculation of the weights.

To demonstrate the possibility of empirical risk minimisation [6] based on a robust estimate and IR-ERM algorithm, here is an example of linear regression problem with a large number of outlies.
We have a straight line through data points with an evenly distributed small error. For the linear regression recovering we use the least squares method, the absolute-error-minimising method, and the robust differentiable estimate minimising technique by means of the M-average through the function

$$\rho_\alpha(r) = |u - z| - \alpha\ln(\alpha + |u - z|) + \alpha\ln\alpha,$$

where $\alpha = 0.001$. Fig. 1 explain advantage of robust linear regression recovery. In both cases, robust differentiable average estimate minimising technique made it possible to avoid the influence of outliers.

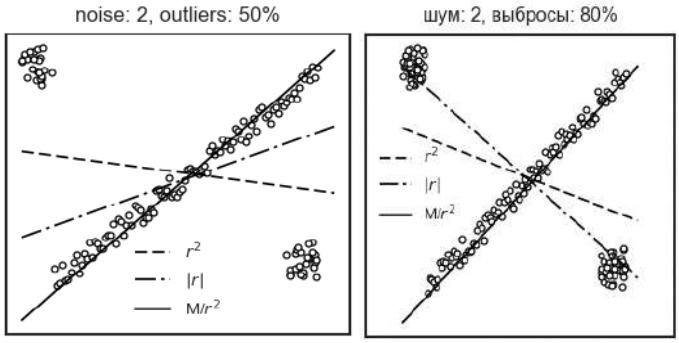


Figure 1. Recovery examples for linear regression with 50% and 80% of outliers from the amount of data without outliers.

3. Minimizing robust sums of functions

Consider a number of summation methods resistant to outliers. All M-averages including the arithmetic mean feature:

$$\frac{\partial \mathsf{M}}{\partial z_1} + \cdots + \frac{\partial \mathsf{M}}{\partial z_N} = 1.$$

But the arithmetic summation features the following important property:

$$\frac{\partial \mathsf{S}}{\partial z_1} + \cdots + \frac{\partial \mathsf{S}}{\partial z_N} = N.$$

It is therefore natural that the proposed summation methods can also maintain this property. Consider the following summation method.

3.1. Least Winsorized Sum and Mean

In the Least Winsorized Sum (LWS) method before summing, all values that are greater than the specified threshold value u are replaced by u, i.e.

$$\mathsf{WS}_u\{z_1, \ldots, z_N\} = \sum_{k=1}^{N} \frac{1}{2}(z_k + u - |z_k - u|).$$

Let's call it WS (Winsorized Sum). It has the following property: if u is the arithmetic mean of $z_1, \ldots, z_N$, then $\mathsf{WS}_u\{z_1, \ldots, z_N\} = z_1 + \cdots + z_N$.

The WM (Winsorized Mean) averaging method is defined as

$$\mathsf{WM}_u\{z_1, \ldots, z_N\} = \frac{1}{N}\mathsf{WS}_u\{z_1, \ldots, z_N\}.$$

We generalize the WS summing method as follows. Let M_ρ be M-average on the basis of a twice differentiable strictly convex function ρ. Denote $\overline{z} = \mathsf{M}_\rho\{z_1, \ldots, z_N\}$. Define

$$\mathsf{WS}_\rho\{z_1, \ldots, z_N\} = \sum_{k=1}^{N} \frac{1}{2}(z_k + \overline{z} - \rho(z_k - \overline{z})).$$

Calculate the partial derivatives:

$$\frac{\partial \mathrm{WS}_\rho}{\partial z_k} = \frac{1}{2}(1 - \rho'(z_k - \overline{z})) + \frac{1}{2}\frac{\partial \mathrm{M}_\rho}{\partial z_k}\left(N + \sum_{l=1}^{N} \rho'(z_l - \overline{z})\right).$$

Since, by definition,

$$\sum_{k=1}^{N} \rho'(z_k - \overline{z}) = 0,$$

then

$$\frac{\partial \mathrm{WS}_\rho}{\partial z_k} = \frac{1}{2}(1 - \rho'(z_k - \overline{z})) + \frac{N}{2}\frac{\partial \mathrm{M}_\rho}{\partial z_k}.$$

Therefore

$$\sum_{k=1}^{N} \frac{\partial \mathrm{WS}_\rho}{\partial z_k} = N.$$

If

$$\lim_{|r|\to\infty} \rho(r)/|r| = 1,$$

then the summation method defined here can be considered as a smooth version of WS.

Now we consider the following problem of the objective function minimizing winsorized mean:

$$Q(\mathbf{w}) = \frac{1}{N}\mathrm{WS}_\rho\{\ell_1(\mathbf{w}), \ldots, \ell_N(\mathbf{w})\}$$

to find the optimal set of parameters $\mathbf{w}^*$. Now write down the gradient:

$$\mathrm{grad}Q(\mathbf{w}) = \sum_{k=1}^{N} v_k(\mathbf{w})\mathrm{grad}\ell_k(\mathbf{w}),$$

where

$$v_k(\mathbf{w}) = \frac{1}{2N}\left(1 - \rho'(\ell_k(\mathbf{w}) - \overline{z}(\mathbf{w}))\right) + \frac{1}{2}\frac{\partial \mathrm{M}_\rho}{\partial z_k},$$

$\overline{z}(\mathbf{w}) = \mathrm{M}_\rho\{\ell_1(\mathbf{w}), \ldots, \ell_N(\mathbf{w})\}$. At that, we note

$$v_1(\mathbf{w}) + \cdots + v_N(\mathbf{w}) = 1.$$

For numerical calculation, we can apply the algorithm IR-SWSM (Iteratively Re-weighted Smoothly Winsorized Sum Minimization) the next version of the IR-SWSM algorithm:

procedure IR-SWSM($\mathbf{w}_0$)
  $t \leftarrow 0$
  repeat
    $z_1 = \ell_1(\mathbf{w}_t), \ldots, z_N = \ell_N(\mathbf{w}_t)$
    $\overline{z}_t \leftarrow \mathrm{M}\{z_1, \ldots, z_N\}$
    for $k = 1, \ldots, N$ **do**
      $v_k = \frac{1}{2N}(1 - \rho'(z_k - \overline{z})) + \frac{1}{2}\frac{\partial \mathrm{M}_\rho}{\partial z_k}$
    end
    $\mathbf{w}_{t+1} \leftarrow \mathrm{argmin}_{\mathbf{w}} \sum_{k=1}^{N} v_k \ell_k(\mathbf{w})$
    $t \leftarrow t + 1$
  until $\{\overline{z}_t\}$ and $\{\mathbf{w}_t\}$ stabilize
end

To illustrate the IR-SWSM algorithm capacity, we consider a neural network with single hidden layer:

$$y = w_0 + w_1 u_1 + \cdots + w_m u_m$$
$$u_j = \mathrm{softplus}(w_{j0} + w_{j1}x_1 + \cdots + w_{jn}x_n),$$

where $\mathrm{softplus}(s) = \ln(1 + e^s)$. NN was trained with the Boston data set.

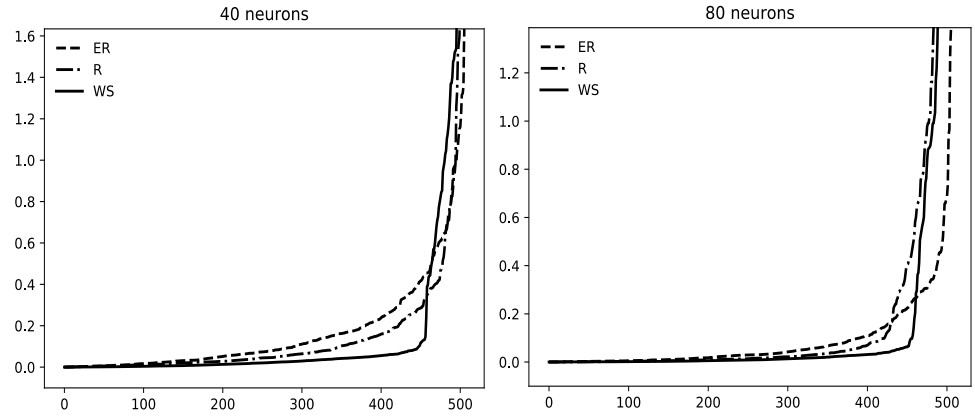


Figure 2. The errors distribution by trained NN with one hidden layer containing 40 and 80 neurons, with respect to the Boston dataset.

For training, error back propagation has been applied, where the mean square error (ER) and mean value of the Huber function with the small parameter (0.001) are minimised, for it to be a continuously differentiable approximation of the module error. IR-SWSM learning algorithm has also been used, where the robust estimate of the error sum WS of squares (WS) are minimised using the function $\rho_{\varepsilon,\alpha}$ type (1), where $\rho_\varepsilon(r) = \sqrt{\varepsilon^2 + r^2} - \varepsilon$, $\alpha = 0.90$.

The Fig. 2 shows the of mean absolute error distribution across the entire dataset. This clearly demonstrate that training neural networks (40 and 80 neurons in the hidden layer) with IR- SWSM algorithm reduce error values to more than 80% of the data.

There is also another experiment with only 7 neurons in the hidden layer. The Fig. 3 shows the of mean absolute error distribution across the entire dataset.

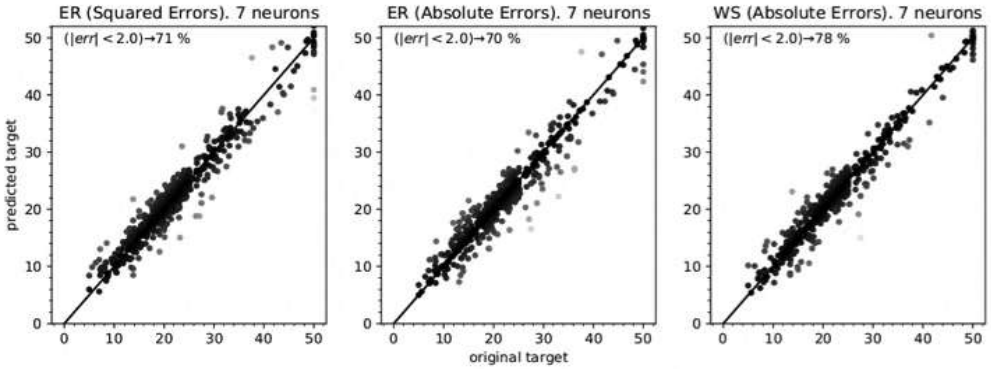


Figure 3. Errors distributions for NN training by boston dataset using three approaches: Least Squares (ER), Least Absolute Errors (ER) and Least Winsorized Squares (WS).

4. Conclusion

In this paper, we propose a method and algorithms for minimizing robust differentiable estimate of means and sums that are potentially resistant to outliers and errors that can lead to a shift in the parameters of the trainees. It is based on minimization of differentiable robust analogs of median, quantiles and winsorized sums of loss functions.

The above approaches are preferable in the cases when application of gradient based minimisation procedures are preferable. For example, these approaches made possible application of weighted variants of back propagation algorithms for NN robust learning. Construction of robust learning algorithms of NN are important in a sense of many applications [12-15]. In particular an iteratively re-weighted procedures are proposed.

In these procedures at each step a weighted variant of back propagation algorithm is used. Examples presented above clearly show that proposed approaches and algorithms can be resistant to a large amount of outliers.

5. References

[1] Rousseeuw P J 1984 Least median of squares regression *American Statistical Association* **79** 871-880
[2] Rousseeuw P J 1987 *Robust regression and outlier detection* (NY: John Wiley and Sons)
[3] Shibzukhov Z M 2017 On the principle of empirical risk minimization based on averaging aggregation functions *Doklady Mathematics* **96(2)** 494-497
[4] Andersen R 2008 *Modern Methods For Robust Regression* (Thousand Oaks: SAGE Publications)
[5] Huber P J 1981 *Robust Statistics* (NY: John Wiley and Sons)
[6] Vapnik V 2000 *The Nature of Statistical Learning Theory. Information Science and Statistics* (Springer-Verlag)
[7] Mesiar R, Komornikova M, Kolesarova A and Calvo T 2008 Aggregation functions: A revision *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models* (Springer, Berlin, Heidelberg)
[8] Grabich M, Marichal J-L and Pap E 2009 Aggregation Functions *Encyclopedia of Mathematics and its Applications* **127**
[9] Beliakov G, Sola H and Calvo T 2016 *A practical guide to averaging functions* (Springer)
[10] Calvo T and Beliakov G 2010 Aggregation functions based on penalties *Fuzzy Sets and Systems* **161(10)** 1420-1436
[11] Yohai V J 1987 High breakdown-point and high effciency robust estimates for regression *The Annals of Statistics* **15** 642-656
[12] Nikonorov A V, Petrov M V, Bibikov S A, Kutikova V V, Morozov A A and Kazanskiy N L 2017 Reconstruction of the images in difractional-optical systems on the base of convolutional neural neworks and deconvolution *Computer Optics* **41(6)** 875-887 DOI: 10.18287/2412-6179-2017-41-6-875-887
[13] Nikitin M Y, Konushin V S and Konushin A S 2017 Neural network model for human recognition by face in video sequences with estimation of frame's utility *Computer Optics* **41(5)** 732-742 DOI: 10.18287/2412-6179-2017-41-5-732-742
[14] Spicyn V G, Bolotova Y A, Fan Hgok Hoang and Bui Thi 2016 Chang Recognition of symbols on the base of wavelet transformations, principal components and neural networks *Computer Optics* **40(2)** 249-257 DOI: 10.18287/2412-6179-2016-40-2-249-257
[15] Poletaev S D and Volotovsky S G 2016 Precision laser recording of microstructures on molybdenum films for generating a diffractive microrelief *Computer Optics* **40(3)** 422-426 DOI: 10.18287/2412-6179-2016-40-3-422-426

Acknowledgments
The work was supported by a grant from the Russian Foundation for Basic Research 18–01–00050.

Algorithm for detecting the latent mastitis state of animals in a dairy farms on the based of data fusion from different types sensors

L V Antonov[1], A A Orlov[1] and A V Astafiev[1]

[1]Murom Institute (branch) Vladimir State University, Orlovskaya street 23, Murom, Russia, 602264

Abstract. The problem of latent mastitis identification in livestock enterprises is analyzed. The necessity of automatizing the mastitis identification process is shown. Biological methods for determining the presence of the disease are considered in the article. Common methods of data fusion for the extraction of an informative trait is analyzed in the work. A new algorithm for identifying mastitis in animals based on data fusion from the livestock enterprise sensors is proposed. The developed algorithm as compared to the conventional method of determining mastitis increases the accuracy of the disease's identification by 6.5 percent.

1. Introduction

The most common non-communicable disease animals contract in dairy plants is mastitis. The disease has two main types: subclinical (hidden) and clinical (open visual symptoms of the disease). The most dangerous of them is subclinical, when the udder and milk produced by the cow both look absolutely normal. Hidden mastitis occurs 5-10 times more frequently than clinical. If subclinical mastitis is not detected in a timely manner, then it goes to the clinical stage after some time. This leads to a situation where the animal must be removed from production for treatment with antibiotics. Animal milk becomes unfit for further processing after medicines are used [1]. The risk of mastitis has a high probability that the animals can get sick at any time of the year and at any stage of the production cycle. For example, it may be in the early days of the postpartum period or in the period of maximum lactation and even in the dry period. The disease is more dangerous in the dry period because the animal ceases to give milk and the condition of the animal's udder is not checked [2]. Mastitis is one of the most common causes of the animals' mass culling. The statistics from [3] shows that 5-35% of culled cows are animals which have mastitis that includes atrophy of parts of the udder. Statistics [1] show that the overall incidence of dairy mastitis in animals is 24-45% in all kinds of farms. The procedure for identifying the disease requires visual veterinary inspection and chemical analysis of milk from different parts of the animal's udder, if possible. The consequences of mastitis are very serious. The next lactation of animals, who have been successfully cured, becomes less productive at 7-32% compared to the average level of the herds' milk productivity [4]. Thus, we meet the topical problem of identifying subclinical mastitis cases in the early stages of the disease.

2. Problem formulation.

The usual method for the detection of subclinical mastitis is an expensive and complex clinical analysis. The essence of this method is a survey of sample milk from the udder of the sick animal.

This helps determine the number of somatic cells. Complex chemical analysis is conducted using expensive reagents [6].

The standard procedure for detection of mastitis includes several different methods used together:

- Clinical study of the udder and milk ejection.
- Taste test of the milk.
- Chemical analysis and measurement of the electrical conductivity of the milk.

Another problem is discussed in [3-4]. Detection of latent mastitis is a very time-consuming task for dairy herds where the number of animals is a few thousand. As stated, the precise identification requires veterinary examination and chemical analysis. This process takes from a few hours to a few days if there is a free laboratory and doctors who are available. Most animals are unexamined in this case. Therefore, there are methods of clipping and reducing the number of test animals in [5]. Early examination of animals suspected of mastitis is formed on the basis of their secondary features. After that, the expert decision about the treatment of each animal is accepted. On the other hand, there is a large number of different sensors installed on the animals. They are used to solve specialized problems [7,12]. Some useful information from the sensors may be lost or not used to solve a number of problems, including the task of identifying mastitis. But the information from some types of sensors can be useful for automatizing solutions to the process of determining latent mastitis on dairy enterprises. The analysis of biological methods shows that visual veterinary inspection and chemical analysis of milk from different parts of the animal's udder is required for accurate identification of the disease [8-10]. High time complexity, the use of expensive equipment and the creation of laboratory facilities within the company are required to perform this analysis. Terms of risky agriculture including tough weather conditions, short duration of daylight hours and the location in the temperate latitudes do not allow for high yields for high profitability and the creation of a company's own laboratory stations. Mastitis disease is the most common disease on farms of any type. The problem of identification of mastitis is one of the main problems. The accurate diagnosis of latent mastitis is time consuming and requires significant financial costs [11]. The creation and application of algorithms using data collected from a variety of sensors installed in the enterprise is offered as a way to reduce costs for the identification of animals suffering from mastitis.

3. Development of algorithm to identify the latent mastitis based on allocation of complex feature.

Data from 800 animals with real livestock enterprises are used as a source in the research. The connection between the traits of mastitis, including milk yield, conductivity and the emergence of the animal's mastitis, has been previously established. These parameters are the input. The presence of mastitis in animals is an output parameter of the mathematical model. It is represented in the database as 0 and 1, and it was determined by the expert (herd manager) at the company. Thus, we have a set of three parameters.

Let us introduce the following notation. We shall accept t as the average number of days in lactation (about 305 days), n as the number of the company's animals. Let us assume $\mu_i(t)$ is the value of milk production index in every day observations, $\eta_i(t)$ is the conductivity value of the index for each day of observation, $v_i(t) = \{0, 1\}$ - expert opinion on the existence or lack of mastitis in the animal. A general view of the mathematical regression model looks as follows:

$$v_i(t) = k_1 \mu_i(t) + k_2 \eta_i(t) + b, \quad (1)$$

where $i = 1..n$.

Thus, the relationship between input and output parameters must be found to assess the regression coefficients k_1 and k_2. Regression analysis allows us to determine the influence of individual independent features on the result (dependent feature). If we shall accept the input parameters $\mu(t)$ and $\eta(t)$ as matrix X with two columns, output parameter $v(t)$ as a vector Y, then the vector of the regression coefficients r calculated by the ordinary least squares is:

$$r = (X^T X)^{-1} X^T Y \quad (2)$$

But the use of the regression model and the least squares method requires that the data were the same size. Therefore, the difference $|\mu_i(t) - \eta_i(t)|$ (where $i = 0,1\ldots n$) should not be large. In our case, the index of milk production is measured in thousands of mL per day, but the electrical conductivity of

milk is measured in several mS/cm. According to the definition of the Euclidean norm it makes one parameter much more significant than the other, when in fact it is not. It is noteworthy that the inputs have different distribution functions. Features must be made independent of time for the mathematical model to produce correct results. Thus, pre-processing of the necessary data must be carried out. Normalization and centering of random functions must be performed.
Let us perform the centering of the function $\mu_i(t)$ according to (3), (4).

$$T(t) = \frac{1}{n}\sum_{i=1}^{n}\mu_i(t) \tag{3}$$

where $i = 1..n$.

$$R_i(t) = \mu_i(t) - T(t) \tag{4}$$

where $i = 1..n$.

The expected value of the function $R(t)$ is 0, and the data are uniformly distributed relative to the x-axis after the conversions (3), (4). Thus, the parameter's values were converted. The distribution of random function $R(t)$ does not depend on the time after the conversion.

Let us perform the normalization of a centered function $R(t)$. Calculation of the standard deviation (SD) for the data of the herd's milk productivity produced by the (5), (6).

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}R(t)^2} \tag{5}$$

$$\hat{\mu}_i(t) = \frac{R_i(t)}{\sigma} \tag{6}$$

where $i = 1\ldots n$.

Then we can calculate the value of the regression coefficients and the free terms of the equation by using (2). Values that were calculated are show in Table 1.

Table 1. The regression coefficients for equation (1).

Regression Coefficient	Value
k_1 (Coefficient of milk production)	-0.0493
k_2 (Conductivity coefficient)	0.1933
b (free term)	0.1022

Thus, the plane that approximates the initial space of the mastitis traits is shown as a set of points in the graph constructed by the input coordinates is found (Fig. 1). The regression model gives the required complex feature, allowing for the estimation of the probability of mastitis in the animals. The regression model to determine the presence of mastitis in the animals is shown in (7). Coefficients are calculated using the least squares method.

$$\nu_i(t) = -0.0493\cdot\hat{\mu}_i(t) + 0.1933\cdot\hat{\eta}_i(t) + 0.1022 \tag{7}$$

where $i = 1..n$.

The practical purpose of the research is the detection of mastitis using the threshold value of the complex feature. The recognition threshold (P) is calculated according to the range of confidence probability, which determines the livestock expert enterprise. Entering the threshold manually allows us to the identify of 100 animals with mastitis because the value of confidence interval is low. Nevertheless, among the sampled animals suspected to disease there are also many healthy cows. Therefore, manual inspection of all the suspected animals is not possible because of the large count of cows in the sample. There is type II error. Thus, the confidence interval may be changed depending on how many animals need to be checked by the workers per day, and what level of identification accuracy is required.

The first step is building a histogram of the probability density distribution for calculating complex mastitis traits. The next stage is obtaining the threshold value P in accordance with the installed confidence interval. Suspicion of disease in the animal is determined when the value of the complex

mastitis trait, based on milk yield and conductivity exceeds the threshold *P*. The animal is entered into a special sample. These methods are based on a single parameter—the electrical conductivity of milk. Thresholds of electrical conductivity to determine the animal's health conditions are given in [6] (Table 2).

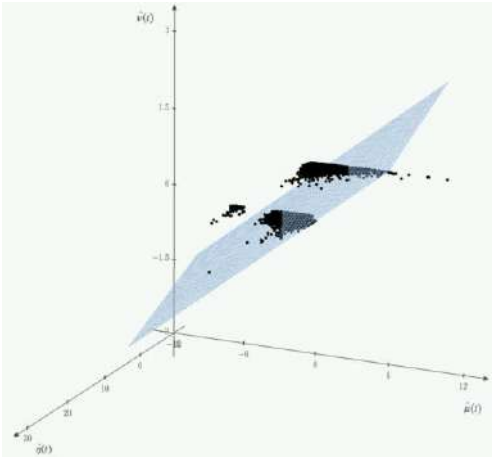


Figure 1. The plane which approximates the feature space.

Table 2. Milk conductivity in healthy and ill quarters (operating instructions).

The Health Status of the Udder	Milk conductivity (mS/cm)
Healthy udder	< 5.5
Suspicion of mastitis udder	5.5 – 6.6
Subclinical mastitis	> 6.5
Clinical mastitis (Milk must not be delivered)	> 8

4. Experimental results.

Currently there are several approaches for preliminary identification of mastitis in animals without using chemical analysis. Generally, mastitis in animals is identified when the milk conductivity exceeds the threshold value of 6 mS/cm. This approach shows good results when you need to create a list of animals suspected of having mastitis disease without expensive equipment [6].

The algorithm based on the complex feature is compared with the algorithm for identifying mastitis in animals using the threshold 6 mS/cm of milk conductivity.

Data obtained for more than 800 cows, for about 305 days from the two sensors' measuring parameters were treated in the research. The objective quantitative results of the algorithms' experimental research for data about daily milk yield and conductivity are shown in Tables 3 and 4. The results of the algorithm were compared with an expert's estimates. The expert estimates that the results about the animal disease are not 100% accurate, because it is carried out at an early stage of the disease. But the application of this approach without long diagnostic procedures increases the efficiency of decisions. Therefore, the expert assessments are the reference data. The percentage of sick animals by the algorithm among expert data about healthy cows is type II error. The percentage of healthy animals by the algorithm among expert data about sick cows is type I error. Type I error is more important than type II error because if a sick animal is not identified, then the disease will become clinical and the animal will be removed from production. Tables 3 and 4 show that the developed algorithm based on the complex feature gives better operation results in comparison with the algorithm of fixed threshold conductivity.

The algorithm developed reduces type I error to 6.5%. It is possible to increase the number of sick animals, which are found at exactly 10 units. Type II error has been decreased to 3%. Thus the algorithm helps achieve a smaller number of false positives for identification of bovine mastitis.

Table 3. Results comparison of algorithm based on the threshold of milk conductivity 6 ms/cm with expert opinion.

Percentage of Animals That are Found by the Algorithm	Expert Opinion	
	Sick Animals, %	Healthy Animals, %
Sick Animals	89.5	9
Healthy Animals	10.5	91

Table 4. Results comparison of developed algorithm based on the complex feature with expert opinion.

Percentage of Animals That are Found by the Algorithm	Expert Opinion	
	Sick Animals, %	Healthy Animals, %
Sick Animals	96	6
Healthy Animals	4	94

Errors of identification are decreased by the help of the additional trait, low level milk yield, for the calculation of complex mastitis traits. The scalable fragment of changes of milk yield values, milk conductivity and the calculated complex mastitis traits during lactation are shown in Fig. 2.

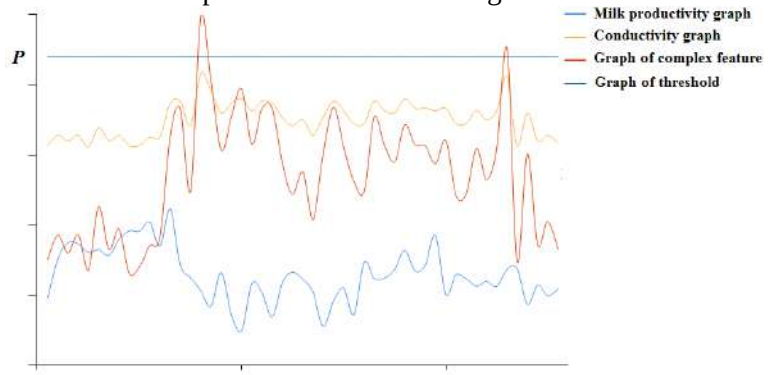


Figure 2. Comparison of the results of algorithms identification latent mastitis in animal.

The graphic fragment of lactation for one of the animals, including changing values measured by the sensors, is shown in Fig. 3. The dynamics of changes in the rate of electrical conductivity shows that there is mastitis in the animal. This is not the same as the expert evaluation. This is a mistake. The graph of the complex trait is stable and does not exceed the threshold *P*, because milk productivity is normal. Thus, type II error is eliminated.

The estimated algorithm for more than 800 cows was produced daily. The value of complex mastitis traits were calculated for each animal daily. Characteristic values exceeding the threshold P were considered abnormal. The confidence interval was 90%. The results of evaluation of the algorithm research are shown in Table 5.

Table 5. A Quantitative comparison of algorithms results of the mastitis identification.

Applied Algorithm	All days	Critical Deviations Found
Identification using a fixed threshold of electrical conductivity 6 mS/cm	131274	14239 (10.8%)
Identification Based on the Complex Feature	131274	13804 (10.5%)

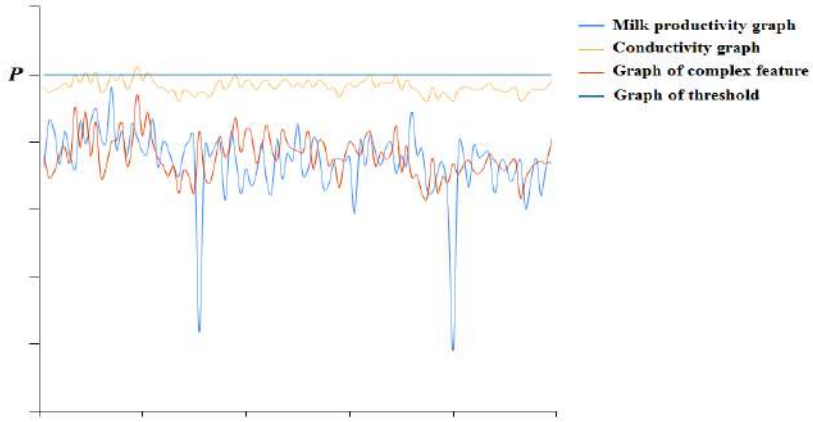


Figure 3. Comparison of the results of algorithms identification latent mastitis in animal.

5. Conclusion

The relevance of timely identification of subclinical mastitis is shown in the work. The process of accurately detecting bovine mastitis is expensive. Furthermore, accurate identification requires the manual inspection of large numbers of animals. There are not enough workers in livestock to solve this problem. Therefore, automation of the disease identification process is necessary. Methods of forming a preliminary list of animals suspected to have mastitis are considered in the article. These methods are based on data analysis from sensors. The approach is based on the dedicated complex traits which allow for identifying the disease proposed in the work. The experimental results show that the developed algorithm for the mastitis detection is more accurate than the traditional approach, which is described in [6]. Use of complex mastitis traits allows for an increase in the accuracy of identifying mastitis by 6.5% and a 3% decrease in false positives. Thus, the proposed algorithm shows promising results. It can be used in program systems for monitoring the production of livestock farms.

6. References

[1] Gianneechini R E, Concha C and Franklin A 2002 Antimicrobial susceptibility of udder pathogens isolated from dairy herds in the west littoral region of Uruguay *Acta Vet Scand.* **43** 31-41

[2] Hillerton J E and Semmens J E 1999 Comparison of treatment of mastitis by oxytocin or antibiotics following detection according to changes in milk electrical conductivity prior to visible signs *J. Dairy Sci.* **82** 82-93

[3] Waage S, Jonsson P and Franklin A 1994 Evaluation of a cow-side test for detection of Gramnegative bacteria in milk from cows with mastitis *Acta Vet. Scand.* **35** 200-207

[4] Antonov L V, Makarov K V and Orlov A A 2015 Development and experimental research on production data analysis algorithm in livestock enterprises *Procedia Engineering* **129** 664-669

[5] Norberg E 2005 Electrical conductivity of milk as a phenotypic and genetic indicator of bovine mastitis *Livestock Production Science* **96** 129-139

[6] Janzekovic M, Brus M, Mursec B, Vinis B, Stajnko D and Cus F 2009 Mastitis detection based on electric conductivity of milk *Journal of Achievements in Materials and Manufacturing Engineering* **34** 39-46

[7] Kate M 2000 Millar respect for animal autonomy in bioethical analysis: the case of automatic milking systems (AMS) *Journal of Agricultural and Environmental Ethics* **12** 41-50

[8] Lovendahl P and Chagunda M 2006 Assessment of fertility in dairy cows based on electronic monitoring of their physical activity *Proceedings of the 8th World Congress on Genetics Applied to Livestock Pro-duction* 496-500

[9] Steeneveld W, van der Gaag L C, Ouweltjes W, Mollenhorst H and Hogeveen H 2010 Discriminating between true-positive and false-positive clinical mastitis alerts from automatic milking systems *Journal of Dairy Science* **93(6)** 2559-2568

[10] Sun D, Li C, Gu C, Chen J, Qu Y, Wang X, Gao J, Wei S, Wang J, Wu R and Guo D 2015 Analysis of Mineral Elements, Metabolism, and Inflammation Indexes in the Plasma of Dairy Cows Suffering from Different Degrees of Lameness *Biological Trace Element Research* **168** 372-379
[11] Zarchi H and Ingi R 2009 Improving Oestrus Detection in Dairy Cows by Combining Statistical Detection with Fuzzy Logic Classification *Proceedings Workshop on Advanced Control and Diagnosis*
[12] Yakovleva T V 2017 Determining the Phase Shift of Quasiharmonic Signals through Envelope Analysis *Computer Optics* **41(6)** 950-956 DOI: 10.18287/2412-6179-2017-41-6-950-956

Optimization of conditions of a heterogeneous catalytic reaction

S N Koledin[1], K F Koledina[1,2], I M Gubaydullin[1,2] and A F Mullayanova[2]

[1]Ufa State Petroleum Technological University, Ufa, Bashkortostan, Russia, 450062
[2]Institute of Petrochemistry and Catalysis, Russian Academy of Sciences, Ufa, Bashkortostan, Russia, 450075

Abstract. A kinetic model of the heterogeneous catalytic reaction of ethanol dehydrogenation to ethyl acetate is considered. The kinetic model is used to solve the problem of optimization of the reaction conditions taking into account the reactant adsorption and desorption on the solid catalyst.

1. Introduction

In the mathematical modeling of heterogeneous reactions, it is necessary to use kinetic equations as the Langmuir-Hinshelwood dependences or the law of acting surfaces [1, 2]. These equations are reported in the scientific literature in relation to the development of kinetic models of, for example, benzene alkylation with ethylene [3] or ethylbenzene dehydrogenation [4].

In the ethanol dehydrogenation to ethyl acetate, bioethanol is used [5]. This is a green chemistry reaction. The target reaction product, ethyl acetate, is used as a solvent, in the production of paint and varnish materials and for other applications. This reaction proceeds to a high conversion and is highly selectivein the presence of copper-zinc-chromium catalysts [6]. The gas-phase reaction on transition metal catalysts is heterogeneous.

2. Kinetics of ethanol dehydrogenation to ethyl acetate

A kinetic study of ethanol dehydrogenation to ethyl acetate has been reported in [6]. The reaction scheme indicating the key reversible steps is presented (Table 1). The kinetic and adsorption parameters have been determined.

Table 1. Scheme of chemical transformations and kinetic equations for ethanol dehydrogenation to ethyl acetate.

Scheme of chemical transformations	Kinetic equations
C_2H_5OH (X_1)↔ CH_3CHO (X_2) + H_2 (X_3)	$w(1) = k(1)*\theta(1) - k(3)*\theta(2)*\theta(3)$
C_2H_5OH (X_1) + CH_3CHO (X_2)↔ $CH_3COOC_2H_5$ (X_4) + +H_2 (X_3)	$w(2) = k(2)*\theta(1)*\theta(2) - k(4)*\theta(4)*\theta(3)$

where X_1 is ethanol C_2H_5OH; X_2 is acetaldehyde CH_3CHO; X_3 is hydrogen H_2; X_4 is ethyl acetate $CH_3COOC_2H_5$; $\theta(i)$ is the fraction of the catalyst surface occupied by i-th component (1), (2), i=1,…,4; *k(j)* are the rate constants of the steps (the units of measurement depend on the order of the particular step); and *w(1), w(2)* are the reaction rates.

$$\theta^* = 1 - \sum_{i=1}^{4} \theta(i), \quad (1)$$

where θ^* is the unoccupied fraction of the catalyst surface at any instant of time.

$$\theta(i) = \frac{b(i) * \frac{x(i)}{V}}{1 + \sum_{i=1}^{4} b(i) * \frac{x(i)}{V}}, \quad (2)$$

where $b(i)$ is the adsorption coefficient of i-th component according to the mathematical model based on the Langmuir-Hinshelwood mechanism for adsorption of reaction components on a solid catalyst, $x(i)$ is the concentration of i-th component, [mol/L], V is the gas-phase volume of the reaction mixture, [m^3], which is determined from the geometric dimensions of the reactor [6].

The activation parameters for the rate constants and adsorption coefficients have been determined earlier [6].

The developed kinetic model of the complex catalytic reaction can be used to optimize the reaction conditions [7-13].

3. Variable parameters and objective functions for optimization of conditions of the catalytic heterogeneous reaction of ethanol dehydrogenation to ethyl acetate

The variable parameters of optimization as appliedto problems of chemical kinetics may include the temperature, type of the catalyst, concentration of the catalyst, pressure and so on. The experimental studies of this reaction were carried out at different temperatures and pressures. These parameters affect the volume of the reaction mixture. For the reaction in question, we take temperature and pressure as the variable parameters. The physico-chemical constraints have been reported previously [6].

In the general form, the optimization criterion based on the kinetic model has the form [14]

$$R(\mathbf{x}, \mathbf{x}^0, t^*, \boldsymbol{\eta}, \mu, T, P) \to \max, \quad (3)$$

where x is the concentration vector of compounds, mol/L; x_0 is the vector of the initial concentrations of compounds, mol/L; η is the vector of compound weights; μ are additional expenses; t^* is reaction time, min; T is temperature, $°C$, P is pressure, atm.

The conditions of ethanol dehydrogenation to ethyl acetate were optimized according to relation (3) considering the following criteria reported previously [15].

1) Yield of the target product x_{prod}, which depends on the temperature and pressure:

$$R_1(T, P) = x_{prod}(T, P) \to \max. \quad (4)$$

The target product of the reaction is ethyl acetate (X_4).

2) Yield of the by-product $x_{by-prod}$, which depends on the temperature and pressure:

$$R_2(T, P) = x_{by-prod}(T, P) \to \min. \quad (5)$$

The by-product of the reaction is acetaldehyde (X_2).

4. Parallelization of the computational process

Mathematical modeling of complex chemical reactions faces the following difficulties:

– there are several hypothetical reaction mechanisms and each of them should be addressed to choose the best one;

– several experiments carried out under different conditions are available (usually more than five); considering the existing experimental error, all of them should be addressedand the best two (or three) should be chosen, that is, those for which the calculated values coincide most closely with the experimental data;

– each of the kinetic parameters is determined ambiguously, being dependent on the correct choice of the initial approximation proceeding from some physicochemical assumptions.

The successive solution of these problems requires a lot of time (from several months to a year). It is proposed to arrange the problems in groups and carry out computations for these groups in parallel, with the computation within each group being carried out sequentially (Fig. 1) [16, 17].

The first group combines all of the mechanisms proposed for the given reaction (Stage 1). For each mechanism, all available experiments for this reaction are considered (Stage 2). For each experiment, the parametric plane is split to search for kinetic parameters (Stage 3).

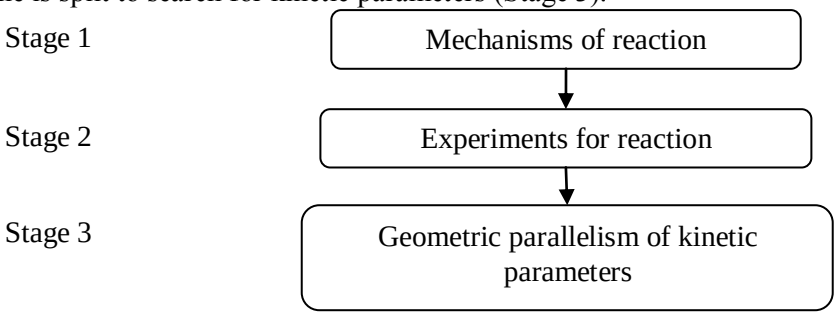


Figure 1. Model of parallelization of the computational process.

For large dimensions of the definition domain, population-based algorithms for solving the optimization problem are used. Parallelization models are employed for these algorithms [18]. The use of the genetic algorithm for solving the optimization problem implies the use of the following parallelization models:

1) Island Model

A multi-population is created as several subpopulations (islands), the number of which is equal to the number of processors used $S = \bigcup_{i=1}^{|P|} S_i$; $|S| = \bigcup_{i=1}^{|P|} |S|_i$ (where S is multi-population, S_i are subpopulations (islands), |P| is the number of processors) (Fig. 2). Each island is treated by a separate processor. During a specified period of time, the subpopulations develop independently and, after that, the islands are synchronized using a special process, in which data exchange takes place.

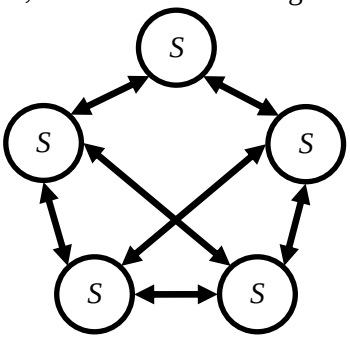


Figure 2. Island model of parallelization of population-based algorithms.

2) Cellular Model

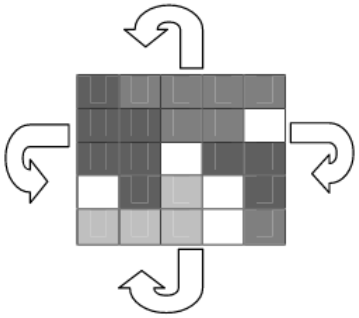


Figure 3. Cellular model of parallelization of population-based algorithms.

The area is split into parts according to the number of processors. The left-hand side is connected to the right-hand side and the upper side is connected to the lower side, thus forming a torus (Fig. 3).

Each process can interact only with four neighbors (above, below, on the left, and on the right). Each cell contains only one solution (individual). Each process will choose the best individual among the neighbors, cross it over with the individual from its own cell, and place one offspring into its cell instead of the parent. The operation of this algorithm brings about effects resembling those in the island model. Initially, all individuals have random fitness (in Fig. 3, it is defined by colors). After several generations, relatively small regions of similar-fitness individuals are formed. As the algorithm operates, these regions grow and compete with one another.

3) Global Worker/Farmer model

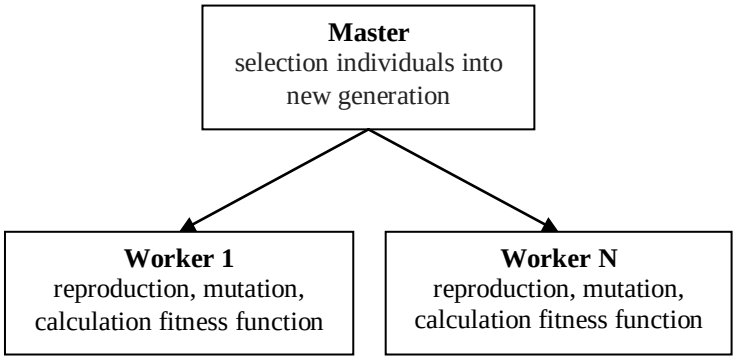


Figure 4. Global Worker/Farmer model of parallelization of population-based algorithms.

"Workers" (workstations) are responsible for reproduction, mutation, and calculation of the fitness function for selecting individuals to the new generation. All individuals created and evaluated by the "workers" are delivered to the "farmer" workstation, which then selects the individuals to the new population in conformity with fitness evaluation. The selected individuals are transferred by the "farmer" to the "workers" stations (Fig. 4).

The reaction conditions were optimized using the genetic algorithm to solve the optimization problem and the island model to parallelize the computational process [19].

The input processor receives a local range of variable parameters. The output processor givesout the optimal values for the variable parameters out of the indicated range. The values from this range are selected and used to solve the direct problem, that is, to solve a system of ordinary non-linear differential equations for determination of the objective functions. The optimal variable parameters for the given processor local range are found (Fig. 5).

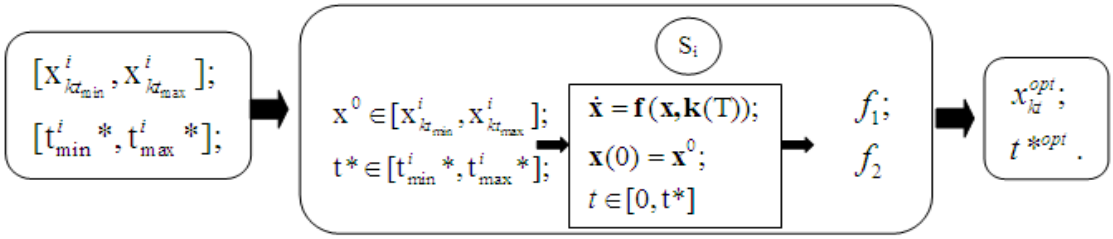


Figure 5. Treatment of the subpopulations by each processor of the island model of parallelization.

The optimized pressure and temperature for ethanol dehydrogenation to ethyl acetate were determined using the island model of parallelization. The time of the computational experiment was estimated (Fig. 6) (the calculations were carried out on a Intel Core I5 7th Gen quad-core PC).

The efficiency of utilizationof processors by the parallel algorithm for problem solution is defined by the relation

$$E_p = \frac{S_p}{p},$$

i.e., it is the mean fraction of algorithm execution time during which the processors are actually occupied by solution of the problem (Fig. 7).

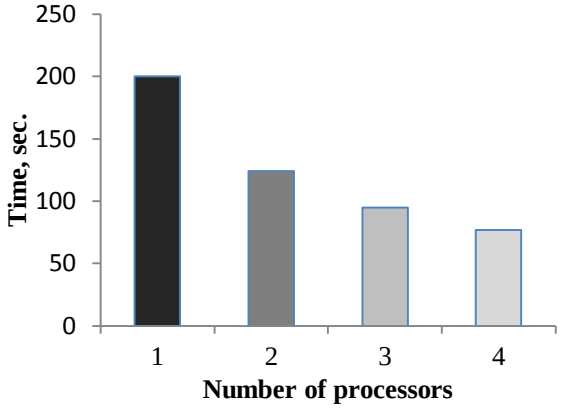


Figure 6. Time of the computational experiment with the genetic algorithm for different numbers of processors.

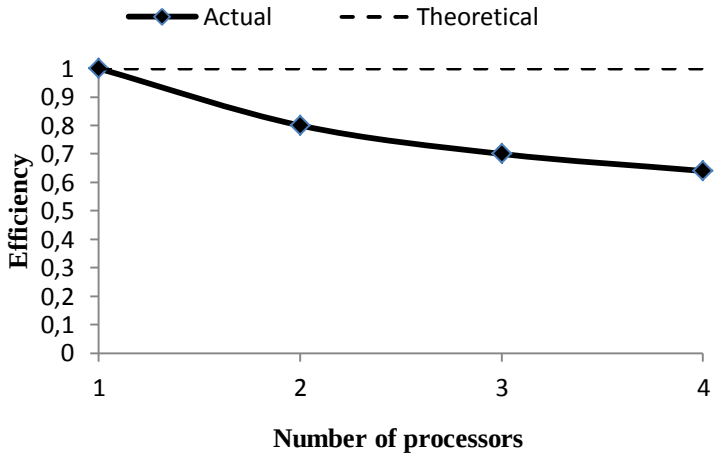


Figure 7. Efficiency of parallel program execution for the genetic algorithm.

The parallelization efficiency for the considered reactions was 65%. Apparently, the ideal efficiency of parallelization is not achieved because of the time spent for data synchronization between the islands.

5. Multiobjective optimization of the conditions of a catalytic heterogeneous reaction

The major challenge in solving an optimization problem in chemistry is that all the theoretical works on optimization have addressed each criterion separately. However, in the last decades, numerous efficient evolutionary algorithms of multiobjective optimization have been proposed. These algorithms take into account all scientific experience in the approximation of the Pareto domination region and genetic algorithms. The computational power has markedly increased, which allows high-throughput computing to be accomplished over reasonable periods of time [20-27]. The multiobjective optimization involves the search for several Pareto-optimal solutions. The set of optimal values of variable parameters is the Pareto set. The objective functions in this set are called the Pareto front. The algorithm used most often to solve the multiobjective optimization problem is the NSGA-II algorithm [22, 23]. According to this algorithm, the generated individuals are ranked, each one being assigned with a particular rank. The non-dominated points have the first rank, the points that are dominated only by first-rank points have the second rank, and so on. The crowding of the obtained individuals is also evaluated; the greater the distance between them, the higher the population diversity. In every

iteration, the offsprings are selectedconsidering the rank and crowding (proximity) of individuals. Subsequently, the best points are chosen in the iteration via crossing over and mutation, which ensures the diversity of the next population. The parents and offsprings are combined into one population corresponding to the best solutions, and so on.

The problem was solved in the information system developed previously [28-31]. The optimization criteria include the yield of the target product and the yield of the by-product. The results of computational experiments are shown in Figs. 8-9.

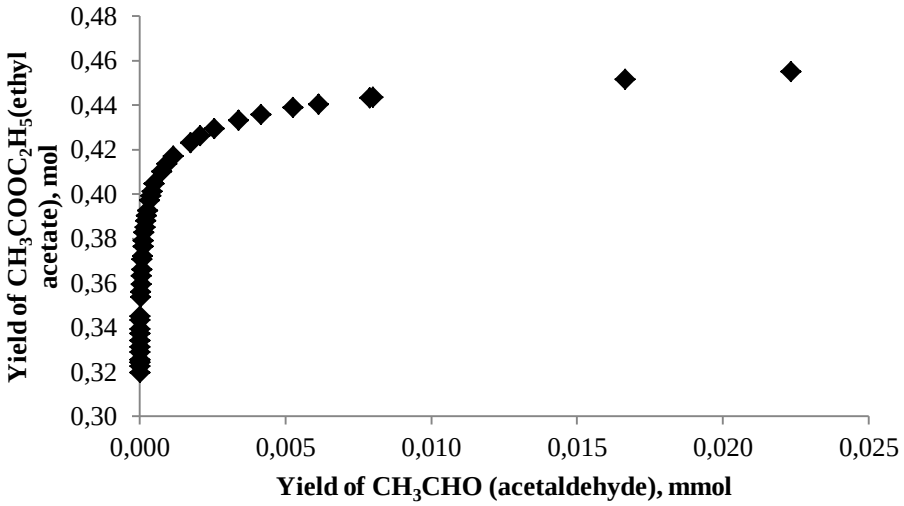


Figure 8. Pareto front approximation with the NSGA-II algorithm for ethanol dehydrogenation to ethyl acetate.

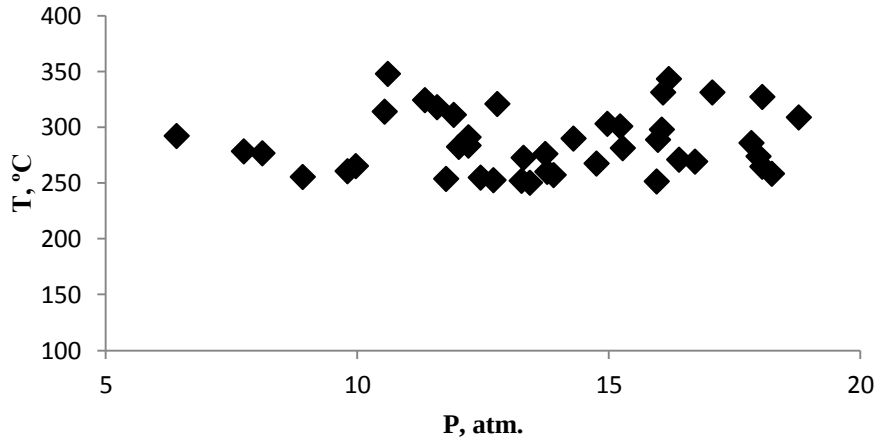


Figure 9. Pareto set approximation with the NSGA-II algorithm for ethanol dehydrogenation to ethyl acetate.

The resulting Pareto front and set approximations for ethanol dehydrogenation to ethylacetate allow the decision maker to choose the reaction conditions through the comparison of the Pareto set and front approximations for the corresponding objective functions (4) and (5) (Fig. 8, 9). Whenever it is necessary to maximize the yield of the target product (the target product is highly valuable) or minimize the yield of the by-product, the corresponding approximated values of the Pareto set and front can be used.

Thus, the optimal conditions for the complex heterogeneous catalytic reaction were studied by multiobjective optimization methods on the basis of a kinetic model of the reaction.

6. References

[1] Slinko M 2007 *Theoretical Foundations of Chemical Technology* **41(1)** 16-34
[2] Froment G 2005 *Catal Rev Sci Eng* **47(1)** 83-124
[3] Inui K, Rurabayashi Tand Sato S 2002 *J Catal* **212(2)** 125
[4] Bolder F 2008 *Ind Eng Chem Res* **47(19)** 4796
[5] Elliott D and Pennella F 1989 *J Catal* **119** 359
[6] Men'Shchikov V, Gol'Dshtein L and Semenov I 2014 *Kinetics and Catalysis* **55(1)** 12-17
[7] Spivak S, Koledina K, Koledin S and Gubaidullin I 2017 *Journal of Applied Informatics* **12(1)** 39-49
[8] Koledina K, Koledin S and Gubaydullin I 2017 *CEUR Workshop Proceedings* **1966** 5-9
[9] Gubaydullin I, Koledina K and Sayfullina L 2014 *Eng J* **18(1)** 13-24
[10] Nurislamova L, Gubaydullin Iand Koledina K 2015 *Reac Kinet Mech Cat* **116(1)** 79-93
[11] Nurislamova L, Gubaydullin I, Koledina K and Safin R 2016 *Reac Kinet Mech Cat* **117(1)** 1-14
[12] Maskov D and Gubaidullin I 2012 *Bulletin of Omsk University* **64(2)** 182-184
[13] Keil F 2013 *Computers & Mathematics with Applications* **65(10)** 1674
[14] Koledina K, Koledin S, Shchadneva N and Gubaidullin I 2017 *Rus J Phys Chem A* **91(3)** 444-449
[15] Koledina K, Koledin S, Schadneva N, Mayakova Y and Gubaydullin I 2017 *Reac Kinet Mech Cat* **121(2)** 425-428
[16] Gubaydullin I, Lind Yu and Koledina K 2012 *Num Methods Prog: Adv Comput* **13 (1)** 236-244
[17] Baynazarova N, Koledina K and Pichugina D 2016 *CEUR Workshop Proceedings* **1576** 425-431
[18] Awrejcewicz J, Lind Y, Gubaydullin I and Koledina K 2012 *Theoretical and Applied Mechanics Letters* **2(4)** 043
[19] Koledina Kand Gubaidullin I 2013 *Science and Education: a scientific edition of the MSTU N E Bauman* **7** 385-398
[20] Karpenko A 2016 *Information and Mathematical Technologies in Science and Management* **2** 8-17
[21] Sobol I and Statnikov R 2006 *The Choice of Optimal Parameters in Problems With Many Criteria: Textbook. Manual for Universities* (Moscow: Drofa) p 175
[22] Deb K, Mohan M and Mishra S 2003 *Evolutionary Multi-Criterion Optimization* **2632** 222-236.
[23] Kalyanmoy D, Pratap A, Agarwal S and Meyarivan T 2002 *IEEE Transactions on Evolutionary Computation* **6(2)** 182-197
[24] Abramson M, Charles A, Dennis J, Digabel J and Digabel S 2009 *SIAM Journal on Optimization* **20(2)** 948-966
[25] Alonso P, Argüeso F, Cortina R, Ranilla J and Vidal A 2013 *J Math Chem* **51(4)** 1153-1163
[26] Alonso P, Argüeso F, Cortina R, Ranilla J and Vidal A 2012 *J Math Chem* **50(2)** 410-420
[27] Canedo M and González-Hernández J 2011 *J Math Chem* **49(1)** 163-184
[28] Vovdenko M, Gubaidulin I, Koledina K and Koledin S 2017 *CEUR Workshop Proceedings* **1966** 20-23
[29] Koledina K and Gubaidullin I 2016 *Rus J Phys Chem A* **90(5)** 914-921
[30] Zainullin R, Koledina K, Akhmetov A and Gubaidullin I 2017 *Kinetics and Catalysis* **58(3)** 279-289
[31] Koledin S, Koledina K, Gubaidullin I and Spivak S 2016 *Chemical Industry today* **10** 24-35

Acknowledgments
The reported study was funded by RFBR according to the research projects № 18-07-00341, 18-37-00015 and by the President of the Russian Federation SP-669.2018.5 stipends.

Myocardial infarction detection using wavelet analysis of ECG signal

N S Davydov[1] and A G Khramov[1]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086

Abstract. The new method of ECG signal segmentation and feature extraction is developed and tested with the use of SVM classification method in this paper. The algorithm based on discrete wavelet transformation and several new steps of beat extraction is proposed in this paper. We found out that the most part of algorithms is not easy to understand and realize. During the developing we made each step of algorithm clear and valid, so anyone can realize it when it needs. The output of preprocessing method is common standard segment, which accumulate all information from ECG signal and, in fact, this segment is the trend line of cardiac cycle in ECG signal. The feature extraction method based on cross wavelet transformation is used in this paper. Cross wavelet spectrum and wavelet coherence are two characteristics which are used for feature extraction. Classification method based on Support Vector Machine and several different transformation kernels is used and sensitivity, specificity and accuracy for each algorithm is calculated. The best result is 0.87, 0.8481 and 0.8603 for sensitivity, specificity and accuracy, respectively.

1. Introduction

The raw signal of ECG is a time series with a specific duration and sampling frequency which depends on type of using device. The first step of ECG analysis is the preprocessing of signal. It can be skipped, but the whole signal is hard to analyzed, and it may take more time than the analysis of a common segment. Main step of signal preprocessing is cardiac cycle extraction. Discrete wavelet transformation is used in this paper for a trend line deleting, noise filtration and cycle extraction [1, 2].

After extraction of the set of cardiac cycles, the average standard segment is calculated. This standard segments represents all information which the signal consists. Each segment is interpolated and normalized according to the requirements of the segment model. A similar procedure is performed for each of the signals.

The next step in the analysis is the cross wavelet transformation of the standard signal segment and the template signal segment of the healthy patient. Cross wavelet transformation is based on continuous wavelet transformation and helps to find the connection between the frequencies of two input signals. It is often used in the analysis of geographic data [3]. Unlike the Fourier transform, which represents the amplitude-frequency characteristic of a signal, a continuous wavelet transformation gives a representation in the time-frequency domain. Thus, the wavelet correlation and cross wavelet spectrum obtained after the transformation make it possible to isolate the frequency-time regions of interest for distinguishing features.

There are many methods for data classification. Many of them are based on convolutional neural networks and give different results depending on the considered features. For example, the method based on shallow convolutional neural networks and the analysis of the second, third and AVF leads

gives sensitivity, specificity and accuracy values of 0.8533, 0.8409 and 0.8454, respectively [4]. There are also methods for classifying tree-based solutions and blind signal separation [5, 6]. However, in this paper, the method of Support Vector Machine will be used to classify the obtained features, which is also widely used in the classification of data [7]. Also in this paper, to obtain a classification result close to that which can be obtained by practical application of the algorithm, the technique of cross-validation data is used. All signals of the electrocardiogram and patient data were taken from the database of the Massachusetts Institute of Technology.

2. Signal preprocessing

The first step is to bring the signal to the required view of a common standard segment is signal segmentation to cardiac cycles, which will allow working with parts containing the most important information about the patient. Segmentation is carried out using a discrete wavelet transform.

2.1. Signal segmentation

The discrete wavelet transform operates as follows: the input signal is analyzed in different frequency bands with different resolutions by decomposing the signal into rough approximation and details. During the transformation, two sets of functions perform their work: scaling functions and wavelets corresponding to low-frequency and high-frequency filters. The signal passes through two types of filters and the output is obtained by the coefficients of the signal decomposition using a high-frequency and low-frequency filter. At the next level, the procedure is repeated for the output of the low-pass filter, but with a lower frequency resolution and only for half the frequency band.

Thus, the time resolution of the signal is reduced by half, the time-based decimation takes place, but the frequency resolution is doubled with each level. This procedure may be performed until there are 2 counts left. For instance, a signal of 512 samples can be decomposed into 8 levels of discrete wavelet transform. As a result of the decomposition at each level, the level coefficients at the output of the high-pass filter and the additional level coefficients will be obtained at the output of the low-pass filter if further decomposition is not performed.

Input signals used in this study have a different length and amplitude. Also, each of them has a wandering trend line that affects the position of each heart cycle and can cause undesirable distortions during further processing. An example of a raw ECG signal with a present wandering trend line is shown in Figure 1.

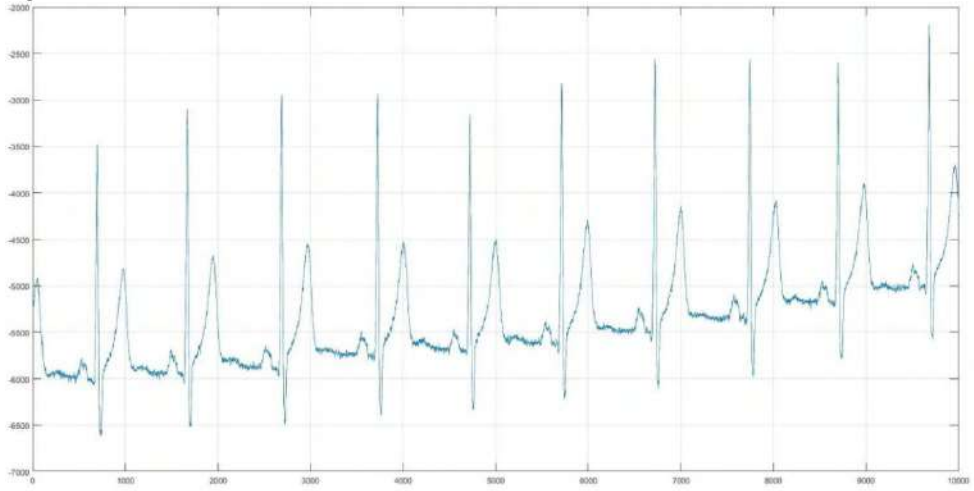


Figure 1. Raw ECG signal with a trend line.

The wandering of the trend line is associated with the physical movements of the chest in the process of patient breathing. This noise is a low-frequency physical noise and its elimination is necessary to obtain a stable signal and its subsequent processing.

The frequency range of the trend line is between 0 and 0.5 Hz. The method used to remove the wandering trend line is based on the expansion using a discrete wavelet transform to 8 levels and the

subsequent reconstruction of the signal at 8 levels without using the additional level coefficients. Due to the fact that the signal was restored from those coefficients that contain all the main details of the original signal without the trend line. Also, to prepare the signal for further processing, it is required to remove high-frequency noise. This can be done by decomposing the signal into 3 levels of discrete wavelet transform and restoring the signal by the coefficients of the third and additional levels, thereby removing high-frequency noise from the signal.

It is necessary to isolate the most important point of each cardiac cycle-the R-peak to perform signal segmentation. The frequency range corresponding to the QRS-complex and the R-peak is between 3 and 40 Hz, which means that the signal decomposition into 5 levels of discrete wavelet transform and subsequent signal recovery by the coefficients of 4 and 5 levels is required. After the signal is restored, it will contain only the peak values at time positions that correspond to the R peaks in the original signal. Detection of peaks occurs by successively searching for a maximum with a sliding window width equal to half of the sampling rate and a maximum value threshold of at least 10% of the amplitude of the maximum point in the reconstructed signal. The result of projecting the found points to the source signal is shown in Figure 2.

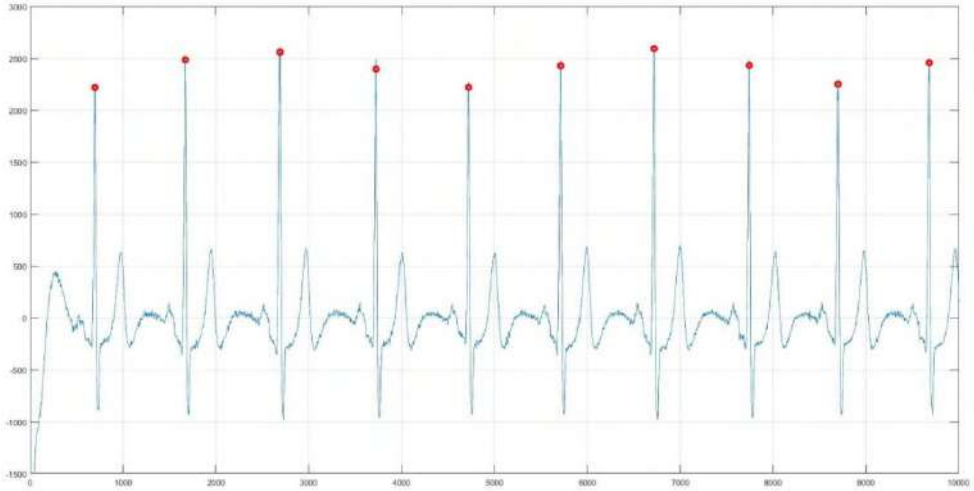


Figure 2. Detected R peaks.

The final step of the segmentation of the ECG signal will be the separation of the signal into cardiac cycles relative to the R-peak. The length of the heart cycle of each patient is different and depends on its pulse and how long the electrocardiogram is taken. In this regard, the following formula is developed to calculate the length of the patient's cardiac cycle:

$$L = \frac{LengthOfSignal}{Pulse}.$$

After length calculating of the heart cycle, it is necessary to isolate it relative to each R-peak. For this, an interval equal to one third of the length of the entire cycle to the left of the R-peak is taken and an interval equal to two-thirds of the length of the entire cycle to the right of the R-peak. However, there is a possibility that the first or last heart beat is not recorded completely, so the presence of this cardiac cycle will interfere with further processing of the signal. To avoid such an error, all cardiac cycles except for the first and the last detected is used in further study.

Thus, the signal is uniformly divided into cardiac cycles of the same length, but for each patient this length will be different. The whole set of segments obtained as a result of dividing the ECG signal is demonstrated in Figure 3.

2.2. The set of cardiac segments processing

The first thing that needs to be done to bring a set of segments to the form of a common standard segment is to interpolate all cardiac cycles and normalize each of them. As the basic requirements for the model of the common standard segment, a length of 1000 samples is chosen and the maximum R-peak value was equal to 1.

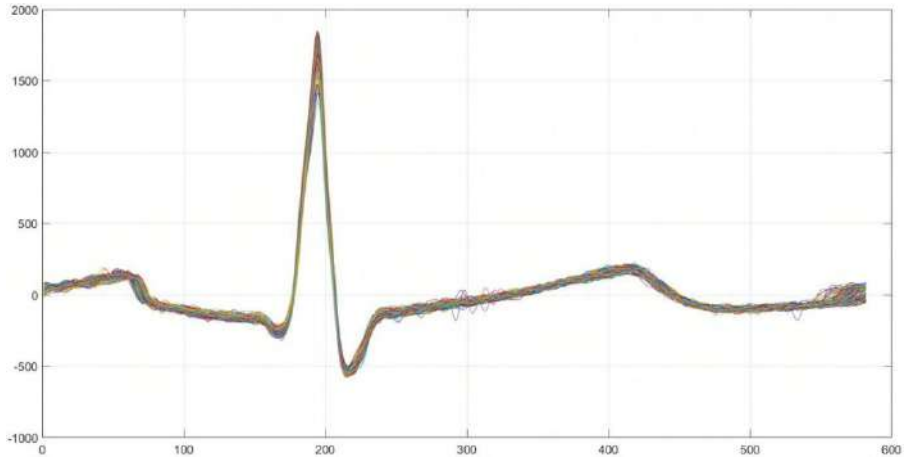


Figure 3. The set of cardiac cycles of ECG signal.

Interpolation is required in order to add points to those segments that are less than the required length and to shrink those segments that are longer than the required length. To perform this operation, we used linear interpolation of two points with fixation of the first and last reference points.

Normalization requires that all segments have the same amplitude range. The difference in the amplitude range is due to the physical characteristics of the patient's bodies and the sensitivity of the instrument.

The final step of bringing the ECG signal to the form of a common standard segment is an averaging of segments set. This procedure is necessary in order to obtain the overall shape of the entire set of segments and to drown out the high-frequency. It is worth mentioning that this step can drown erroneously selected cardiac cycles, which are not. For example, if any physical noise with frequency equal to the frequency of the QRS complex was recorded during the ECG removal, it can be defined as part of the cardiac cycle. In this regard, averaging will significantly reduce the contribution of the defective cycle to the general information segment of the signal. Also, in the case of pathological signals, various heart murmurs and heart damage can affect the waveform. However, the detection of these damages can be based on the selection of the form – the average standard segment, and the subsequent finding of the greatest deviation from this segment.

The result of final steps of signal processing is illustrated for set of segments from Figure 3 on Figure 4.

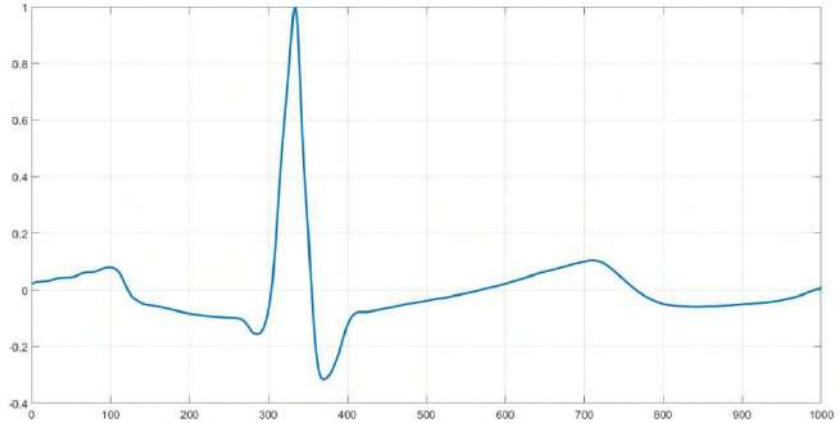


Figure 4. Common standard segment of ECG signal.

3. Features extraction

Continuous wavelet transformation is an efficient way to get time-frequency characteristics of input signal. Also it is possible to frequency comparation of two different signals by performing a cross wavelet transformation. The result of the cross wavelet transformation will be two characteristics: cross wavelet spectrum and wavelet coherence.

The cross wavelet spectrum shows at what frequencies and at what time point the signals are different or similar. If the signals are different, then the image on the wavelet spectrum is darker, and in the other case lighter, which will further distinguish the signs of the difference between the forms of the common standard segment and the template. The cross wavelet spectrum of the two signals is calculated by the following formula:

$$\Psi^{xy} = \Psi^{x}\Psi^{y*}.$$

The wavelet coherence shows how the frequencies of the two signals under consideration are interrelated, since the wavelet transformation treats the signals precisely in the frequency-time range. The wavelet coherence of the two signals is calculated by the following formula:

$$R_{\tau}(s) = \frac{S(s^{-1}\Psi_{\tau}^{xy}(s))}{S(s^{-1}|\Psi_{\tau}^{x}(s)|) \cdot S(s^{-1}|\Psi_{\tau}^{y}(s)|)}$$

To calculate features for classification, it is necessary to sum the values of the cross wavelet spectrum and wavelet coherence, thereby reducing the dimensions of the available features in the form of a spectrum and a coherence.

The first step is the summation over the time interval, which corresponds to the place of interest in the heart cycle. In the case of myocardial infarction, the QT interval is considered, which is between 250 and 750 time points. After this procedure the scale vectors of chosen time interval is taken and illustrated in Figure 5.

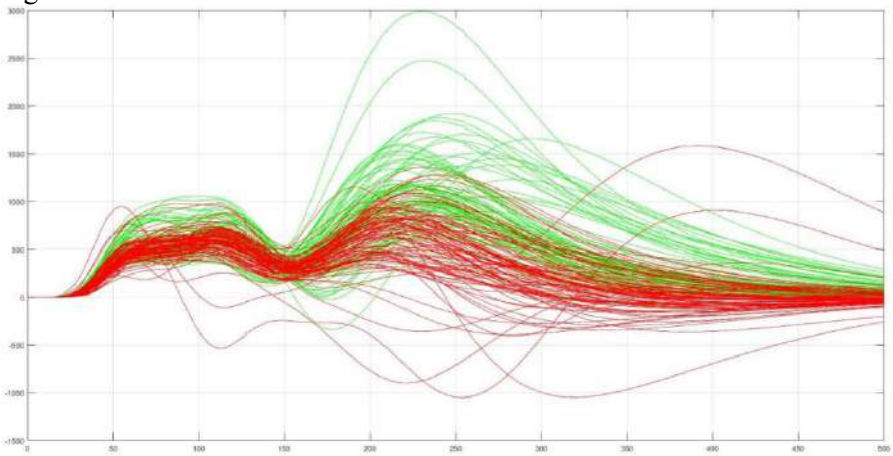


Figure 5. The scale vector of QT interval.

The next summation must go on a scale, however, the selection of the most correct interval is not obvious, therefore, an interval from 150 to 400 scale is chosen, but later, after the classifiers are formed, additional numerical experiments will be carried out in order to determine the most effective summation interval by scale. The final characteristics are formed by the following formulas:

$$pp = \sum_{s=150}^{400} \sum_{\tau=250}^{750} \Psi^{xy}(\tau, s),$$

$$pa = \sum_{s=150}^{400} \sum_{\tau=250}^{750} R(\tau, s).$$

The characteristic is a point in a two-dimensional space with coordinates *(pp, pa)*, which will allow visualizing the feature space and the classification process.

4. Classification of extracted features

An experimental study of the developed algorithms for signal processing and features extraction is tested with use of Support Vector Machine with a different type of hyperplane constructions and kernel functions.

The input data for the experiment are 179 ECG records with a sampling frequency of 1000 Hz. 100 records belong to patients with myocardial infarction, and 79 of them belong to healthy patients. In

each ECG record, a signal was taken from the first standard lead, since it contains the most common information and the most common signs of myocardial infarction. All data were taken from the database of the Massachusetts Institute of Technology, provided by the Physikalisch-Technische Bundesanstalt.

Each of the records has been pre-processed, the noise is filtered out. The resulting data set was classified by three kinds of trained classifiers and cross-validated using the leave-one-out method. The result for accuracy of each classification is provided in the table 1.

Table 1. Accuracy of classification.

Method	Kernel functions						
	Linear	Quadr.	3 order	4 order	5 order	6 order	Gaussian
QP	0,8603	0,8380	0,8547	0,8492	0,8603	0,8547	0,8324
SMO	0,8547	0,8324	0,8324	0,8492	0,8659	0,8547	0,8268
LS	0,8268	0,8659	0,8380	0,8492	0,8380	0,8436	0,8380

From the obtained experimental results it is clear that the classification based on the algorithm of sequential minimal optimization (SMO) with a kernel function in form of a fifth-order polynomial and the method of least squares with a quadratic kernel turned out to be most efficient. Classification by a polynomial of the fifth order is given as an example in Figure 6.

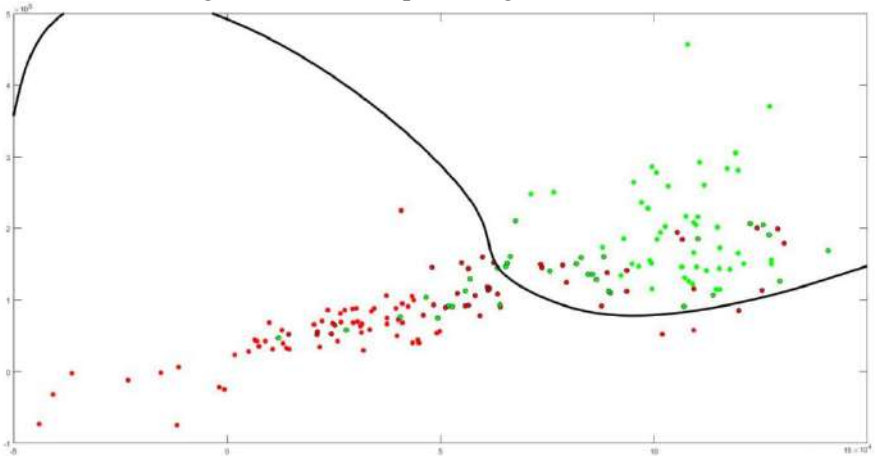


Figure 6. Classification by SVM with a fifth-order polynomial kernel.

Further, it is required to compare the obtained results with other methods of detecting myocardial infarction. It is worth noting that different articles use different methods for distinguishing signs from the ECG signal, and in some of them the entire signal is analyzed, so that a direct comparison of the result of classification with each of them is not completely correct. Also, the difference between the represented algorithms for extracting features can be that the algorithm considers only one ECG lead out of twelve, while the other algorithms are based on several leads.

The first algorithm considered for comparison is based on shallow convolutional neural networks (SCNN), in which the entire signal in three leads is analyzed and the differences between healthy and unhealthy signals are searched. Also, in this algorithm, cross-validation of the data was carried out. The following algorithm is based on the method of blind signal separation. The problem of blind signal separation (BSS) and noise is an extremely urgent and complex mathematical task. Algorithms based on the method of support vectors have different directions. In one of the articles considered, the ECG signal is reduced to a three-dimensional form and the three-dimensional signs of the disease are calculated and the subsequent application of the multilayer support vector method (MSVM) is used. The comparison of these method in table 2.

Table 2. Comparison with other methods.

Methods	Sensitivity	Specificity	Accuracy
This study	0,8700	0,8481	0,8603
SCNN	0,8533	0,8409	0,8454
BSS	0,9837	0,9404	0,9677
MSVM	0,8682	0,9105	0,8543

5. Conclusion

In this study, a method for extracting features of myocardial infarction from an ECG signal with the use of a cross-sectional wavelet transformation was developed. As a result of the work, the features were obtained, on the basis of which classification was carried out by the method of support vector machine with different approaches to constructing a separating hyperplane and various transformation kernels. The best result was obtained using the method of support vectors, which was based on sequential minimal optimization with a transformation kernel in the form of a polynomial of the fifth degree. As a result, the sensitivity values (0.8700), specificity (0.8481) and accuracy (0.8603) were calculated. Classification was carried out using the method of cross-sectional cross-checking data to obtain the result, which most adequately shows the work of the algorithm in practical application.

As a further development of the study, it is planned to use neural networks and gradient-boosting methods to improve the result of the classification. Also during the research it was noted that taking the middle segment as a common standard is not always true. In this connection, a study will be carried out when the segment that is most different from the average is taken as the general standard segment. However, in this case, a more efficient method for filtering the signal from physical noise will be required.

6. References

[1] Haddadi R, Abdelmounim E and Hanine M 2014 Discrete wavelet transform based algorithm for recognition of qrs complexes *World of Computer Science and Information Technology Journal* **4** 127

[2] Chandel S and Singh K 2016 ECG denoising using wavelet transform and soft thresholding *International Journal of Advanced Research in Computer Science and Software Engineering* **6** 370

[3] Grinsted A, Moore J C and Jevrejeva S 2004 Application of the cross wavelet transform and wavelet coherence to geophysical time series *Nonlinear Processes in Geophysics* **11** 561

[4] Tahsin R and Celia S 2017 Detection of inferior myocardial infarction using shallow convolutional neural networks *IEEE Region 10 Humanitarian Technology Conference* **1** 718

[5] Khatun S and Morshed B I 2017 Detection of Myocardial Infarction and Arrhythmia from Single-Lead ECG Data using Bagging Trees Classifier *Electro Information Technology* **1** 520

[6] Devika M G and Aneesh R P 2016 Myocardial infarction fetection using hybrid BSS method *International Conference on Communication Systems and Networks* **1** 167

[7] Dhawan A, Wenzel B, George S, Gussak I, Bojovic B and Panescu D 2012 Detection of acute myocardial infarction from serial ecg using multilayer support vector machine *34th Annual International Conference of the IEEE EMBS* **1** 2704

Development of data model for the functioning of production active elements based on information interaction

I N Khaimovich[1,2] and V M Ramzaev[1]

[1]Samara University of Public Administration International Market Institute, G.S. Aksakova Street 21, Samara, Russia, 443030
[2]Samara National Research University, Moskovskoe shosse 34A, Samara, Russia, 443086

Abstract. The article discusses the design and technological preparation of production as a hierarchical active system "center – designers - technologists". A functioning model of the production of the active element, which is a model of decision making, is described in a formalized form. The most effective state, chosen by each active member as a team of a relative subdivision, may be different from the planned states that are defined in terms of the criterion which characterizes given type of production preparation in general, which leads to contradictions in the system. The article introduces quantitative assessment of inconsistency between each team and the center, and on this basis the coordinated interaction problem is stated. The authors discuss the informative way of solving the coordinated interaction problem from the standpoint of team interests: the availability of information resources in the form of PDM/PLM – the management system of product life cycle with implemented Big Data technology will allow designers and technologists to create adequate representation of actions of another team and reach optimum equilibrium. The mechanism of effective design and engineering solutions formation is connected with the mechanism of versions and organization of information processing level Product Event Management.

1. Introduction

In modern conditions, a unified information space (UIS) is actively used in the organization of machinery production, which includes all information related to the product and provides information support for all stages of the product life cycle, i.e. the UIS should become the only reliable source of latest data. The foundation of information data of the enterprises is the data obtained in the process of design and technological preproduction. As when making agreed decisions between design and technological services, it is necessary to make multiple changes, the number of design and technological documentation undergoes multiple version creation of documents in the enterprise database [1, 2]. Particularly relevant the mechanism of version creation becomes in the aerospace industry, since, on average, the aircraft consists of several million products made in different regions of the globe from different materials [3, 4], and the solution for each product is adopted using the version mechanism. To integrate such data volumes, the BIG DATA mechanism is used in product data management systems (PDM systems), for example, in the TEAMCENTER software.

To work out the version mechanism, it is necessary to place all agreed decisions between designers and technologists in the database. Let us consider necessary and sufficient conditions for ensuring coordinated interaction between the units.

The relevant objective of design and engineering production preparation (DEPP) management is the task of coordination of resource flows between separate shops and sites. The solution of this particular task is connected with defining the control actions which are plan targets for permissible set of resource flows [5, 6]. However recently to solve the problem of coordination of resource flows it is necessary to achieve balanced interaction between subsystems of DEPP based on the estimation of their activities to fulfill the required administrative objectives. The procedure of coordination in this case is understood as sequential control and achievement of balance between local targets of subsystems of DEPP which are complied with the common (global) goal of the system [7]. So far as it is supposed that active systems have their own goals, then the task description and control mechanisms work out will be made in terms of systems with active elements (AE).

Let us consider the control scheme with the help of design and engineering production preparation and formalize control problem. Hierarchical control system of DEPP, presented in the figure 1 is realized in the form of organization systems (OS), including control subsystem of upper level (center), controlled and simultaneously controlling subsystems of AE of lower level, as well as external environment. The center is the leader of OS, the active elements are the teams of designers (DAE) and production engineers (technologists) (TAE), and the external environment is the outlet market or Ordering Customer. The subject to control is the product project including the stages and resources of life cycle (LS) from designing to production. The object is product specified by Y which are characterized by a set of global parameters Q, a list of target requirements to the product (consumer property of a product), specified in the technical project on a product:

$$Q = Q(q_i^{kon}, q_j^{tec}), \quad i = 1,...,L; \quad \text{j=1,...,M}, \tag{1}$$

where q_i^{kon}, q_j^{tec} are the parameters of LC of a product according to the production design and to the technology process, L is the number of requirements to the design, M is the number of requirement to the technology process. In the present case vector Q depends only on the product design Y, i.e. in (1) vector q^{kon} is included in the explicit form, and vector of technological parameters q^{tec} - in implicit form. At any specific time t LC of a product Y is in the condition of $\mathrm{Y}(Q(...,t))$. In this regard the global target function of the center is to achieve the condition of

$$Y_{opt}^{Nesh} = Y(\mathrm{Q}(\mathrm{q}_i^{kon}, \mathrm{q}_j^{tec}, \mathrm{n}, \mathrm{t})) \tag{2}$$

for the time $t \le t_\Sigma$, when the current parameters $\mathrm{Q}(\mathrm{q}_i^{kon}, \mathrm{q}_j^{tec}, \mathrm{n}, \mathrm{t}) \approx Q$ achieve or exceed target values Q. In formula (2) n is the target figure of the amount (line) of product Y. To achieve a global task the center controls the lower level and generates target figures $x = \{Q, n, t_\Sigma\}$, and controlled subsystems of lower systems perform their implementation, developing their own tasks within AE K_{xi} and T_{xi} - for DAE and TAE correspondingly (figure 1). The management of AE is fulfilled by the groups of people - teams which are able to achieve goals independently, the latest are man-machine systems providing production yield Y_t. The supervisory activity of the center is carried out by monitoring of qualitative characteristics Y_t, which are connected with some controlled variables in design and technology of a product in DEPP [8]. If we consider the LC system processes from the point of view the document – centric approach then the output quality in DEPP may be indirectly estimated in terms of volume of project (design and technological) documentation with minimum

amount of changes. To achieve the necessary level of quality it is required to optimize target functions of TAE and DAE.

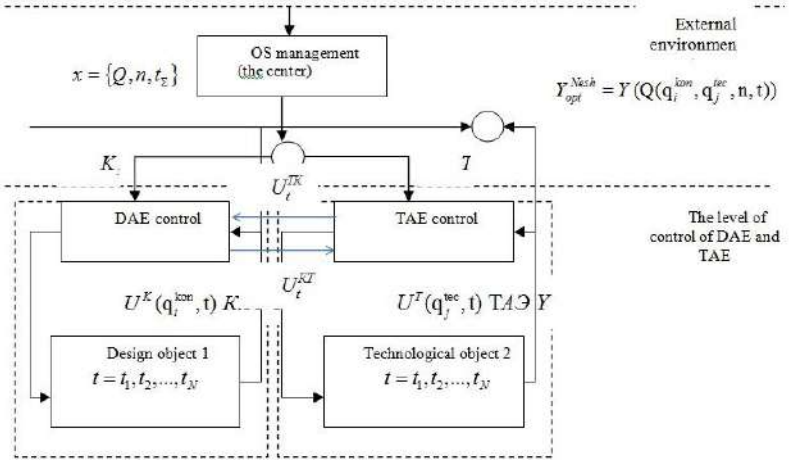


Figure 1. The structure of DEPP management system with independent elements.

Let us present a model for DEPP control system in formalized way, in this case to describe its functioning it is possible to apply the theory of reflexive games or network programming methods [9, 10, 11] as it will be shown below.

Control in DEPP is represented in the form of multivariable mathematical model:

$$\begin{cases} t_{\min} = \arg\min\limits_{t} Y_{opt}^{Nesh} \\ Y_{opt}^{Nesh} = Y\left(\mathrm{Q}(\mathrm{q}_i^{kon}, \mathrm{q}_j^{tec}, \mathrm{n}, \mathrm{t})\right) \\ \mathrm{q}_i^{kon} \le (\mathrm{q}_i^{kon})_{gr}; \mathrm{q}_j^{tec} \le (\mathrm{q}_j^{tec})_{gr} \end{cases}, \tag{3}$$

where $(\mathrm{q}_i^{kon})_{gr}, (\mathrm{q}_j^{tec})_{gr}$ are the boundary parameters values of design and technology.

As previously stated consumer characteristics Q of a product Y are shown in the design parameters q^{kon} and are realized if the technology parameters q^{tec} are achieved, though the latter is not included in Q in explicit form. In formalized form the connection between the parameters Q of a product and included parameters on the stages of design (construction) and production (technology) may be presented as follows (4) и (5):

$$Q = [q_i^{kon}] \cdot M \ , \tag{4}$$

where M is the matrix of positive impact coefficients, $m_{ij} \ge 0$.

Not evident nature of connection between Q and q^{tec} may be defined if we add the correlation matrix K between design and technological parameters according to (5):

$$[\Delta q^{tec}] = K \cdot [\Delta q^{kon}]^T, \tag{5}$$

where $\Delta q_i^{kon}, \Delta q_j^{tec}$ are the parameters value changes $q_i^{kon}, q_j^{tec}, k_{ij}$ is the correlation coefficient $-1 \le k_{ij} \le 1$.

At the moment there is no clear recommended practice to calculate the correlation matrix. To define its coefficients for instance expert evaluation method may be suggested. The given data for analysis and calculations k_{ij} are the technical specifications contained in the design documentation of a product and parameters of standard techniques and procedures with reference to facilities of the specific production.

It should be noted that between target functions DAE and TAE to achieve required level of parameters q_i^{tec}, q_j^{kon} there is often a contradiction , for instance, the requirement of weight reduction of a product, given in its design i.e. the increase of q_i^{kon} reduces its technological effectiveness i.e. reduces the value of q_j^{tec}. Consequently, if $k_{ij} < 0$, in order to make the system self-regulating i.e. to make it reach the equilibrium state according to the Nash [10], it is necessary to add the coordination process between designers and technologists, as there are some contradictions between target functions of development $q_i^{kon} \to \max$ и $q_j^{tec} \to \max$.

Let suppose that there is *Y* an object of production – a design of a product, then $\Omega(Y)$ is seen as external environment, where $Q(\{q_m^{kon}, q_n^{tec}\})$ are the parameters of production object, $w(\{p_i^{kon}, p_j^{tec}\}, \mathrm{Q})$ – production object display in design and technological documentation, where p_i^{kon}, p_j^{tec} are the action by designers and technologists, which lead to the results q_m^{kon}, q_n^{tec}. If $\{p_i^{kon}, p_j^{tec}\} \in P_{opt}$ are the combined actions, then Nash equilibrium is true in the following way:

$$E_N(\theta) = \{\{p_i^{kon}, p_j^{tec}\}_{i\in I, j\in J} \in P_{i,j} \mid \forall i \in I, \forall j \in J, \forall\{\mathrm{y}_i^{kon}, \mathrm{y}_j^{tec}\} \in P_{i,j}$$
$$f_i(\theta_i, \{p_i^{kon}, p_j^{tec}\}_1, ..., \{p_i^{kon}, p_j^{tec}\}_n) \geq f_i(\theta_i, (\{p_i^{kon}, p_j^{tec}\}_1, ..., \{p_i^{kon}, p_j^{tec}\}_{i-1}, \{\mathrm{y}_i^{kon}, \mathrm{y}_j^{tec}\}_i, \quad (6)$$
$$\{p_i^{kon}, p_j^{tec}\}_{i+1}, ..., \{p_i^{kon}, p_j^{tec}\}_n)\}, \ \theta \in \Omega^n(Y).$$

Thus, for the formalized DEPP system provided that the systems works within the time interval when it gradually gets through the states of $t_1, ..., t_n$, the necessary requirement for coordination of the interaction between DAE and TAE (3) is the achievement of Nash equilibrium.

In order to reach the Nash equilibrium according to the reflexive game theory [11] players (in our case they are designers and technologists) must have an adequate concept of the other team's actions. The availability of information resource of PDM/PLM for control system of a product LC Y is the requirement for accomplishment of (6) and, thus of (3).

Let us define the sufficient condition for accomplishment (3), i.e. let us carry out the justification of conditions and mechanisms of system self-regulation. Besides the performance of plan targets from the center K_{xi}, T_{xj} the players (DAE and TAE) begin to cooperate with each other, for this purpose they fulfill some actions $U^{KT}(\mathrm{q}_i^{\mathrm{kon}}, \mathrm{q}_j^{tec}, \mathrm{t})$ и $U^{TK}(\mathrm{q}_i^{\mathrm{kon}}, \mathrm{q}_j^{tec}, \mathrm{t})$ shown in figure 1. As an example of such actions from DEPP common practice we may refer to the stage of technologic coordination of design documentation at the machine-building enterprises. It is evident that to reach equilibrium at (6) the actions $U^{KT}(\mathrm{q}_i^{\mathrm{kon}}, \mathrm{q}_j^{tec}, \mathrm{t})$ and $U^{TK}(\mathrm{q}_i^{\mathrm{kon}}, \mathrm{q}_j^{tec}, \mathrm{t})$ must be task-oriented. The task orientation within the given context is understood as such actions of players $p_i^{kon} = p_i^{kon}(U^K, \mathrm{U}^{KT}), p_j^{tec} = p_j^{tec}(U^T, \mathrm{U}^{TK})$ at (6), so that K_{xi}, T_{xj} will be performed in a balanced manner. In other words DAE and TAE must be interested in outcomes of the opposite team's actions.

The task orientation $U^{KT}(\mathrm{q}_i^{\mathrm{kon}}, \mathrm{q}_j^{tec}, \mathrm{t})$ and $U^{TK}(\mathrm{q}_i^{\mathrm{kon}}, \mathrm{q}_j^{tec}, \mathrm{t})$ may be achieved by means of introduction the mechanism of mutual interest from players in DAE and TAE in achieving the goal from the center, for instance, with the help of incentive mechanism with project time limit t_Σ. In this case the incentive mechanism serves as sufficient condition for accomplishing (3). For illustrative purposes let us consider the simplified scheme of such mechanism implementation based on incentive payment systems [12].

2. Incentive mechanism in coordinated interaction with time limit $t \le t_{\Sigma}$.

The common form of the incentive model of production active elements in DEPP can be given as follows:

$$\begin{cases} V_n = f(K_{изм}(D_n)) \to \max \\ \sum_{j=1}^{m1}\sum_{k=1}^{m2} \delta_{jk}\Delta t_k \le t_{\Sigma} \\ q_i^{kon} \le (q_i^{kon})_{gr}; q_j^{tec} \le (q_j^{tec})_{gr} \end{cases}, \quad (7)$$

where V_n is the value of incentive resource, $K_{изм}$ is the number of document change, D_n is the working paper, n is the number of documents, m1 is the number of employees in design and technological departments, m2 is the number of documents with LC elements, t_{Σ} is the rate time at DEPP, δ_{jk} is the competency matrix element of an employee, Δt_k is the development efforts change of a document thanks to the coordinated decisions based on $U^{KT}(q_i^{kon}, q_i^{tec}, t)$ and $U^{TK}(q_i^{kon}, q_j^{tec}, t)$.

In accordance with document oriented approach to DEPP it is necessary to minimize the number of document changes (design and technological) by means of using the incentive resources V_n, optimizing organizational resource ($U^{KT}(q_i^{kon}, q_j^{tec}, t)$ and $U^{TK}(q_i^{kon}, q_j^{tec}, t)$). As a result an employee of a department may work out more documents (D_n), due to the optimization of a document LC (reduction of number and cutting time). Simply stated, every project document goes through the following stages of LC: D (DI, S, C, S_2), where DI is the development with labour intensity (t_{dev}), S is the statement (t_{stat}), C is the change (t_{ch}), S_2 is the statement (t_{stat}). During the optimization process due to the change reduction the competency matrix suffers the redistribution (table 1).

Table 1. Competency matrix of organizational system with parameter δ_{jk} :δ_{jk} = 1, if an employee works with a document; 0 – the other way.

Employees	Documents (on project stages) with document LC elements							
	D_p^1	D_y^1	$D_и^1$	D_{y2}^1	D_p^2	D_y^2	$D_и^2$	D_{y2}^2
1	1	0	1	0	1	0	1	0
….								
k	0	1	0	1	0	1	0	1

The work with documents is combined for a specific employee if the amount of work is less than counterpart, i.e. $\Lambda = \{O_j \le K_i t_j Z_i\} = 1$, then the work j is fulfilled by I employee; if it is 0, then we add some work to the employee, where K is the number of employees working on the project, Z is the employment of workers, t is the labour intensity of document development.

The incentive system may be the motivation of change quantity reduction in design and technological departments: 100 % of change reduction leads to benefit increase at 10%, i.e. the following dependence is formed: $V = f(K_{ch})$, where V is the amount of benefit, K_{ch} is the number of document changes (figure 2).

The dependence of incentive resource value from reduction of change number can be illustrated by the following example. There are two employees from different departments with competences stated in Table 1. With coordinated interaction according to (3) the first employee reduced the time spent on design document development at 10 days, while the second employee reduced it on technological document development at 20 days, the whole project took 150 days (t_{Σ}).

It is as follows: $\sum_{j=1}^{m1}\sum_{k=1}^{m2} \delta_{jk}\Delta t_k = 1\cdot 10+1+1+1+1\cdot 20+1+1+1 \le t_{\Sigma} = 150.$

The amount of benefit for coordinated interaction may be defined from target function (7) and the graph from the figure 2. As time on development was reduced at 30 days (or 20% from 150 days of total days for project), then according to the graph the amount of benefit payment must be 20% from basic amount *V*.

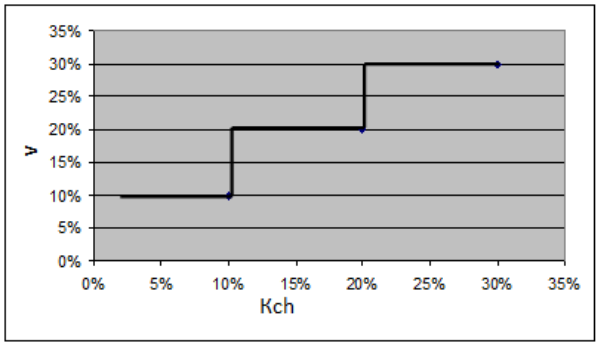


Figure 2. The dependence of benefit amount (V) from reduction of numbers of document changes (K_{ch}).

3. Implementation of the incentive mechanism in the design and technological preparation of production using the information processing level

The incentive mechanism developed above with the coordination of interests in the DEPP can be realized in the context of a single information space of the enterprise. For implementation, it is necessary to develop a structure for the information processing level through the integration of data models of product data management systems (PDM) and event management systems (PEM).

The data model in the PDM system should consist of product design data in the form of a tree with part and assembly units, attribute data and relationships between the elements in the form of project tree, the geometry data of the product in the form of 3D models. Also, the data model should contain a set of design and technological documentation and data on notifications of changes and preliminary notice of a change after working out the “Coordination” unit with the incentive mechanism.

The model of classes in the PEM system can be deployed using Big Data technology to store information about changes in configuration of structures and technologies for all parts and assembly units of the product, data on the events of version changes and task execution, as well as creating cloud services for processing and analysis data.

Big Data technology consists of two components: a connected cluster system and a programming interface. Basically the work of these technologies is connected with three basic principles. Firstly, the data is evenly distributed on the internal disks of many servers, and in the systems of design and technological preparation for production data on the design and technology are on different servers, and when working with an aviation cluster they are on different continents. Secondly, it is not the data which is transferred to the processing program, but the program is transferred to the data. The third principle is that data is processed in parallel, and this possibility is embedded architecturally in the program interface. Thus, instead of the usual "database-server" concept, there is a cluster of many inexpensive nodes, each of which is both a repository and a data processor, and the concept of a “database” is absent. This system has two important characteristics: any complex analysis of a large amount of data is reduced to processing them on the local disks of the server, i.e. the maximum possible reaction time is known in advance; the system is scaled symmetrically and linearly, i.e. the processing time does not depend on their volume [13].

At production enterprises under the concept of CALS, these classes of systems will constitute a single information space of the enterprise and store all information related to information support for all stages of the product life cycle, that is, they will become the only reliable source of up-to-date data.

The accumulated information in the field of design and technological preparation of production at the enterprise is an important asset, but the time spent searching for huge amounts of data does not justify the production process. In aviation and automotive enterprises, the need to introduce large data processing technologies was previously lacking, as the data for designers and technologists were structured, and a relational data model and a database management system existed to work with the structured data flow, but with a huge flow of unstructured data associated with data on operation of equipment in technological processes, in notifications of changes after the detection of defects in prototypes of products obtained from the power of audio and streaming data from a variety of

counters, etc., the old approaches have stopped working. The relevance of this problem has led to the ever growing interest in Big Data class technologies and the emergence of new modules in PDM systems (for example, TEAMCENTER: Big Data).

These modules will allow us to modernize the single information space of the enterprise, taking into account the provision of the required quality level, both the products manufactured and the production processes in general, in the conditions of constant product improvement through notifications of changes, shortening the terms of design and technological preparation of production and through the use of modern methods of operational planning, based on the current information using data protection tools in the form of tables with a access rights differentiation to information.

The purpose of integration of information resources of the enterprise is to ensure effective interaction between designers and technologists, as well as the accumulation of structured and unstructured data, their distribution and quality processing, and the subsequent compilation of statistics based on them.

The general structure of the new class model in the PDM system and in the PEM system is shown in figure 3.

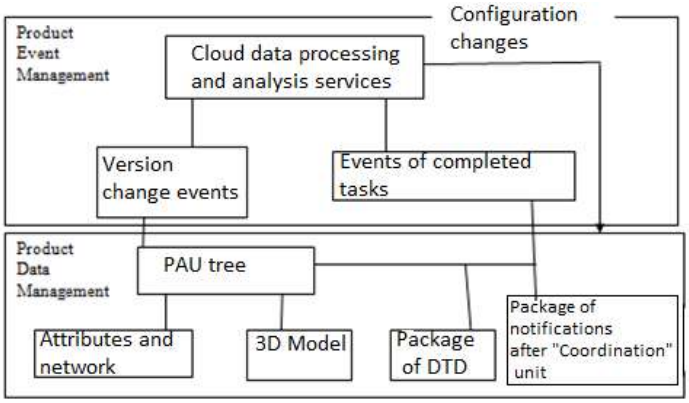


Figure 3. Data presentation level in single information space.

4. Results and discussion

The process of DEPP may be described as a hierarchical active system “center –designers - technologists" in DEPP; however a functioning model of the production of the active element in the system is a model of decision making. Task oriented decision, chosen by each active member as a team, lead to the definite state of the system and may be different from the planned states that are defined by the center. The latter may lead to contradictions in the system in making practical engineering decisions in modeling of materials [14, 15, 16]. Because of that the quantitative assessment of inconsistency in decision making between teams was introduced, and the coordinated interaction task was stated. Based on the Nash equilibrium it was stated that it is possible to achieve balanced solving of contradiction problem between active elements of the system – the designers (DAE) and the technologists (TAE). Due to this fact it is justified that mutual consideration of goals of active elements apart from the goals of the center is the required condition when controlling DEPP in mode of self-regulation.

The achievement of coordinated balance (on actions and goals of subsystems) may be reached with the help of incentive mechanism introduction for active elements. The task of coordinated cooperation in DEPP may be solved by choosing the additional effect function for active elements, connected with benefit payments for reduction changes in documents, which is a variable of their target functions.

5. References

[1] Agafonov A A and Myasnikov V V 2016 *Computer Optics* **40(2)** 275-283
[2] Vasin Yu G and Yasakov Yu V 2016 *Computer Optics* **40(6)** 919-928
[3] Grechnikov F V and Khaimovich A I 2015 *CEUR Workshop Proceedings* **1490** 364-375

[4] Grechnikov F V, Dem'yanenko E G and Popov I P 2015 Russian Journal of Non-Ferrous Metals 56(1) 15-19.
[5] Geraskin M I 2017 CEUR Workshop Proceedings 1903 92-99
[6] Haimovich I N and Frolov M A 2016 *Key engineering materials* **684** 487-507
[7] Morozov V V, Kirichenko A S and Khaimovich I N 2013 *Kuznechno-shtampovochnoe proizvodstvo Obrabotka materialov davleniem* **3** 42-48 (in Russian)
[8] Balashov V G 2002 *Autom. Remote Control*. **12** 92-95
[9] Burkov V N and Burkova I V 2014 *Autom. Remote Control*. **3** 73-86
[10] Burkova I V 2009 *Autom. Remote Control*. **10** 15-21
[11] Novikov D A 2012 *Autom. Remote Control*. **1** 3-23
[12] Novikov D A 2008 *Publishing House of Physical and Mathematical Literature* 139
[13] Leontiev S V, Novikov D A and Petrakov S N 2002 *Autom. Remote Control*. **7** 107-116
[14] Zalozhnev D A, Kalashnikov A O and Novikov D A 2007 *Autom. Remote Control*. **3** 142-153
[15] Ivashchenko A V, Dvoinina O V and Lednev A M 2015 *Intellect. Innovations. Investments.* **2** 58-64
[16] Kirichenko A S and Khaimovich I N 2011 *Bulletin of the Samara State University* **2** 276-281 (in Russian)

Data Modelling to Analyze How the Cities in the Volga Region Correspondent to the Digital State Format

I N Khaimovich[1,2], V M Ramzaev[1] and V G Chumak[1]

[1]Samara University of Public Administration International Market Institute, G.S. Aksakova Street 21, Samara, Russia, 443030
[2]Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. The article suggests the methodology for assessing the readiness of Volga region municipal entities to introduce digital state. The authors worked out a model of statistic tests based on multiple probability theoretic and statistical modelling of parameter values of technological elements of digital economy. This model will allow the Volga region cities to define their possibility to participate in the State Program of Digital Economy, to choose cities most suitable for introduction of modern technologies, to identify the main shortcomings hindering their integration into the program. This research may be of interest to experts in the field of digital economy and Big Data management.

1. Introduction

At present, Russia is carried out a transition to digital development of the state and the digital economy. In the context of fierce competition among municipal entities, the issue of introducing information technologies becomes especially urgent. Municipal entities (ME) - the city districts are key elements of the territorial organization of the country's economy, where new business, financial and cultural centers are formed, which stimulate change. The implementation of modern forms and tools of the digital economy in the ME is currently constrained by the following problems: restricting access to the digital systems in municipalities, the depressed state of the ME economies, the lack of social involvement of local government in the management process, and the lack of bases for building the digital economy.

At the same time, further development of the ME is impossible without attracting new intellectual services, one of which is the technology of the digital state. The following questions arises:

- Are MEs ready to introduce digital state technologies, i.e. to the elimination of intermediaries in services, the implementation of direct transactions in the economy, etc?
- What is the methodology which is used to check the readiness.

When managing the development of the regional economy in the modern digital format, it is necessary to solve the following tasks:

- consider the main ratings of the conformity assessment of the ME to the principles of the digital state;
- develop a methodology for applying, through analysis of compliance data, to the principles of the "digital state";
- analyze the cities of the Volga region using this method.

2. Analysis of existing ratings in the concept of “digital state”

The concept of "digital state" includes the following concepts:

- "smart economy", which includes a high level of indicators in the areas of innovation, employment, trade, productivity, physical infrastructure;
- "smart environment" that links air quality index, water supply, noise level, environmental quality, biodiversity, power economy;
- "smart society and culture", which includes education, health service, security, housing, culture, social involvement.

All these indicators should interact only through information and communication technologies (ICT).To analyze the readiness and attract investment in the ME, the "Smart City" rating is developed by the United Nations [1] and along with that the Russian rating for the assessment of sustainable urban development was developed [2]. Analysis results are given in table 1.

Table 1. Ratings of “digital” state in Russia and abroad.

Ratings	Indicators
Rating of sustainable development of cities in RF	Economy, municipal infrastructure, social sphere, environmental situation
Smart city system of indicators	Economy (physical infrastructure, innovations, employment, trade, productivity, information and communication technologies). Environment (air quality, water supply, noice level, environmental quality, biodiversity, power economy). Society and culture (education, health service, security, housing, culture, social involvement)

Let's consider in detail the rating of sustainable development of Russian cities, proposed by SGM Agency. This rating includes more than thirty indicators that characterize ME: economy, municipal infrastructure, social sphere, ecology. The respondents of the selection are the administrative centers of Russia. The advantage of this rating is the balance of the indicators under consideration, since unbalancing adversely affects the sustainable development of the ME. The disadvantages include the fact that high performance does not always determine the leading position of the ME in the country. Every year the leading cities are the same in this ranking: Moscow, St. Petersburg and Ufa. The administrative centers of the Volga region are annually referred to the outsiders. The shortcomings in the context of the topic under consideration include the isolated character of the indicators of this rating to the topic "digital state". For example, the indicator "demography", consisting of the criteria "natural growth rate", "migration growth rate", "demographic burden" only indirectly affects the use of ICT in all areas of the city.

Next, consider the smart cities metrics system developed by the United Nations Economic Commission for Europe. The degree of readiness to introduce new ICTs is assessed through the city's innovation indicators to improve the living standards of the population. This system assesses the effectiveness of activities and services to meet future generations in various aspects of activities.

This system consists of the following three blocks:

Block 1. Economics: ICT infrastructure, innovation, employment, trade - e-commerce, trade - export / import, productivity, physical infrastructure - water supply, electricity, health service, transport, buildings.

Block 2. Environment: air quality, water supply, noise level, environmental quality, biodiversity, power economy.

Block 3. Society and culture: education, health, safety – consequence management, security - emergencies, security - ICT, housing, culture, social involvement. The last indicator includes calculation of levels of public participation, gender equality of incomes, the ability of people with special needs, attractiveness for qualified personnel, the Gini coefficient.

The advantage of this system is the detailed calculation of all indicators. The disadvantage is that this assessment can be applied only to European countries.

After a detailed analysis of European and Russian ratings and evaluation systems, a methodology was developed for assessing cities to introduce technologies of the "digital state", taking into account all the features of the Volga region [3-6].

3. Involvement model of ME of Volga region in “digital economy” technology

Despite a large number of studies in this field [7-11], there is no unified methodology for assessing the involvement of ME in “digital state” technology. International ratings and systems are only being tested and they answer specific questions: is the economy of the city stable; does it have elements of a "smart" city. They do not assess any indicators of intellectualization of the urban environment.

Further, a model of statistical tests will be considered based on repeated probability-theoretical and statistical modeling of the parameter values of the "digital state" concept. This model will be associated with the analysis of a large amount of data, which will require the automation of evaluation indicators, and it can be based on the Big Data technology [12,13].

The method of using database mining technique in order to support business objectives is as follows:

1. Formation of a big data set in the hadoop from the twitter using the filter “Samara region”, revealing the number of calls;
2. Division of the formed set into various filters associated with the performance measures of the involvement of the ME in the "digital state" technology;
3. Monitoring of the stream analysis of unstructured data sources using filters;
4. Development of a program in Scala Programming Language for working with filtering in the field of Big Data;
5. Debugging and program testing with a set of practical data;
6. Analysis of calculation results.

The social network “Twitter” is used in order to receive data, since it is an “open” product, its application does not require additional investment, and 50% of Internet users have profiles in this program. Twitter is the second most popular network among users worldwide, second only to Facebook. However, unlike Facebook, which does not provide open access to its data, Twitter provides such access; there are no restrictions on access to the server's data sets. Users of this social network exchange mainly textual information, which is an undoubted advantage in processing. Twitter is not an object network and most widely reflects public opinion on many issues of interest, so data processing from this social network was best possible for the formation of small business zones in the region.

To work with BIG DATA in social networks it is necessary to use methods of data collection, processing and analysis. The data is collected in real time, within a certain geolocation, or within the entire network, according to certain patterns. Information of interest for analysis is: location, date and time, content, “author” of the content (user), communication between users. Data collection in social networks can be performed using the following tools: Apache Hadoop, Biglnsights (IBM), Cloudera, Hortonworks, and Storm. Hortonworks was chosen to carry out research on ME involvement in the "digital economy". Twitter Application (apps.twitter.com) was used, in which the following key parameters were defined and refined: API key, API secret, Access token, Access token secret.

To collect data using Hortonworks, Twitter App, the Flume service configuration file was used in the Hortonworks Sandbox virtual machine. After installing the virtual machine Hortonworks_ Sandbox version 2.3 and the Flume service settings, the system is ready to download data from Twitter. To view and download files, go to the HDFS folder where the data is process. The types of the HDFS file structure in the Hortonworks virtual machine when solving the task of ME involvement is shown in figure 1.

The collected data must be structured (i.e. processed) in accordance with the MapReduce paradigm. MapReduce is a framework for performing distributed tasks using a large number of computers that form a cluster.

Using MapReduce helped to structure the data stream from social networks by the criteria: fonts, text size, color, link to user profile, location, time and so on.

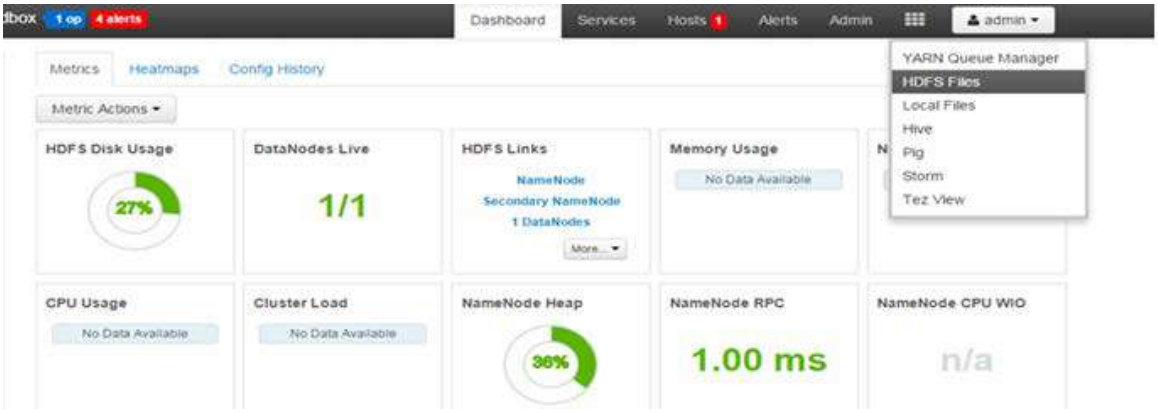


Figure 1. Visualization of HDFS in Hortonworks when downloading files to solve the problem of ME involvement in the "digital economy".

To determine the data for ME analysis, for our study it is necessary to collect the data of the following types: placement, text, language and time. In order to extract only this information, you can use the MapReduce technology built into the Hortonworks Sandbox tool. For data processing we use DBMS Hive in Hadoop environment, which allows performing operations on data and their analysis by SQL-sequel like. To do this, it is necessary to create a file for processing and creating the necessary hivedll.sql. tables.

Run this file using the following command: Hive_f hiveddl.sql. Structured data will be placed in table 2.

Table 2. Type of headings for the analysis of structured data in tasks for the estimation of ME indicators.

A	B	C	D	E	F
Data/Time	Time/Zona	language	Text	location	Sentiments

This table was obtained from social network "twitter" data.

Thanks to the BIG DATA technology, it is possible to store and update data in the file system "hadoop" for the filter "Samara region" (filter1 = {Samara region}). Then it is necessary to filter this area according to the basic parameters for estimating the ME, by setting, for example, the following filters: Filter2 (economy) = {roads, goods}; Filter3 (environment) = {forest, air}; Filter4 (society and culture) = {nightclub, concert, session, hangout}.

It is possible to obtain graphs of the number of users accessing filters (i.e., the value of the ME internetization indicator) from the data collection time.

The time of data collection from the Internet in BIG DATA technology is unlimited.

As a result, we receive a dynamic change in the information in real time from the Internet, which allows us to monitor the stream analysis of unstructured information (the technology of In-Memory Data Processing and Stream) by filters with minimal investment. To implement this method, a program was written in Scala Programming Language:

```
val file = spark.textFile("hdfs://… ")
val errors=file.filter(line=>line.contains("Samara region"))
//count all the data
errors.count()
//count data mentioning Filter
errors.filter(line=>line. contains("concert")).count()
//Fetch the filter as an array of string
errors.filter(line=>line. contains("doctor consultation")).collect()
```

After the work of the program we obtain a dynamic change of parameters in the BIG DATA environment, which allow us to determine some of the indices of ME involvement in the "digital economy" taking into account unstructured information. This method of collecting information for estimating parameters can also be used for other social systems and sites, and also statistics of official sources posted on the Internet can be used to collect information

Thus, a tool is proposed for data collection in the ME system of indicators in the “digital economy”.

The system of indicators of ME involvement in the introduction of “digital state” technologies is shown in figure 3.

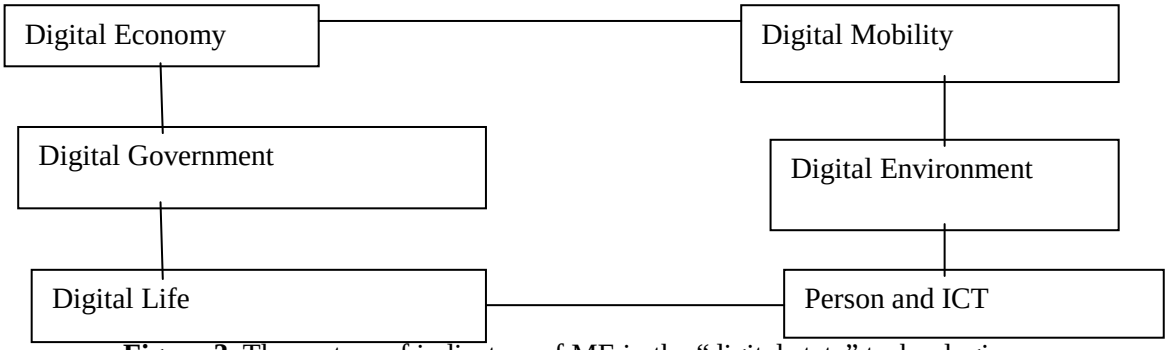


Figure 3. The system of indicators of ME in the “digital state” technologies.

The indicators in the above mentioned indicator system include the following:

- Digital Economy: the indicator for innovation, entrepreneurship, the city's competitiveness, the indicator for producibility, the labor market, the indicator for financial independence;
- Digital Mobility: local transport system, (inter-) national accessibility, ICT infrastructure, transport system stability;
- Person and ICT: an indicator for intelligence, lifelong learning, an ethnic variety;
- Digital Life: cultural and entertainment facilities, health status, individual security, housing quality, educational institutions, tourist attraction, social cohesion;
- Digital Government: political awareness, public and social services, effective and transparent administration;
- Digital Environment: air quality (without pollution); environmental awareness, sustainable resource management.

These indicators have expert values. To compare different indicators, it is necessary to standardize the values from the samples of several cities. The study uses the standardization method of z-transformation using the following formula:

$$z_i = \frac{x_i - \bar{x}}{S},$$

where $\bar{x}$ is the average value in the sample, S is the standard variation in the sample. This method converts all values of the indicators into standardized values with an average of 0 and a standard deviation of 1. The method has the advantages of considering heterogeneity within groups and maintaining its metric information. In addition, high sensitivity to changes is achieved.

In order to obtain results by indicator level, indicators and final result for each city, it is necessary to summarize the values at the level of the indicator. To aggregate the corresponding group of indicators by domain, we also take into account the coverage factor of each indicator. A certain result from the indicator covering all cities weighs more than the indicator covering only, for example, 6 cities. In addition to this slight correction, the results were aggregated at all levels without any weighting. Aggregation was added, but divided by the number of values added. This allows us to include cities that do not cover all indicators. Their results are calculated from the available values. However, it is necessary to ensure good coverage of all cities in order to obtain reasonable results.

Some indicators can be not only expert, but also calculated, they include the indicator of manufacturability of production, the index of innovation, the indicator of Internetization, the index of intellectualization, the indicator of financial independence, the index of energy efficiency, the indicator of the introduction of creative technologies. These groupings can allow us to obtain quick management decisions depending on the average values of the indicators. If the average is in the range of 3.7 and above, then the ME is ready for the introduction of digital economy technologies. If the average value is in the range from 2, 5 to 3.7, then the ME has an average level of readiness. If the

average value is in the range from 1.95 to 2.5, then the ME has a satisfactory level of readiness. If the average is below 1.95, then the ME is not ready to introduce digital economy technologies [14].

4. Results and discussion

According to the methodology developed above, there were carried out some calculations of some key figures and indicators of readiness of cities in the Volga region to implement the technologies of the "digital state".

Table 3 shows the calculations of the relative and absolute values of the ME in the Volga region, taking into account the z - transformation.

Table 3. Indicators of municipal entities of Volga region taking into account the z-transformations.

Digital Economy	ME (abs)	variation	ME (rel)
Indicator for innovation	3	0,547722558	0,912871
Entrepreneurship	3		0,912871
Competitiveness	2		-0,91287
Indicator for producibility	3		0,912871
Labour market	2		-0,91287
International integration (indicator for financial independence)	2		-0,91287
Total	15		
Person and ICT			
Indicator for intelligence	1	1,707825128	-1,0247
Lifelong learning	2		-0,43916
Ethnic variety	3		0,146385
Openness	5		1,317465
Total	11		
Digital Mobility			
Local transport system	2	2,217355783	-0,56373
(Inter-) national accessibility	1		-1,01472
ICT- infrastructure	4		0,338241
Transport system stability	6		1,240216
Total	13		
Digital Life			
Cultural and entertainment establishments	6	0,975900073	1,610235
Health conditions	5		0,58554
Personal safety	3		-1,46385
Energy index	4		-0,43916
Educational institutions	4		-0,43916
Tourist attractiveness	5		0,58554
Social cohesion	4		-0,43916
Total	31		
Digital Government			
Political awareness	3	0,577350269	-0,57735
Public and social services	3		-0,57735
Effective and transparent administration	4		1,154701
Total	10		
Digital Environment			
Air quality (without pollution)	4	1,154700538	0,57735
Ecological awareness	4		
Sustainable resource management	2		
Total	10		

Further the figures 4-9 show the histograms of digital city indicators for the cities of Samara and Ulyanovsk.

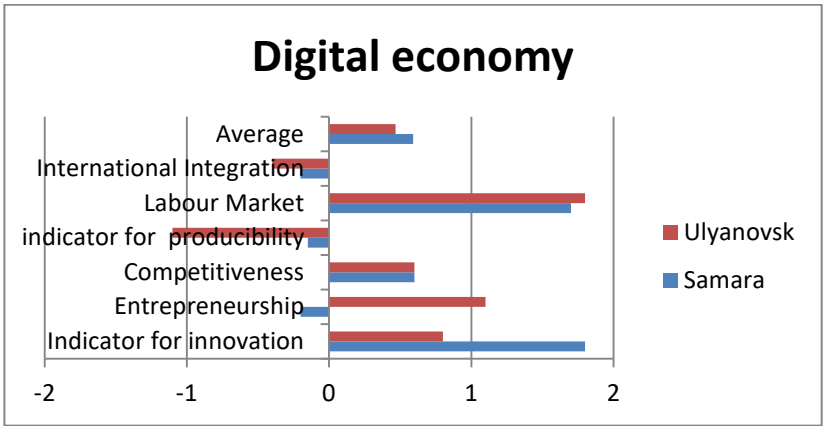


Figure 4. Indicators of "Digital Economy" for Samara and Ulyanovsk cities.

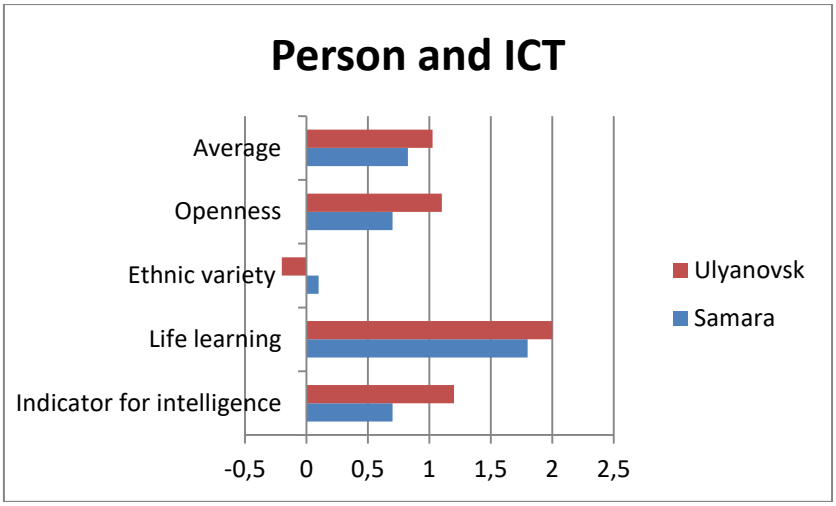


Figure 5. Indicators of "Person and ICT" for Samara and Ulyanovsk cities.

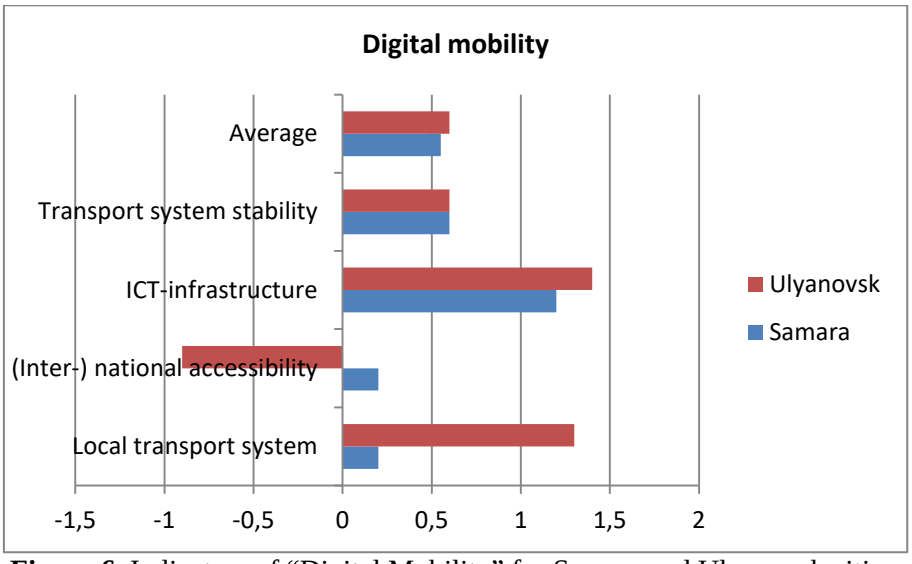


Figure 6. Indicators of "Digital Mobility" for Samara and Ulyanovsk cities.

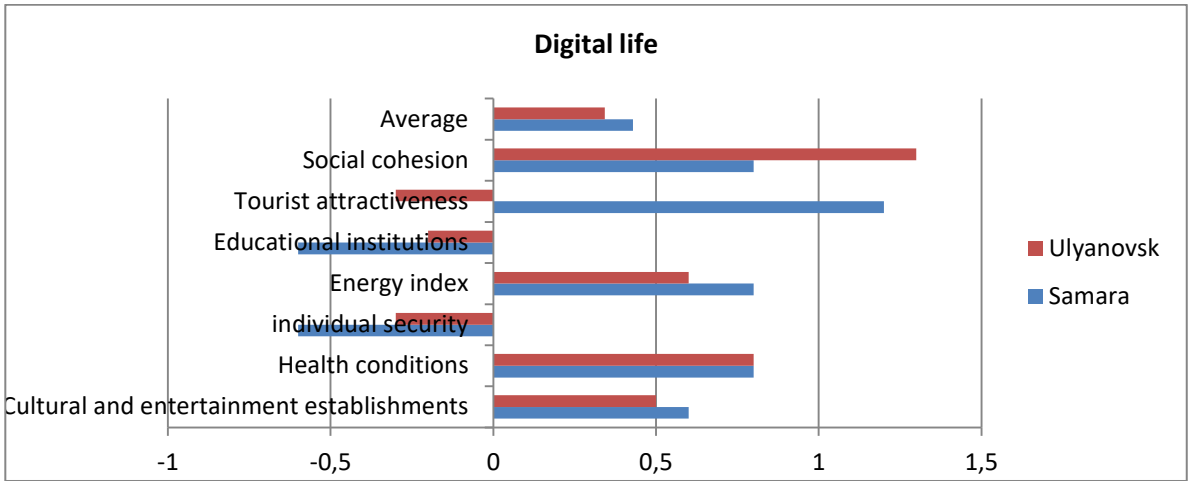


Figure 7. Indicators of “Digital Life” for Samara and Ulyanovsk cities.

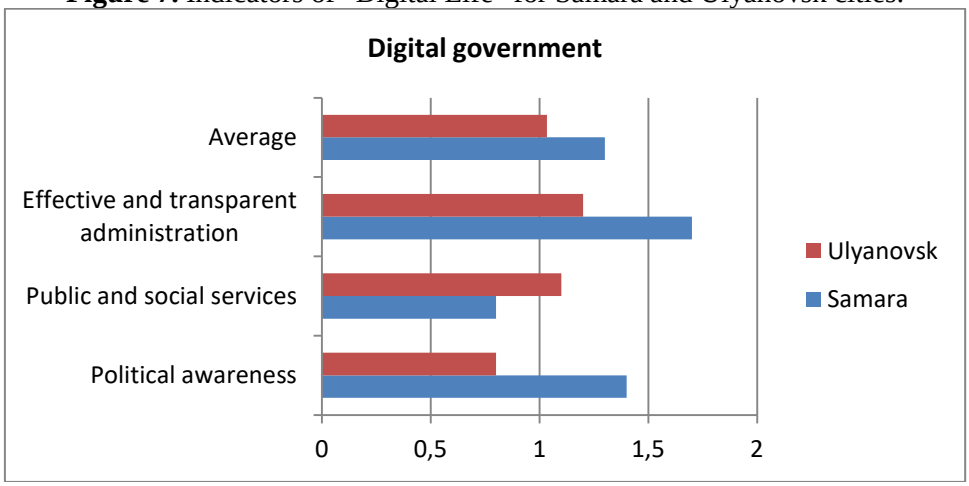


Figure 8. Indicators of “Digital Government” for Samara and Ulyanovsk cities.

Then applying a similar method it is possible to obtain level of indicators for the cities of Ulyanovsk and Samara (figure 10). From the histogram data, it can be seen that Samara is ahead of Ulyanovsk in the indicators of “digital production”, “digital economy”, “digital life”, but the indicators of "people and ICT", "digital environment", "digital life" are better for the Ulyanovsk municipal entity. In general, the willingness to introduce ICT in both cities is the same.

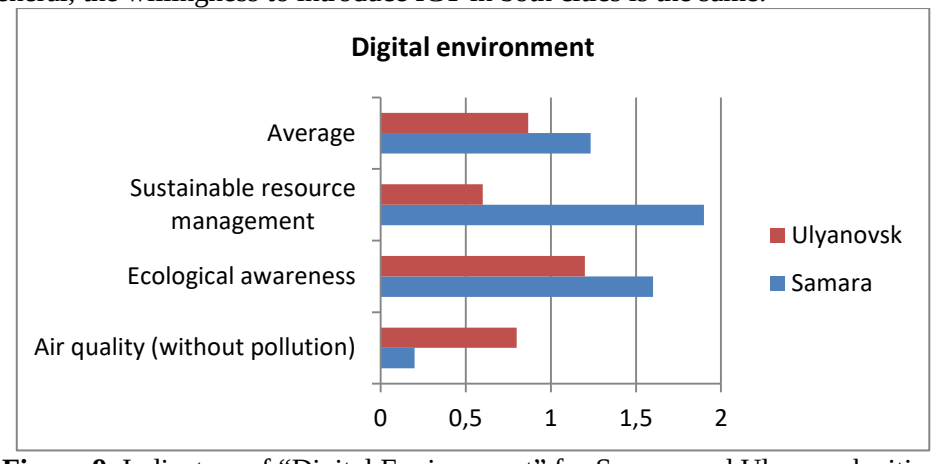


Figure 9. Indicators of “Digital Environment” for Samara and Ulyanovsk cities.

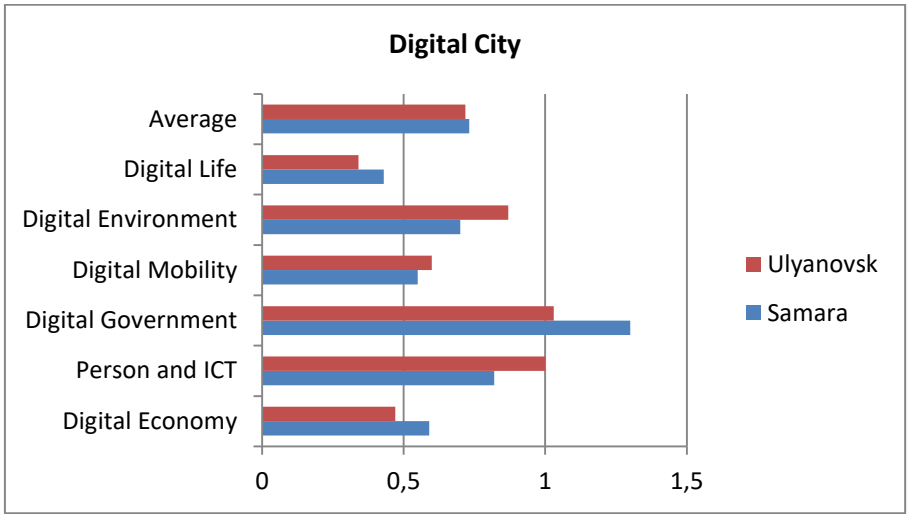


Figure 10. Indicators of “Digital City” for Samara and Ulyanovsk cities.

Let us carry out a comparative analysis of indicators of the "digital city" for municipalities with a lower level of readiness, i.e. located below the zero level, i.e. requiring significant investment in the introduction of ICT. A comparative graph of the study results is shown in figure 11.

As a result, it is possible to identify the main trends for investing in the ME of Volga region on the basis of a comparative analysis of the "digital city" indicators. Figure 11 shows that investing in a region with indicators below the zero level is not profitable. Many cities of the Volga region belong to this zone, for example, Zhigulevsk. It is better to invest in the ME with a level of readiness above the zero level, for example, Ulyanovsk. These cities correspond more to the concept of the "digital city", they are almost ready to introduce the technologies of the "digital state".

Thus, the assessment model will allow to determine the level of development of municipal entities, which are ready to implement the digital state, to identify shortcomings in the group "which is not ready for implementation", will improve the performance of the ME on the basis of a detailed analysis of data of all major cities of the Volga region.

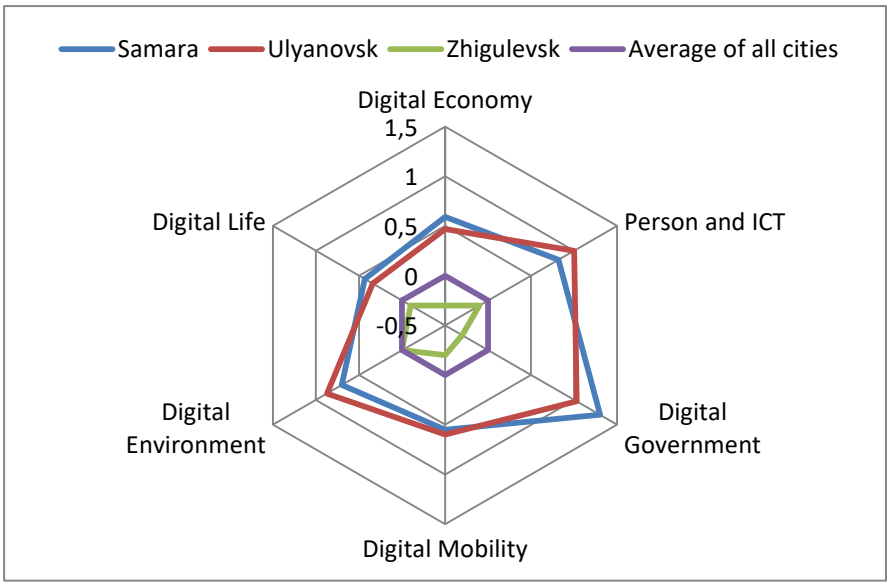


Figure 11. Graph of indicator level of Digital City for Volga region cities.

5. References

[1] Access mode: http://www.unece.org/fileadmin/DAM/hlm/documents/2015/ECE_HBP_2015_4.ru.pdf (9. 11.2017).

[2] Access mode: http://agencysgm.com/projects/Рейтинг%20устойчивого%20развития-2015.pdf (9.11. 2017)
[3] Terekhin E A 2017 *Computer Optics* **41(5)** 719-725 DOI: 10.18287/2412-6179-2017-41-5-719-725
[4] Afanasyev A A and Zamyatin A V 2017 *Computer Optics* **41(3)** 431-440 DOI: 10.18287/2412-6179-2017-41-3-431-440
[5] Vorobiova N S,Sergeyev V V and Chernov A V 2016 *Computer Optics* **40(6)** 929-938 DOI: 10.18287/2412-6179-2016-40-6-929-938
[6] Boori M S, Kuznetsov A V, Choudhary K K and Kupriyanov A V 2015 *Computer Optics* **39(5)** 818-822 DOI: 10.18287/0134-2452-2015-39-5-818-822
[7] Akaslan D and Taskln S 2016 *4th Int. Istambul Smart Grid Congress and Fair* (New York: IEEE Press)
[8] De Domenico M, Lima A A and Gonzalez M C 2015 *EPJ Data Science* **1** 1-11
[9] Glebova I S, Yasnitskaya Y S and Maklakova N V 2014 *Mediterranean J. of Social Sciences* **12** 129-133
[10] Ishkineeva G, Ishkineeva F and Akhmetova S 2015 *Asian Social Science* **5** 70-73
[11] Khatoun R and Zeadally S 2016 *Communications of the ACM* **8** 46-57
[12] Khaimovich I N, Ramzaev V M and Chumak V G 2016 *CEUR Workshop Proceedings* **1638** 864-872
[13] Khaimovich I N, Ramzaev V M and Chumak V G 2015 *CEUR Workshop Proceedings* **1490** 327-337
[14] Komarevtseva O O 2017 *R-Economy* **3** 32-39

Block algorithm for the joint difference solution of the d'Alembert and Maxwell's equations

L V Yablokova[1] **and D L Golovashkin**[1,2]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086
[2] Image Processing Systems Institute - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001

Abstract. A characteristic feature of mathematical modeling at the present stage of development is the consideration of the architecture of the computer system, not only for stage of compiling a computer program, but also during the development of a numerical method and synthesis of the mathematical model. This method significantly broadens the researcher's ability to search for the optimal mapping of the numerical method to the mentioned architecture, in the sense of accelerating computations. In this paper, this idea is illustrating by examples of the basic mathematical model of computational electrodynamics and optics, Maxwell's equations, and the FDTD. This modification allows to reducing the data exchange rate between the operational and cache memory due to the greater number of arithmetic operations per one grid function in solving the d'Alembert equation. On the other hand, freely use the technologies FDTD method and ready-made software implementations for setting the incident wave, imposing the absorbing layers, taking into account the dispersion of the medium.

1. Introduction

Despite the old beginning of researches [1] in the field of the numerical solution of the equations of Maxwell, interest in this subject domain only increases over time. This remark first of all belongs to application of a method of final differences [2] – to version of the numerical decision most popular so far the specified equations. In the seventies the last century [3] he has received own abbreviation of FDTD (Finite-Difference Time-Domain) under which it is widely known to this day.

The relevance of development of a FDTD method is due to several reasons. First, broad demand in new subject domains: to nanophotonics [4], radiobiology [5], etc. owing to community of mathematical model. Indeed, by means of Maxwell's equations are described any processes connected with the wave nature of electromagnetic radiation without restrictions. Secondly, the necessity of taking into account the features of modern computing architectures: cluster [6], vector [7], etc. If previously the performance of computers was increased mainly due to the increase in the clock speed of the central processor, but now it is increasing using various methods of parallel processing of data.

In the study proposed by the authors of this work, the emphasis in synthesizing a new variant of the FDTD method is made on the basis of the hierarchical structure of computer storage devices, in particular, on the possibility of optimizing communications between the operational and cache memory of the processor. Specified subject sparingly illuminated in the scientific press in force the interdisciplinary nature of the task. Experts in calculus mathematics traditionally are interested in other

problems of the theory of differential schemes: increase in an order of approximation of a differential task [10], designing of mobile net areas [11]. In turn, developers of mathematical software that implements the FDTD method, seeking to keep pace with the development of "large" modern forms of parallel and distributed computing: using processors [12] the clouds [13].

Nevertheless, fresh works appear on the block algorithms for the difference solution of the basic equalizations of electrodynamics [14-16]. This testifies to the attention to "small" forms of organization of calculations, taking into account the inconspicuous architectural features of the processors and demonstrates significant acceleration at the expense of such an account. In computational practice, it has long been possible, with the example of matrix calculations, to manage the duration of calculations by changing the block sizes [17, 18] in block algorithms. Unfortunately, the features of the theory of difference schemes (the absence of the need for multiplication of dense matrices, the traditional emphasis on stability problems) until recently hampering the penetration of blockiness into this subject area.

Another obstacle to the development of block methods is the complexity of controlling the loading/unloading of data into the processor's memory cache during the computing process. Load (but not unload) statements are present only in Assembly language and their use is not prescriptive: data may not be loaded into the cache after they are executed if the block is already there.In more common programming languages, the management of communications between the operational and cache memory is achieved indirectly by receiving a partitioning of the cyclic constructs referred to as tiling [19, 20].

The application of this technique for the joint difference solution of the equations of d’Alembert and Maxwell and is devoted to the proposed publication.

2. Peculiarities of the joint difference solution of the d'Alembert and Maxwell’s equations

The theory of the joint difference solution of the d'Alembert and Maxwell’s equations is described in detail in [21, 22], the results of his experimental research, significant for the chosen subject. The following software and hardware tools were used during the experiments: the Intel Core i7-3770 processor, the Ubuntu 16.04.1 operating system (the 4.4 kernel), the gcc 5.3 compilerand Meep 1.3 (compiled for comparison by the same compiler) is a free and open-source software package for simulating electromagnetic systems via the finite-difference time-domain (FDTD) method. Actually being a reference for a wide range of researchers [23]. The grid area was chosen with the dimensions of 10000 × 10000 nodes in space and 200 nodes in time, which provided sufficient memory load (as for real computational experiments in nanophotonics) for an acceptable duration of calculations.

In the case of simulating the propagation of TM-mode (in terms of the fundamental work [24]) differential solution of the equations of d'alembert and Maxwell differ from each other on the magnitude of the machine precision and converge to the analytical. When using the MEEP package, the duration of the calculations made 124.75 sec., calculations on the author's software implementation of the difference solution of Maxwell's equations lasted 112.71 sec., for the difference solution of the d’Alembert equation lasted 41.34 sec. Acceleration in the case of the difference solution of the d’Alembertequation in 41.34 sec. and 2.73 times, respectively, can not be explained by a decrease in computational complexity by 10% in the difference solution of the d’Alembert equations compared to that for Maxwell. The authors associate the observed effect with an increase in the ratio of the number of arithmetic operations that fall on the calculation of the differential pattern, to the amount of memory involved in working with the same pattern. For a two-dimensional Yee scheme, the specified value is 3/2 (three operations are performed on two different field projections); for the first and second difference equations and 7/4 (7 operations on 3 field projections and the value of dielectric permeability) for the third; for the second scheme – 9/3 (9 operations on two different time layers of the same projection and the value of dielectric permeability), which apparently leads to a significant decrease in the intensity of communications between the operational and cache memory.

The joint difference solution combines the advantages of both approaches: acceleration for the dalamber equation and the developed Toolkit (overlays of absorbing layers, the setting of the incident wave, etc.) for the Maxwell’s equations. The Maxwell’s equations were experimentally solved in PML layers [24] with a thickness of 100 knots along the edges of the grid area, and the dalamber equation

was solved in the center. The duration of the transactions in this case amounted to 47.47 sec. and slightly superior to the previous result (in which the calculations in PML layers was not carried out), obtained by accelerations of 2.63 times compared to the Meep package and 2.37 times compared to the author's implementation of the Yee scheme.

Further, the authors use the observed effect of reducing the duration of calculations with a decrease in the intensity of communication between the operational and cache memory for further acceleration of calculations in the case of a joint solution.

3. Block algorithm for the joint difference solution

To date, block algorithms for the difference solution of Maxwell's equations [14] (2009) and for d'Alembert [16] (2015) are known. The method of transition to cloud computing with diamond toroidal block shape (Diamond Torre Algorithm) proposed in the last work served as the basis for the synthesis of the author's block algorithm of joint difference solution. The transition to cloud computing consists of the following two steps.

The first one performs calculations on the*h* time layers of the grid area to update the values of the grid functions in the sub-area of the PML location. In figure 1 it can be said "bucket handle". Due to the information dependence on the iteration space coinciding with the grid area, the values in the nodes adjacent to the distance *h* and smaller to the PML, which are already related to the wave equation, are also to be calculated at this stage.So in the node separated from the absorbing layers on *k* other nodes ($k \leq h$) the value of the grid function on the *k*time layer will be formed.All such units will make up the "bucket wall" in figure 1.

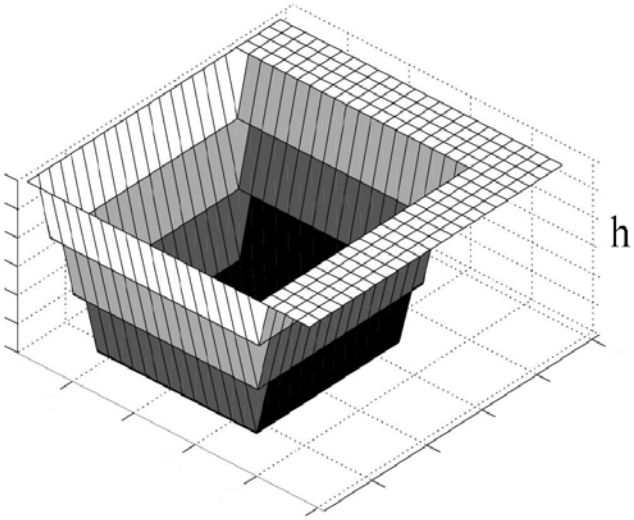


Figure 1. Distribution of grid functions on time layers before block stage of calculations. From white (the layer at the maximum height and near) to black (the layers at a lower altitude and around it).

The second stage is characterized by the organization of calculations in the internal volume of the "bucket" according to the algorithm, which the authors called the wave by analogy with the method of transition to the block of [19]. As shown in figure 2, during the computing process inside the bucket, the values of the grid functions are calculated from left to right in such a way that the front of the process has the form of an inclined plane.In the nodes of the grid area before it, the function values are not yet defined, in the nodes of the grid area after it are found on the layer *h*. On the very surface of the front, values are calculated on different layers in ascending order with increasing height and decreasing node abscissa. After the second stage of the algorithm, the first one comes again, and their alternation continues until the values of the grid functions on all layers of the region are found.

In contrast to the diamond toroidal algorithm from [16], such an approach is easier to implement and leads to a slight but stable reduction in the duration of calculations compared to the other method of transition to blockage.

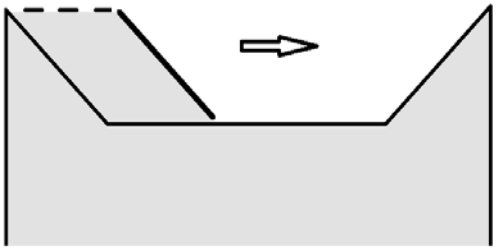


Figure 2. The distribution grid functions at the temporary segments on a block calculation step.

Table 1. The dependence of the duration of computations (sec.) the height of the block *h*. T_1 is the duration for the wave algorithm, T_2 is the duration for the diamond toroidal algorithm. The value *V* is the volume of the block in megabytes.

h	2	4	8	10	20	40	50	100
T_1	42.63	36.88	33.88	33.49	32.54	36.83	40.80	48.42
T_2	42.75	37.15	34.13	33.57	33.64	37.16	40.90	48.44
V	0.45	0.9	1.79	2.24	4.49	8.97	11.22	22.43

It is noteworthy that the best results for both algorithms are achieved for *h*=20the maximum wave height (or torus) at which the block still fits into the cache memory as a whole (the volume of L3 cache for Intel Core i7-3770 is 8 MB). Indeed, at*h*>20 the block is no longer placed in the cache and you have to load it in parts, and for *h*<20 the fast cache memory is not used entirely. Both of these circumstances lead to an increase in communications and, as a consequence, the total duration of calculations.

4. The study of the silicon Fresnel's lens

Of particular interest is the calculation and creation of elements to control the radiation of the Novosibirsk free electron laser (NovoFEL) [25], which is the most powerful source of terahertz radiation. By now, focusing elements in the form of binary diffraction lenses made of silicon [26, 27] are known for it, which are characterized by the known advantages and disadvantages of binary optics. Diffraction elements favorably with refractive [28] much smaller thickness, high radiation resistance, greater manufacturability in use and manufacture. However, the focusing element with the binary relief mentioned in [26, 27] is characterized by low efficiency (often not exceeding 40% [28]), in which most of the energy of the illuminating beam does not fall into the focus area.

However, diffraction elements with a continuous profile, free from this drawback, have been known for a long time. Moreover, the calculation of the binary lens itself is traditionally [28] preceded by the calculation of the Fresnel's lens, which is such an element. The problem of forming a continuous relief on a silicon plate has not yet been solved technologically, the processes of mechanical and chemical processing of silicon are considered to be sufficiently coarse to obtain an uninterrupted profile with the necessary accuracy. However, advances in plasma etching of the diamond surface (not inferior to silicon in hardness and inertness to various kinds of effects) [29] allow us to hope for the rapid improvement of the technology of reactive-ion etching of silicon demonstrated so far in the manufacture of binary relief [30].

The model, numerical method and software complex presented in [22, 31, 32] and based on the joint difference solution of Maxwell's equations with the use of block algorithm of organization of calculations are chosen as a tool for calculating diffraction on various silicon lenses. Based on [26,27,30], we assume the refractive index of silicon n=3.42 for the wavelength λ=141 μm, to which we further give all distances. Thus, the aperture of the refractive lens put equal 101.74 λ, the thickness and the radius of curvature of 10.33λ and 130.41λ, respectively. Then, according to the geometric optics [33], the focus area will be located at a distance f=50.87λ from the right pole. Calculation of Fresnel's lens traditionally assumed the height of its relief equal to *h*=λ/(n-1), which is unacceptable in this case. For a lens with such a high numerical aperture, half of its Fresnel's zones will be less than the width of the wavelength, which is why the work of the element is strongly unpredictable. Taking

h=2.07λ, Fresnel's own lens was calculated (figure3), consisting of 5 zones, the extreme of which is characterized by an acceptable width of 5.18 λ.

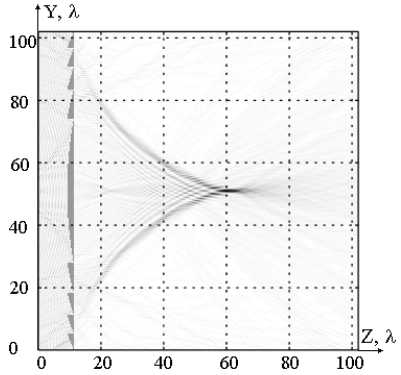


Figure 3. The intensity distribution in the computational region for a Fresnel's lens without technological errors in the manufacture of.

Figure 3 shows the lens focusing of the radiation incident from left to right. On figure 4, there is a noticeable drop in the intensity I (normalized to the intensity of the incident beam) for the lens under discussion as compared to the refractive one. The intensity after the first one is δ=0.9 from the intensity value after the second one in the coordinates of the maxima. The latter are shifted by 4.73 λ and δ=2.47 λ to the left for refractive and diffraction lenses, respectively.

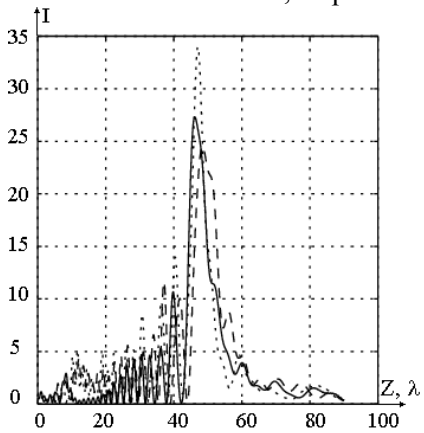


Figure 4. Intensity distributions on the main optical axis behind the right pole for refractive (continuous curve) lens, Fresnel's lens without manufacturing process error (dotted line) and Fresnel's lens manufacturing process error of 0.15 λ (points).

Of particular interest is the modeling of elements with manufacturing process errors [34], especially since such studies have not yet been conducted for silicon elements. Let's assume that the height error is associated with overheating of the plate surface (figure5), leading to an uneven increase in the height of the profile. On figure 5, the value of the discussed value is taken to be equal to λ for better clarity, then consider the more real smaller values (table 2). Indeed, the maximum height error for the spent process is unlikely to exceed 10%.

Table 2. Characteristics of the studied lenses with manufacturing errors.

Characteristics	Profile height error					
	0.05λ	0.1λ	0.15λ	0.2λ	0.25λ	0.3λ
I_{max}	24.86	24.74	34.03	28.98	23.93	21.07
f	47.54	47.15	47.05	46.82	46.96	46.42
δ	0.91	0.91	1.25	1.06	0.88	0.77

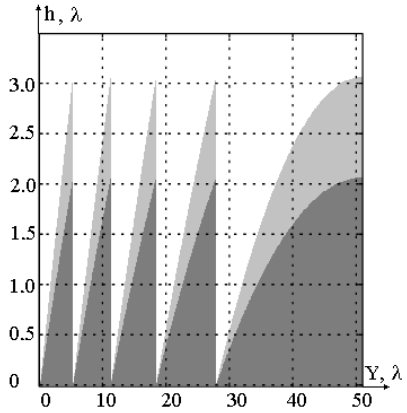


Figure 5. Estimated profile of a lens of Fresnel (dark) and a profile with a technological error of manufacture (dark and light) in case of the maximum error on height in λ.

The results of modeling the Fresnel's lenses with manufacturing process errors presented in table 2 were somewhat unexpected. The natural assumption about the negative influence of errors on the profile height on the focusing efficiency was justified only for sufficiently large error values exceeding a quarter of the wavelength. The deviation of 0.15λ (figure 4, table 2) led to an increase in intensity in the focus area by a quarter compared to the refractive lens. This may be due to the high Fresnel reflection coefficient of silicon, and, consequently, to the significant fluctuations of the intensity of the transmitted wave with a slight change in the height of the profile. Curiously, for a plane-parallel plate, the maximum intensity also fell by a width of 0.15λ.

5. References

[1] Cron G 1944 Equivalent circuit of the field equations of Maxwell *Proc. IRE* **32** 289-299
[2] Yee K S 1966 Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media *IEEE Trans. Antennas Propag* **14** 302-307
[3] Taflove A and Brodwin M 1975 Numerical solution of steady-state electromagnetic scattering problems using the time-dependent Maxwell's equations *IEEE Transactions of Microwave Theory and Techniques* **2(8)** 623-630
[4] Gavrilov A V, Golovashkin D L, Doskolovich L L, Dyachenko P N, Khonina S N, Kotlyar V V, Kovalev A A, Nalimov A G, Nesterenko D V, Pavelyev V S, Shuyupova Y O, Skidanov R V and Soifer V A 2014 *Diffracrive Nanophotonics* (Boca Raton: CRC Press) p 679
[5] Perov S Yu and Bogacheva E V 2014 Theoretical and experimental dosimetry in the evaluation of the biological action of electromagnetic fields of wearable radio stations. Flat phantoms *Radiation biology: Radioecology* **54(1)** 57-61
[6] Guiffaut C and Mahdjoubi K 2001 A parallel FDTD algorithm using the MPI library *IEEE Antennas and Propagation Magazine* **43(2)** 94-103
[7] Vorotnikova D G and Golovashkin D L 2014 CUBLAS-aided Long Vector Algorithms *Journal of Mathematical Modelling and Algorithms in Operations Research* **13(4)** 425-431
[8] Malysheva S A and Golovashkin D L 2016 Implementation of the FDTD algorithm on GPU using a pyramid method *Computer Optics* **40(2)** 179-187 DOI: 10.18287/2412-6179-2016-40-2-179-187
[9] Vorotnikova D G and Golovashkin D L 2017 Difference solutions of the wave equation on GPU with reuse of pairwise sums of the differential template *Computer Optics* **41(1)** 134-138 DOI: 10.18287/2412-6179-2017-41-1-134-138
[10] Xiao F 2007 High-order accurate split-step FDTD method for solution of Maxwell's equations *Electronics Letters* **43(2)** 72
[11] Fidel B, Heyman E, Kastner R and Zioklowski R W 1997 Hybrid ray-FDTD moving window approach to pulsepropagation *Journal of Computational Physics* **138(2)** 480-500
[12] *FDTD solver* (Access mode: http://www.acceleware.com/fdtd-solvers)
[13] *EMA3D Version 3.3 parallel modification* (Access mode: http://www.ema3d.com)

[14] Orozco D and Guang G 2009 Mapping the FDTD Application to Many-Core Chip Architectures *Parallel Processing* 309-316
[15] Grosser T and Cohen A 2013 Split Tiling for GPUs: Automatic Parallelization Using Trapezoidal Tiles *Proc. of the GPGPU* **6** 24-31
[16] Perepelkina A Yu and Levchenko V D 2015 Diamond Torre Algorithm for High-Performance Wave Modeling *Keldysh Institute Preprints* **18** 20
[17] Golub G H 1996 *Matrix Computations* (Baltimor: Johns Hopkins University Press) p 726
[18] Demmel J W 1997 *Applied Numerical Linear Algebra* (Baltimor: Johns Hopkins University Press) p 435
[19] Wolfe M 1986 Loops skewing: The wavefront method revisited *International Journal of Parallel Programming* **15(4)** 279-293
[20] Wolfe M 1989 More Iteration Space Tiling *Proceedings of the ACM* 655-664
[21] Golovashkin D L and Yablokova L V 2012 Joint finite-difference solution of the d'Alembert and Maxwell's equations. One-dimensional case *Computer Optics* **36(4)** 527-533
[22] Golovashkin D L, Buldygin E Y and Yablokova L V 2014 Joint finite – difference solution of the d'Alembert and Maxwell's equations. Two–dimensional case *Computer Optics* **38**1 20-27
[23] Oskooi A F, Roundyb D and Ibanescua M 2010 Meep: A flexible free-software package for electromagnetic simulations by the FDTD method *Computer Physics Communications* **181** 687-702
[24] Taflove A and Hagness S 2005 *Computational Electrodynamics: The Finite-Difference Time-Domain Method* (Boston: Arthech House Publishers) p 1006
[25] Knyazev B A, Kulipanov G N and Vinokurov N A 2010 Novosibirsk terahertz free electron laser: instrumentation development and experimental achievements *Measurement Science and Technology* **21** 13
[26] Agafonov A N, Volodkin B O, Kaveev A K, Knyazev B A, Kropotov G I, Pavelyev V S, Soyfer V A, Tukmakov K N, Tsygankova E V and Choporova Yu Yu 2013 *Optoelectronics, Instrumentation and Data Processing* **49(2)** 98-105
[27] Agafonov A N, Vlasenko M G, Volodkin B O, Gerasimov V V, Kaveev A K, Knyazev B A, Kropotov G I, Pavelyev V S, Palchikova I G, Soyfer V A, Stupak M F, Tukmakov K N, Tsygankova E V and Choporova Yu Yu 2013 Diffraction lenses for powerful bundles of terahertz radiation *News of RAS Series physical* **77(9)** 1360-1362
[28] Golovashkin D L, Doskolovich L L, Kazansky N L, Kotlyar V V, Pavelyev V S, Skidanov R V, Soyfer V A and Khonina S N 2007 *Diffraction computer optics* (Moscow: Fizmatlit) p 736
[29] Karlsson M and Nikolajeff F 2003 Diamond micro-optics: microlenses and antireflection structured surfaces for the infrared spectral region *Optics Express* **11(5)** 502-507
[30] Agafonov A N, Volodkin B O, Kaveev A K, Knyazev B A, Kropotov G I, Pavelyev V S, Tukmakov K N and Choporova Yu Yu 2014 Control of a terahertz laser radiation of cross and modal composition by means of elements of binary silicon optics *Computer Optics* **38(4)** 763-769
[31] Yablokova L V and Golovashkin D L 2018 Block algorithms for the joint difference solution of the D'Alembert and Maxwell equations *Computer Optics* **42(2)** 320-327 DOI: 10.18287/2412-6179-2018-42-2-320-327
[32] Certificate of state registration of the computer program № 2017613903 *Joint difference solution of the equations of D'Alembert and Maxwell*
[33] Butikov E I 2003*Optics* (St. Petersburg: Nevsky Dialect) p 480
[34] Golovashkikn D L,Volkov A V, Eropolov V A, Kazanskiy N L, Karpeev S V, Moiseev O Y, Pavelyev V S, Artushenko V G and Kashin V V 2007 Studying fabrication errors or the diffraction grating on the end face of a silver-halide fiber *Optical Memory & Neural Networks Information Optics* **16(4)** 263-268

Acknowledgments
The research leading to these results has received funding from the Russian Science Foundation grant №16-47-630560-r_a.

An approach to analysis of the similarity of DNA-sequences

B F Melnikov[1], E A Melnikova[1], S V Pivneva[1] and M A Trenina[2]

[1]Russian State Social University, Wilhelm Pieck str. 4, Moscow, Russia, 129226
[2]Togliatti State University, Belorusskaya str. 14, Togliatti, Russia, 445020

Abstract. This paper describes algorithms, corresponding computer programs and the results of computations, supplementing results published earlier. We consider the multiple sequence alignment problem, which can be nominated by a central problem in computational biology. For it, we continue to consider some different versions of so-called "triangular norm" defined on the set of triangles formed by the different distance between genomes computed by different algorithms.
Besides, one of the problems considered in biocybernetics is the problem of reconstructing the distance matrix between DNA sequences, when not all the elements of the matrix under consideration are known at the input of the algorithm. In this connection, the problem arises that the developed method of comparative evaluation of algorithms for calculating the distances between sequences should be used for another problem, i.e., for reconstructing the matrix of distances between DNA sequences.
In this paper, we consider the possibility of applying the method of comparative evaluation of the algorithms for calculating the distances between a pair of DNA strings that we developed and studied earlier for the reconstruction of a partially filled distance matrix. The restoration of the matrix occurs as a result of several computational passes. Estimates of unknown matrix elements are averaged in a special way using so-called risk functions, and the result of this averaging is considered as the received value of the unknown element.

1. Introduction and motivation

In this paper, we describe the continuation of research on the issues that we started in [1, 2, 3, 4, 5]. We describe algorithms, corresponding computer programs and the results of computations, supplementing results published earlier. We consider the multiple sequence alignment problem, which can be nominated by a central problem in computational biology. For it, we continue to consider some different versions of so-called "triangular norm". (The name "triangular metric", also sometimes encountered in our previous papers, is also quite possible and does not contain errors, we will not give detailed comments on this thing. However, the "norm" in our case is some more correct, both when we speak about the whole matrix, and when we speak only about the only triangle.) Such norm defined on the set of triangles formed by the different distance between genomes computed by different algorithms. Thus, in this paper the word "metric" will be used only as the distance between the genomes, and the word "norm" as an indicator of the badness of a certain set of such distances.

In previous cited papers, we described Panin's metric and some algorithms for its improvement. In this paper, we consider some variants of investigation of various variants of the triangular norm. Let us remark, that both these directions are interrelated. Namely,

we improve the interpretation of the Panin's metric in the same way as it is done in some interpretations of genetic algorithms: we are trying to achieve a combination of parameters, for which the triangular norm reaches a minimum value (or is close to it).

Considering this second problem, i.e., the study of variants of the triangular norm, we proceed as follows. We consider incorrect variants of obtaining the triangle inequality, which for matrices of the order of about 50×50 is violated in the two most successful metrics (including the Panin's metric earlier developed by us) in less than 1% of cases. It is important to note that in order to improve the decrease in the quantitative index of the badness of the entire matrix of considerations, we also, first of all, consider triangles in which the triangle inequality is not violated, but in which the badness value is relatively large. Possible improvements are related to the use of neural networks that were not used by us in previous calculations. In this case, neural networks solve the inverse problem: we improve (reduce) the overall badness of the matrix of distances between genomes, forcibly changing the previously obtained distances; further, we try to reflect these forced changes in the original algorithms for calculating distances.

Thus, to determine the distance between genomes, we need *heuristic* algorithms; and, if possible, they should not require too much time. There are various such algorithms, but their obvious disadvantage is getting a few *differing results* when using different heuristic algorithms applied to the distance calculation between *the same pair of DNA strings*. Therefore, the problem of *quality evaluation* of the metrics used (distances) arises; and based on the results obtained in solving this problem, one can draw conclusions about the applicability of a particular algorithm for calculating distances to various applied studies. A possible approach to determining the quality of metrics was given in [4]; also partially we consider these approaches below in this paper.

However, we used the same approach for a completely different problem; it is as follows. Even heuristic algorithms require often large time-consuming costs: for example, to construct a matrix of the order of 50×50 into which the distances computed by the Needleman – Wunsch algorithm ([6] etc.) are recorded, it takes about 28 hours (at a processor clock frequency of about 2 GHz, see [4]). Therefore, one of the problems considered in the biocybernetics is the problem of *restoration* of the distance matrix between DNA sequences (below we simply shall write "DNA matrix"), in which *not all* elements of the matrix under consideration are known at the input of the algorithm, see [7, 8], compare also [9, 10]. In connection with all this, another problem arises: to use the developed method of comparative evaluation of algorithms for calculating distances between sequences for a completely different purpose, namely, for the briefly described *problems of reconstructing the distance matrix* between DNA sequences. For this problem, in this article we consider the application of the previously developed *method of comparative evaluation of distance calculation algorithms* between a pair of DNA strings.

Using this approach (i.e., using the method of comparative evaluation of algorithms for calculating distances to matrix reconstruction), the reconstruction itself occurs as a result of several computational passes. On each of the passes, for some of the as-yet-unfilled (unknown) elements of the matrix, different estimates are obtained; these estimates are averaged in a special way-and the result of averaging is taken as the value of the unknown element. From the physical point of view, the applied averaging gives the position of the center of gravity of a one-dimensional system of bodies whose mass is specified by a special function, i.e., the risk function, see [11, 12]. We note that earlier we used risk functions in completely different subject areas; these areas were always connected with auxiliary algorithms related to multicriteria optimization.

Below, the matrices to be reconstructed are referred to as *incompletely filled with distance matrices*. We *enter this term for the matrix*, from which a number of elements are "crossed out".

2. Preliminaries

Thus, like our previous papers, we consider the square symmetric matrix of distances between genomes. There it is necessary to note the following. First, the genomes we choose random enough and took them off the site [13]. Second, like previous works, we consider in fact *three* variants for each of the considered problems:

- for very distant species, including, for example, a mammal "Bison bison" and a reptilia "Apalone spinifera" (we use the official scientific Latin names), see detailed the species' list by the link for direct download [14];
- for a sufficiently close species (human and apes);
- and also for human races (in fact, they can be considered as subspecies of a biological species).

Let us note, looking ahead, we believe that our a theoretical construction is best applicable for more distant species, however, acceptable results are obtained also in two other cases.

Thus, we look at algorithms of comparing the quality of different algorithms that calculate the distance between two genomes. Apparently, all these algorithms are based on the use of various versions of the Levenshtein distance (or Levenshtein metric), see [15] and very many other following papers. It is very important to note, that for the simplest formulation of the problem (to make the strict calculating the value of Levenshtein metric for two given genomes), we unfortunately obtain a very long-running algorithm (program): it has to do with the actual length of the strings of genomes. Therefore, in each of the actual algorithms (see [16, 17, 18, 19] etc.), computation of distances between genomes in reality is a heuristic extension of the exact algorithm for calculating the Levenshtein distance. And apparently, the approach closest to our one is given by [20]; let us note, a little running forward, that it also uses a version of the branch-and-bound method.

Thus, the considered in our previous papers Panin's metric is no exclusion, It also is a heuristic algorithm for calculating the close version of such metric; it is an optimization problems, see [21] etc. However, we used in it a special approach (so called *multi-heuristic approach for discrete optimization problems*), and for it, we use the same heuristic as in very different problems. From many such problems, let us mention two ones only:

- the classical traveling salesman problem (however, we consider our own approach to this problem, and, most importantly, our original way of specifying the input data, different from the traditional geometric placement, see, e.g., [22]);
- and the problem of state-minimization for nondeterministic finite automata, see, e.g.,[23, 24].

3. The triangular norms: their study and possible attempts of improvement

It was justified in our previous works cited above, there is desirable that in the matrix of distances between genomes, any of the resulting triangles be close to an acute angled isosceles one with two angles exceeding 60 degrees. Several various empirically selected numerical characteristics describing such differences are also given in our works, see [4] etc. However, in the previous articles we did not consider detailed examples, let us consider them in this section.

In [14], the results of calculations for several metrics and several norms are presented. In this section, we shall consider the Panin's metric only, and 3 norms ("badnesses" for the triangle under consideration). Thus, for each triangle with the sides $a \geq b \geq c$ and the angles $\alpha \geq \beta \geq \gamma$ we considered the following norms:

$$\mathrm{bad}_1 = (\alpha - \beta)/\pi\,, \quad \mathrm{bad}_2 = (\alpha - \beta)/\alpha\,, \quad \mathrm{bad}_3 = (a - b)/a\,.$$

In case if $\alpha \geq 90^o$, we have considered each norm by the maximum possible (1.0) or even usually exceeding this value. We assigned an even greater value to the value of badness in the case when the three considered sides do not form a triangle at all (that is, they do not satisfy the triangle inequality); let us note, running ahead, that similar situations is happened for any of metrics considered by us.

The resulting value of the norm *of the whole matrix* was considered as the arithmetic mean of the norms of all triangles. We note that for the matrices of distances between genomes (usually from 30×30 to 50×50), the number of triangles is

$$\frac{30 \cdot 29 \cdot 28}{2 \cdot 3} = 4060$$

for dimensions 30, and 19600 for dimension 50; from these values, it is clear that the calculations we need are quite difficult.

Thus, let us consider the part of the table given on the page titled "Panin's metrics" of [14] (they are designed as an `xlsx`-file and are available for the direct download), see the table on Fig. 1. (The names of the considered species can be found there on the page titled "Types of animals".)

No. of genomes	1	2	3	4	5	6	7	8	9	10	...	15	...	25	...
1	0	2200	2904	5996	2580	4149	4206	6336	3057	3222	...	3182	...	2300	...
2	2200	0	2638	5998	2860	3922	4416	6000	3373	2982	...	2962	...	2150	...
3	2904	2638	0	6068	2890	4037	4639	6414	3647	3202	...	3201	...	2703	...
4	5996	5998	6068	0	5647	5849	5918	6066	5858	5508	...	5596	...	5618	...
5	2580	2860	2890	5647	0	4426	4145	6445	3274	3589	...	3533	...	3142	...
6	4149	3922	4037	5849	4426	0	4682	6492	4397	3996	...	4006	...	3919	...
7	4206	4416	4639	5918	4145	4682	0	6581	4230	4586	...	4577	...	4571	...
8	6336	6000	6414	6066	6445	6492	6581	0	5893	5731	...	5776	...	5950	...
9	3057	3373	3647	5858	3274	4397	4230	5893	0	3651	...	3579	...	3447	...
10	3222	2982	3202	5508	3589	3996	4586	5731	3651	0	...	1985	...	2953	...
...	...	...	...	...	...	...	...	...	...	...	0	...	...	...	...
15	3182	2962	3201	5596	3533	4006	4577	5776	3579	1985	...	0	...	2940	...
...	...	...	...	...	...	...	...	...	...	...	...	...	0	...	...
25	2300	2150	2703	5618	3142	3919	4571	5850	3447	2953	...	2940	...	0	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	0

Figure 1. The part of the table of Panin's metric.

Let us choose species with the numbers 5, 15, 25; while doing so, we specifically chose exactly the three of those considered, where the metric gives rather poor results (that means the following: the badnesses of Panin's metric give relatively worse results than other metrics comparing most other triangles of the matrix under consideration). For all 5 considered metrics, we obtain the following 5 triangles corresponding to the species with the numbers 5, 15, 25:

1) sides 3533, 3142, and 2940 (Panin's metric);
2) sides 1215, 1179, and 704 (van der Loo's 1st metric);
3) sides 4379, 4036, and 4029 (van der Loo's 2nd metric);
4) sides 2943, 2674, and 2492 (Pages' 1st metric);
5) sides 3444, 3046, and 2838 (Pages' 2nd metric)

(here, the numbers correspond to numbers of metrics of [4, 14]). Let us consider these 5 triangles and also 3 other ones, see Fig. 2:

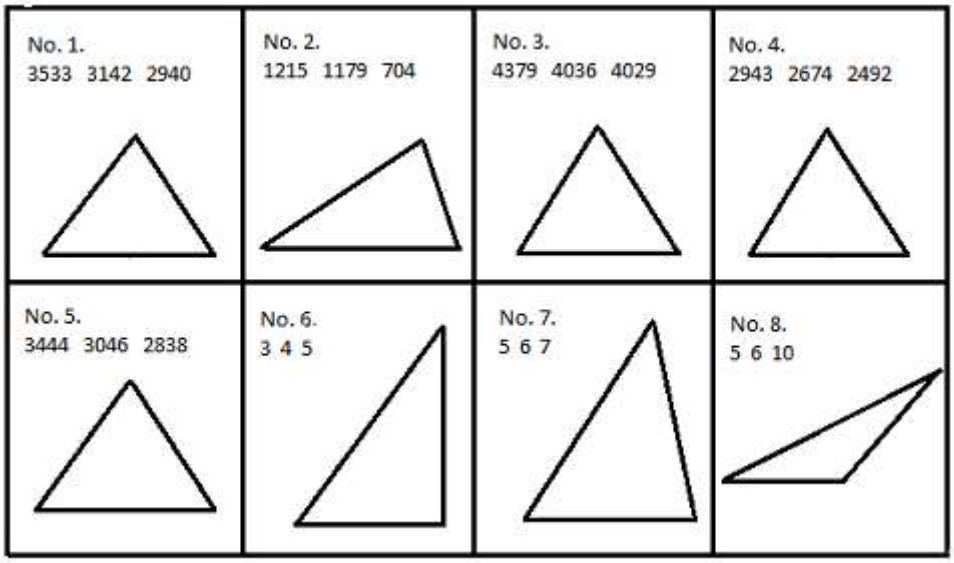


Figure 2. Examples of triangles of metrics (No.No. 1–5) and 3 other ones.

Let us note once again, that we only need the relative lengths of the sides of the triangle.

The further calculations yield the following auxiliary values and values of badnesses, see Table 1.

Table 1. The values of badnesses and the auxiliary calculations.

No.	a	b	c	$\cos(\alpha)$	$\cos(\beta)$	$\cos(\gamma)$	$\alpha,^o$	$\beta,^o$	$\gamma,^o$	bad_1	bad_2	bad_3
1	3533	3142	2940	0.33	0.54	0.62	70.1	57.2	51.8	0.076	0.194	0.111
2	1215	1179	704	0.25	0.34	0.83	75.4	70.2	34.2	0.029	0.068	0.027
3	4379	4036	4029	0.41	0.541	0.544	65.8	57.19	57.05	0.048	0.130	0.078
4	2943	2674	2492	0.35	0.53	0.61	69.4	58.1	52.5	0.062	0.160	0.091
5	3444	3046	2838	0.32	0.54	0.62	71.6	57.0	51.4	0.081	0.203	0.116
6	5	4	3	0	0.6	0.8	90	53.1	36.9	–	–	–
7	7	6	5	0.2	0.54	0.71	78.4	57.1	44.4	0.119	0.272	0.143
8	10	6	5	−0.65	0.89	0.93	130.5	27.1	22.3	–	–	–

We can see, that three bad orderings for all five triangles give the same sequence of metrics. Once again, we mention (above, this thing has already been said, but in connection with other facts) that this ordering differs significantly from the ordering of the badness considered for complete matrices, see the results of the calculations below and, in more detail, in [4]. However, in all our calculations (in any case, with the exception of less than 1% of all triangles, i.e., including ordering for complete matrices), such ordering turns out to be the same for all three norms (badnesses).

Another option for investigating the comparative characteristics of norms is the following one (we will continue to consider 30 species, and the matrix of distance between genomes, given in [14]). We choose two metrics and for any one fixed norm we arrange 4060 triangles in order of increasing values of this norm. At the same time, when reading that both these norms are good (that is, they give acceptable results), we should ideally obtain *identical sequences* of triangles. Actually, one of these two sequences of triangles is obtained from the other by some sequence transpositions of neighboring elements. Since, as we have noted, the number of triangles in our case is 4060, then the maximum possible number of transpositions of neighboring elements is

equal to

$$\frac{4060 \cdot 4059}{2} = 8\,239\,770.$$

Let us give concrete results of work (Table 2) for the first norm (value "bad_1") only.

Table 2. The joint study of pairs of metrics for the selected first norm ($\mathrm{bad}1$).

Pairs	$\{1,2\}$	$\{1,3\}$	$\{1,4\}$	$\{1,5\}$	$\{2,3\}$	$\{2,4\}$	$\{2,5\}$	$\{3,4\}$	$\{3,5\}$	$\{4,5\}$
Transpositions	175199	163214	169720	175267	154301	159737	160614	180413	179384	181700
Percentages	2.13	1.98	2.06	2.13	1.87	1.94	1.95	2.19	2.18	2.21
Correlation	0.957	0.960	0.959	0.957	0.963	0.961	0.961	0.956	0.956	0.956

Let us give some comments to this table. The pair of metrics is selected in the first line. The number of transpositions of neighboring elements (for obtaining the monotone sequence) is given in the second line. The percent of the maximum possible number of transpositions of neighboring elements (8 239 770) is given in the third line. We do *not* use Spearman's rank correlation coefficient (and some others correlation coefficients used in similar problems); instead of them, we use the linear function having value 1 for 0 transpositions and value -1 for 8 239 770 transpositions.

Still, we note that the two other norms give somewhat worse results; but for them, the number of transpositions does not exceed 600 000 (i.e., less than 7.5%), and, therefore, calculated by our method rank correlation coefficient is more than 0.85.

Thus, all three norms almost always give similar results. Therefore we often use the singular for the word "norm": "some different versions of so-called triangular norm" etc.

4. About one method of DNA matrix reconstruction

In this section, one of the methods of comparative analysis of various algorithms for calculating distances between DNA sequences is presented, and on the basis of this, a *method for reconstructing an incompletely filled matrix is developed.* In order to carry out this comparative analysis, we propose for the resulting algorithm to calculate the distances between genomes, consider all possible triangles, because ideally they should be acute-angled isosceles.

To answer the question of how "correct" is the matrix obtained as a result of some heuristic algorithm, we propose to use the "characteristic of the departure" of the obtained triangles from "elongated isosceles" triangles; i.e., the "badness", below we shall write this term without quotes. In this case, as one of the variants of the badness, the following formula can be used:

$$\sigma = \frac{\alpha - \beta}{\gamma} \tag{1}$$

where α, β and γ are the angles, and we admit, that $\alpha \geq \beta \geq \gamma$ [4]. In the opinion of the authors of this paper, this formula best characterizes the requirements described by us. Here is the *informal* explanation for this: The closer the triangle to the isosceles triangle, the less the difference between α and β, and in the ideal case, the numerator is 0; in accordance with our assumptions, an obtuse (or rectangular) isosceles triangle cannot be obtained. The performance of the properties of acute angles increases the denominator. Consequently, the approximation of a triangle to an isosceles triangle reduces the numerator and increases the denominator in the formula, i.e., σ tends to 0.

When calculating the badness of the entire matrix for each recovery option, we can:

- either summing the corresponding badness over all possible triangles of the matrices in question;

- or take the maximum badness for these triangles.

In the future, we propose to consider other approaches to calculating the badness of the entire matrix.

To determine the unknown element, we consider all the possible triangles formed from elements of this matrix for which one of the sides is unknown. For each such triangle, from the condition that it is isosceles acute-angled, we get *one of the possible values* of this unknown side. Next, we calculate the final value of this side (an unknown element) in a special way. Namely, for its calculation, on the basis of all the estimates obtained, the element is assumed to be equal to the arithmetic mean of all the values obtained. As an alternative, we can exclude the largest and smallest of the values obtained.

With a large number of missing elements, the matrix of the triangles with two known sides will be small, so the restoration of the matrix in one pass is usually impossible. When restoring the matrix on the second and subsequent passages, we can either use only the elements of the matrix of the last pass, or use all the matrices obtained in the previous passes. In the second case, with each successive passage in the matrix, there are more and more elements calculated approximately. Therefore, when evaluating an unknown element, it is possible to use the analog of the risk function, which will adjust the weight of the elements *depending on the pass number*.

When using the so-called *static* risk function, the weight of the elements with each pass decreases with the same coefficient, and to estimate the unknown element of the matrix, formula

$$E = \frac{c_0 E_0 + c_1 E_1 + \ldots + c_k E_k}{c_0 + c_1 + \ldots + c_k}, \tag{2}$$

where:

- E_i is the value of the matrix element, calculated on i-th pass;
- $c_0, \ldots, c_k$ are some specially chosen coefficients.

In practice (see [5]), good results are achieved when the following formulas are used for the coefficients:

$$c_0 = 1, \quad c_i = p\, c_{i-1}. \tag{3}$$

By [11, 12], the risk function can be also *dynamic*: when using it, we take averaging, depending on the "rough estimate" of the final value: whether it is "good", "middle" or "bad". Besides, we can also consider the *sequence* of such dynamic risk functions, where at each stage we rely on the value obtained in the previous step as such a "rough estimate". In our case, to evaluate the unknown element of the matrix of distances between DNA strings, the formula

$$\frac{\sum_{i=1}^{k} a_i f(a_i)}{\sum_{i=1}^{k} f(a_i)}, \tag{4}$$

is used; $f(x)$ is some specially chosen decreasing function.

5. The formal description of the algorithm

Let us consider the detailed *formal description* of the algorithm briefly considered before.

Algorithm 1 (Restoring a matrix using a static risk function)

Input: Incompletely defined matrix $A = a_{ij}$ (all elements equal to zero outside the main diagonal are assumed to be unknown).

Used auxiliary variables: b_i is the array of unknown element estimates.

Description of the algorithm.

Step 1: We set $s := 1$ (the number of the pass).

Step 2: We count h, i.e. the number of elements of the upper triangle, which are equal to 0.

Step 3:
if $a_{ij} = 0$ and $i \neq j$ then
  begin
    $kol := 0$ {We count the number of triangles,
      built on the unknown element under consideration}
    for $k := 0$ to n do begin
      if $k \neq i$ and $k \neq j$ and $a_{ki} \neq 0$ and $a_{kj} \neq 0$ then
        begin
          $kol := kol + 1$; $c_0 := 1$; $c_s := c_{s-1} \cdot p$;

$$E_{ki} := \frac{c_0 E^0_{ki} + \ldots + c_s E^s_{ki}}{c_0 + \ldots + c_s};$$

$$E_{kj} := \frac{c_0 E^0_{kj} + \ldots + c_s E^s_{kj}}{c_0 + \ldots + c_s};$$

          if $E_{ki} > E_{kj}$ then $b_{kol} := E_{ki}$ else $b_{kol} := E_{kj}$
        end;
      end;
  end;

$$a_{ij} := \frac{b_1 + \ldots + b_{kol}}{kol}.$$

Step 4: We count h_1, i.e., the number of elements of the upper triangle, which are equal to 0 after the next pass.

Step 5:
if $h_1 = 0$ then goto Output 1;
if $h_1 = h$ then goto Output 2;
$s := s + 1$;
goto Step 2.

Output 1: Filled matrix A.

Output 2: Matrix A cannot be filled.

End of description of the algorithm.

After execution of the algorithm for performing a comparative analysis of the results of the reconstruction of the matrix, we use such an indicator as the residual; it characterizes the deviation of the resulting matrix from the original one. We calculate the residual on the basis of the natural metric

$$d = \frac{\sqrt{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (a_{ij} - \widetilde{a_{ij}})^2}}{n(n-1)/2}, \tag{5}$$

where:

- $\widetilde{a_{ij}}$ are elements of the matrix obtained as a result of applying some algorithm for calculating the distances between a pair of genomes (in our case, the Needleman – Wunsch algorithm);
- a_{ij} are elements of the matrix restored as a result of the above algorithm.

Some examples of operation of the algorithm and corresponding values of residual are given in [5].

6. Conclusion

Thus, a very small change of the given matrix greatly reduces its badness. The values marked in red are chosen by us manually. However, at the present time we have a neural network that implements such an algorithm; we are going to describe this neural network in the following publication. But the following is much more important: these "red" changes give *input* information to another neural network, the one that computes the constants for the metric. Thus, along with one "loop" of algorithms already described in previous section, we get one more. Table 3 above already takes into account similar improvements in the metric.

As we said before, the received results of our computer programs are designed as an `xlsx`-file and are available for the direct download by [14]. The whole article i s actually a comment to this file.

We also note that different "places" of different metrics for different cases (i.e., the case of distant animal species, the case of close animal species and the case of subspecies) talk about the need to continue research in this direction.

Also in the near future we expect to develop other approaches for comparative analysis of various algorithms for calculating distances between sequences, as well as describe the algorithms for reconstructing matrices based on these approaches. At present, we are working on comparing two of such algorithms, both for application in the "normal" problems of DNA analysis, and in the problems of DNA matrix reconstruction close to those considered in the present paper.

7. References

[1] Melnikov B and Panin A 2012 On a parallel implementation of the multi-heuristic approach in the problem of comparison of genetic sequences *Vektor Nauki of Togliatti State University* **4(22)** 83-86 (in Russian)
[2] Makarkin S, Melnikov B and Panin A 2013 A parallel implementation of the multi-heuristic approach in the task of comparing genetic sequences *Applied Mathematics* **4(10)** 35-39 DOI: 10.4236/am.2013.410A1006
[3] Melnikov B, Pivneva S and Trifonov M 2017 Multiheuristic approach to compare the quality of defined metrics on the set of DNA sequences *Modern Information Technologies and IT Education* **13(2)** 89-96 (in Russian) DOI: 10.25559/SITITO.2017.2.235
[4] Melnikov B, Pivneva S and Trifonov M 2017 Various algorithms, calculating distances of DNA sequences, and some computational recommendations for use such algorithms *CEUR Workshop Proceedings* **1902** 43-50
[5] Melnikov B and Trenina M 2018 On a problem of the reconstruction of distance matrices between DNA sequences *International Journal of Open Information Technologies* **6** 1-13 (in Russian)
[6] Needleman S and Wunsch Ch 1970 A general method applicable to the search for similarities in the amino acid sequence of two proteins *Journal of Molecular Biology* **48(3)** 443-453
[7] Eckes B, Nischt R and Krieg T 2010 Cell-matrix interactions in dermal repair and scarring *Fibrogenesis Tissue Repair* **3(4)** 1-13 DOI: 10.1186/1755-1536-3-4
[8] Midwood K S, Williams L V and Schwarzbauer J E 2004 Tissue repair and the dynamics of the extracellular matrix *International Journal of Biochemistry & Cell Biology* **36(6)** 1031-1037
[9] Evdokimova N I and Kuznetsov A V 2017 Local patterns in the copy-move detection problem solution *Computer Optics* **41(1)** 79-87 DOI: 10.18287/2412-6179-2017-41-1-79-87
[10] Evsutin O O, Shelupanov A A, Meshcheryakov R V and Bondarenko D O 2017 An algorithm for information embedding into compressed digital images based on replacement procedures with use of optimization *Computer Optics* **41(3)** 412-421 DOI: 10.18287/2412-6179-2017-41-3-412-421
[11] Melnikov B and Radionov A 1998 On the choice of strategy in nondeterministic antagonistic games *Programming and Computer Software* **5** 55-62 (in Russian)
[12] Melnikov B 2001 Heuristics in programming of nondeterministic games *Programming and Computer Software* **5** 277-288

[13] *Nucleotide (The Nucleotide database)* (Access mode: http://www.ncbi.nlm.nih.gov/nuccore)
[14] Melnikov B *The processed results of the computer calculations* (Access mode: http://bormel.ru/BorMel-DNA.xlsx)
[15] Levenshtein V 1966 Binary codes capable of correcting deletions, insertions, and reversals *Soviet Physics Doklady* **10(8)** 707-710
[16] Winkler W 1990 String comparator metrics and enhanced decision rules in the Fellegi-Sunter model of record linkage *Proceedings of the survey research methods sections, American Statistical Association* **4(22)** 354-359
[17] Pages H, Aboyoun P, Gentleman R and DebRaoy S *Biostrings: String objects representing biological sequences, and matching algorithms* (Access mode: https://rdrr.io/bioc/Biostrings/)
[18] Morgan M, Anders S and Lawrence M *ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data* (Access mode: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2752612/)
[19] Van der Loo M 2014 The Stringdist Package for Approximate String Matching *R Journal* **6** 111-122
[20] Althaus E, Caprara A, Lenhof H-P and Reinert K 2006 A branch-and-cut algorithm for multiple sequence alignment *Mathematical Programming* **105** 387-425
[21] Melnikov B 2006 Multiheuristic approach to discrete optimization problems *Cybernetics and Systems Analysis* **42(3)** 335-41
[22] Makarkin S and Melnikov B 2013 Geometrical methods for solving the pseudo-geometric version of the traveling salesman problem *Stochastic optimization in informatics* **9(2)** 54-72 (in Russian)
[23] Melnikov B 2000 Once more about the state-minimization of the nondeterministic finite automata *Journal of Applied Mathematics and Computing* **7(3)** 655-662
[24] Melnikov B and Tsyganov A 2012 The state minimization problem for nondeterministic finite automata: the parallel implementation of the truncated branch and bound method *Proceedings 5th International Symposium on Parallel Architectures, Algorithms and Programming (Taipei)* 194-201

Acknowledgements
The authors of the article express their gratitude to Vladislav Dudnikov (Togliatti State University, Russia) for his help in preparing this paper. The reported study was partially supported by RFBR according to the research project No. 16-47-630829.

The ensemble of algorithms for coronary heart disease detection based on electrocardiogram

V N Guryanova[1]

[1]Lomonosov Moscow State University, Leninskie Gory 1, Moscow, Russia, 119991

Abstract. Coronary heart disease (CHD) is the leading cause of death in the world. This disease can be asymptomatic for a long time and over time can progress and result in death. Today electrocardiogram (ECG) can be done at home with the help of special equipment from CardioQvark. In this paper the possibility of CHD detection based on such ECGs was explored. Different approaches to the classification of such electrocardiograms were surveyed. New algorithms and modifications to existing algorithms were proposed. A new method – the ensemble of different algorithms – has shown the best performance.

1. Introduction

Coronary heart disease(CHD) [1] is a group of diseases, that is defined by lack of oxygen supply to the heart muscle through the coronary arteries. According to World Health Organization, this disease is the leading cause of death in the world.

At the initial stages of the disease, most people do not show any symptoms of this disease. It is very important to identify the CHD in time to slow the course of the disease and prevent the patient's death.

Traditionally, CHD can be detected with the help of specialists and a number of tests. It should be noted that these tests take a significant amount of patient's time, and also require high qualification of the specialist who will conduct them. Since there are very few specialists and the number of potential patients is growing every year, the task of automatically determining CHD is extremely urgent at the present time. Currently, it is extremely relevant to create a device that will help determine the disease or its probability at home. Such devices will allow the person to be sent to a doctor in case of a high probability of having a CHD.

An electrocardiogram (ECG) – is a signal that reflects the electrical activity of the heart. ECG is one of the most affordable ways of diagnosing heart disease now due to its non-invasiveness and low cost. Currently, there are many different studies that show that the ECG can be used to determine CHD [2], [3], [4], [5].

CardioQvark (project site: www.cardioqvark.ru) has created a device in a form of a smartphone case that allows you to make ECG measurements at home. The CardioQVARK device is a portable electrocardiograph in the form of a smartphone case (iPhone 5 / 5s / SE / 6 / 6s), allowing to register data of bio-electrical activity of the heart from the first ECG lead and next leads: aVR, aVL, aVF, Vi (i = 1 ... 6) using the patient's cable. In this work, the possibility of CHD detection based on such ECGs from the first lead was explored.

There are different approaches to ECG classification problem. Some of them were surveyed in this work. New algorithms and modifications to existing algorithms were proposed. In order to improve the performance of the classification, an ensemble of 5 different methods was built. Each of these methods will be described below.

2. Data Description

All research was conducted on the basis of the following medical centers: NGHCI "Semashko Central Clinical Hospital 2", Federal State Scientific Institution "Petrovsky Russian Scientific Center of Surgery", Federal State-Funded Health Care Institution "City Clinical Hospital 4 Health Care Moscow Department", Federal State-Funded Health Care Institution "Moscow Clinical Scientific Center of Moscow Department Health Care", State Autonomous Health Care Institution of the Moscow Region "Clinical Center for Restorative Medicine and Rehabilitation".

A voluntary anonymous study included patients over 18 years of age. Annotated impersonal electrocardiograms (ECG) were recorded from the first ECG-lead using a CardioQVARK cardio monitor. The duration of each recording was 5 minutes. The measurement was taken in the sitting position, with support for the back, hands on the knees or on the table. Data was collected in dynamics of 3-10 observations with an interval of at least 12 hours between measurements.

The sample that was used for this task consists of 1798 cardiograms. It contained 1055 cardiograms of healthy patients and 743 cardiograms of patients with CHD. The sampling frequency was 1000 Hz.

Signals were preprocessed before applying machine learning algorithms. For preprocessing, a low-pass and high-pass Butterworth filters [6] of the second order were used. For low-pass filter cutoff frequency was 0.3 Hz. For high-pass filter cutoff frequency was 15 Hz. The signal trend was extracted using a median filter [6]. Then the trend was subtracted from the preprocessed signal.

3. Algorithms Description

Below are descriptions of the algorithms that were used to build the ensemble.

3.1. The algorithm based on the HRV signal

The idea which was used for this algorithm was described in the article [2]. In the ECG signal, R-peaks can be distinguished, which correspond to the person's pulse [6] . ECG signal is used to create heart rate variability signal (HRV signal). It is calculated as follows.

- R-peaks are computed.
- The intervals between two R-peaks (RR-intervals) are measured.
- Each value of the RR-interval is converted to 60/RR.

The main idea of this method is a construction of various groups of features from HRV signal.

The first group of features includes various entropic features, which indicate a measure of unpredictability in the signal. The following types of entropies are used: approximate entropy, sample entropy, and Shannon entropy. Each type of entropy is described in detail below.

Approximate entropy is calculated as follows. Here $x = (x_0, x_1, ..., x_{N-1})$ is the HRV signal of length N.

- The integer m and the real r are fixed.
- A set of vectors of the form $\overline{x_i^m} = (x_i, x_{i+1}, ..., x_{i+m-1})$, where $i \in [0, N-m]$ is composed.
- The values $C_i^m(r)$ are calculated as follows:

$$C_i^m(r) = \frac{\left|\overline{x_k^m} : d\left(\overline{x_i^m}, \overline{x_k^m}\right) \leq r, k \in [0, N-m]\right|}{N-m+1},$$

where

$$d\left(\overline{x_i^m}, \overline{x_k^m}\right) = \max_{a \in [0, m-1]} \left|\overline{x_i^m}(a) - \overline{x_k^m}(a)\right|,$$

where the $\overline{x_i^m}(a)$ – component a of the vector $\overline{x_i^m}$.

- The values $\Phi^m(r)$ are defined as:

$$\Phi^m(r) = (N - M + 1)^{-l} \sum_{i=1}^{N-m+1} \log(C_i^m(r)).$$

- Approximate entropy (ApEn) is defined as

$$\text{ApEn}(r) = \Phi^m(r) - \Phi^{m+1}(r).$$

In this paper, the approximate entropy was realized for $m = 10$ and $r = 0.2\text{std}(x)$, $\text{std}(x)$ – standard deviation of the signal x.

The Sample entropy is calculated as follows.

- Vectors $\overline{x_i^m}$ of length m and vectors $\overline{x_i^{m+1}}$ of length $m + 1$ are formed similar to those that were formed in approximate entropy.
- The values A and B are calculated as follows:

$$A(r) = \left|(\overline{x_k^{m+1}}, \overline{x_l^{m+1}}) : d(\overline{x_k^{m+1}}, \overline{x_l^{m+1}}) \le r,\ 0 \le k \le l \le N - m - 1\right|,$$

$$B(r) = \left|(\overline{x_k^m}, \overline{x_l^m}) : d(\overline{x_k^m}, \overline{x_l^m}) \le r,\ 0 \le k \le l \le N - m\right|,$$

where d is determined as in approximate entropy.

- Sample entropy (SampEn) is defined as

$$\text{SampEn (r)} = -\log \frac{A(r)}{B(r)}.$$

In this paper, the sample entropy was determined for $m = 10$, $r = 0.2\text{std(x)}$.

The Shannon entropy is calculated as

$$\text{ShanEn} = \sum_{f=1}^{k} p_f \log p_f,$$

where k is the number of different elements in the signal x, p_f is the frequency of the element f in the signal x.

The following group of features is based on a recurrence plot. This plot shows the frequency and duration of repetitions in the signal. The element (i, j) of the given plot is defined as 1, if $||x_i - x_j|| < \varepsilon$, or as 0 otherwise, where x is the HRV signal.

Based on this plot, the following features are calculated, N – the number of elements in the HRV signal, $l_{d\,\min}$ – the minimum length of the plot diagonal, $l_{v\,\min}$ – minimal length of the vertical line, $l_{d\,\max}$ – maximum length of the diagonal, $l_{v\,\max}$ – maximum length of the vertical line:

- Density of points (REC)

$$\text{REC} = \frac{1}{N^2} \sum_{i,j=0}^{N} R(i, j).$$

- The percentage of points that form the diagonal lines (DET)

$$\mathrm{DET} = \frac{\sum_{l=l_{d\min}}^{l_{d\max}} lP(l)}{\sum_{i,j}^{N} R(i,j)},$$

  $P(l)$ – is the number of diagonals of length l.
- The average length of the diagonals (L_{mean})

$$L_{\mathrm{mean}} = \frac{\sum_{l=l_{d\min}}^{l_{d\max}} lP(l)}{\sum_{l=l_{d\min}} P(l)}.$$

- Entropy of diagonal lines (EN_d)

$$\mathrm{EN}_d = -\sum_{l=l_{d\min}}^{l_{d\max}} p_l \log p_l,$$

  where p_l is the frequency of diagonal lines of length l.
- Entropy of vertical lines (EN_v)

$$\mathrm{EN}_v = -\sum_{l=l_{v\min}}^{l_{v\max}} p_{vl} \log p_{vl},$$

  where p_{vl} is the frequency of vertical lines of length l.

For the calculation of these features, the PyRQA library [7] was used.

Another group of features that is used for this approach is a group of features based on the Poincare plot. The Poincare plot is constructed as follows: for the signal $x = (x_0, x_1, ..., x_{N-1})$, the plot consists of the points

$$(x_0, x_1), (x_1, x_2), ..., (x_i, x_{i+1})$$

and so on. In this case, RR-intervals were used as a signal x. The following features are constructed:

- The standard deviation of distances from the points of the plot to the line $y = x$. This feature describes the local variability of RR-intervals.
- The standard deviation of distances from the points of the plot to the line $y = -x - 2RR_{\mathrm{mean}}$. RR_{mean} – is the average value of RR-intervals. This feature describes the long-term variability of RR-intervals.

The feature, which is based on the detrended fluctuation analysis. This method allows determining the self-dependence of the signal. The following cumulative sum is defined as:

$$x_{\mathrm{cumsum}}(t) = \sum_{i=1}^{t} (x(i) - \mu),$$

x is a signal consisting of RR-intervals, μ is the mean of the signal x.

The data is segmented with a window of size Δn. On each segment polynomial is found for the data, which most accurately represents it (usually linear). The union of all such polynomials forms a function $x_{\Delta n}(t)$, which is an approximation of the original function $x_{\mathrm{cumsum}}(t)$. Then there is the following function:

$$F(\Delta n) = \sqrt{\frac{1}{N}\sum_{t=1}^{N}[x_{\text{cumsum}}(t) - x_{\Delta n}(t)]^2},$$

where N is the length of the signal consisting of RR-intervals.

The feature is the slope of the line $\log F(\Delta n)$ to $\log(\Delta n)$. More information about this approach is written in [8]. In this paper, the implementation of this feature was computed using publicly available software Nonlinear measures for dynamical systems or nolds, version 0.3.2, which can be downloaded from (https://pypi.python.org/pypi/nolds).

The next feature that was used in this approach is the correlation of dimensions. This feature is a quantitative characteristic of the signal trajectory and is defined as follows.

- Vectors $\overline{x_i^m}$ of length m are similar to those that were formed in approximate entropy are constructed.
- The value g is defined as:

$$g(r) = \left|(\overline{x_k^m}, \overline{x_l^m}) : d(\overline{x_l^m}, \overline{x_k^m}) \le r,\ 0 \le k \le l \le N - m - 1\right|,$$

  $g -$ is the number of pairs of vectors whose distance is less than or equal to r.
- The value $C(r)$ is defined as

$$C(r) = \frac{g(r)}{N^2},$$

  where $N -$ is the length of the signal x.
- The correlation of dimensions ($D2$) is defined as

$$D2 = \lim_{r \to 0} \frac{\log C(r)}{\log(r)}.$$

In this paper, the implementation of this feature was computed using publicly available software Nonlinear measures for dynamical systems or nolds, version 0.3.2, which can be downloaded from (https://pypi.python.org/pypi/nolds).

The gradient boosting from the xgboost package [9] was used as the classifier in this approach.

3.2. The algorithm based on 3 different feature spaces

This algorithm is a mixture of 3 different feature spaces, which were previously used in the classification of biomedical signals.

The first group of features consists of the parameters of the Hjorth's parameters: activity, mobility, complexity [10]. These parameters were originally used as features for electroencephalograms. Later, these parameters were used in many works, including the classification of ECG signals [11].

The second group of features consists of statistical signal features: mean, standard deviation, signal minimum, signal maximum, skew, kurtosis, selective quantiles of order: 0.1, 0.25, 0.5, 0.75, 0.9, sums and sums of signal values squares that are above / below certain values of quantiles: 0.1, 0.25, 0.5, 0.75, 0.9.

The next group of features was suggested by Uspenskiy for the disease detection by patient's ECG [12]. To calculate these features, it is necessary to calculate the amplitudes of the R-peaks ($A(n)$), the distances between the R-peaks ($T(n)$), and the arctangent of their ratio

$$\alpha(n) = \text{acrtg}\ \frac{A(n)}{T(n)}.$$

Table 1. Signal encoding for Uspenskiy features

	A	B	C	D	E	F
$A(n+1) - A(n)$	+	−	+	−	+	−
$T(n+1) - T(n)$	+	−	−	+	+	−
$\alpha(n+1) - \alpha(n)$	+	+	+	−	−	−

It is assumed that the values of $A(n)$, $T(n)$ are not important, but signs of their increments are. The method of signal encoding based on all possible signs of increments of these quantities is presented in Table 1.

After the code representation of the signal is received, the three-gram selection is performed. The feature space is the number of occurrences of each of the possible three-grams in a given code sequence derived from the signal.

The logistic regression from the scikit-learn package [13] was used as the classifier in this approach.

3.3. The algorithm based on R-peak's neighborhood

The idea used in this algorithm was described in the article [14] for determining the state of the patient's heart in which he should be sent to the cardiac service. The feature space for this approach is constructed as follows.

- The neighborhoods of the signal R-peaks are allocated: 200 points before R-peak and 500 after.
- The averaged neighborhood is used as a feature space.

Neural network with the architecture described in Table 2 was used as the classification model for this algorithm.

Table 2. Neural Network Structure for the algorithm based on R-peaks neighborhoods

Input Layer	Shape = (700)
Dense Layer	Units Number = 90 Activation Function = sigmoid
Dense Layer	Units Number = 1 Activation Function = sigmoid

In this work, the neural network was implemented using the libraries Theano [15] and Lasagne [16].

3.4. The algorithm based on wavelet transformation

The idea for this algorithm was described in paper [17] for epilepsy detection based on ECG-signal and for identification of a person based on ECG-signal.

Wavelet signal transformation is a convolution of the signal with functions $\Psi(t)$, called wavelets. Such wavelet functions should possess specific properties:

$$\int_{-\infty}^{+\infty} \Psi(t)dt = 0 \quad \int_{-\infty}^{+\infty} |\Psi(t)|^2 dt < \infty.$$

Wavelet transformation allows achieving signal compression, with good performance of reproduced original signal [6].

All wavelet functions used in a wavelet transformation can be represented with prototype function $\psi(t)$ using scaling and shift. In case of discreet wavelet transformation all wavelet function can be written as:

$$\psi_{m,n}(t) = \frac{1}{\sqrt{2^m}} \quad \left(2^{-m}t - n\right).$$

In discrete wavelet transformation of $\psi^*_{m,n}$ function can be separated into two parts. The two types of functions correspond to coefficients of approximation and detail coefficients. In the present work approximation coefficients are used for new representation of the signal, and Daubechies wavelets [18] are used as a wavelet function.

After getting the coefficients of wavelet transformations, local segments are extracted from the signal. The segments are extracted by moving a window of certain length w with step s, with all elements within one window going into a separate segment. After such a procedure every signal is represented as a set of local segments.

All local segments in the training set are separated into k clusters using k-means algorithm. After this, every segment is replaced with the number of the cluster it belongs to. This way every signal is represented as a text of codewords, with each word representing the certain cluster.

In the implementation of described algorithm these parameter values were used: $w = 100$, $s = 30$, $k = 200$. K-means algorithm implementation from scikit-learn package was used.

In the train sample every local segment is replaced with the cluster it is the closest to. That means that for every local segment s_i in test sample cluster c it belongs to is determined using this formula:

$$c = \text{argmin } d(b_j, s_i), \quad d(b_j, s_i) = \sqrt{\sum_{k=1}^{w} (s_i^k - b_j^k)^2},$$

where b_j is a cluster center j, s_i^k (b_j^k) is a kth element of local segment i (cluster j).

Using this approach and transforming the input signal into text it is possible to use natural language processing algorithms. The paper's authors that suggested such encoding used the bag of words as a feature space. Feature description is a number of occurrences of each code word in a specific signal.

Features based on word2vec technology were used in order to extract dependencies between local segments in the signal. This approach was suggested by Google and it allows to use context-aware text processing, reducing the dimensions of the data [19].

Word2Vec model was trained using the length of embedding vector equals to 80. A mean of all vectors in the signal was used as a feature. The model was trained using package gensim [20]. The logistic regression from scikit-learn package [13] was used as a classification algorithm.

3.5. The algorithm based on bispectrum

Bispectrum is a function of two variables f_1 and f_2 that specify the frequencies, expressed by the following formula [21]:

$$B(f_1, f_2) = X(f_1)X(f_2)X^*(f_1 + f_2),$$

where $X(f)$ is the Fourier transform of the signal, and $X^*(f)$ is the complex conjugate of it. The signal bispectrum is usually calculated using a fast Fourier transform. A detailed description of the algorithm for the bispectrum computation of the signal can be found in [22].

During calculation the bispectrum of the signal, we obtain a two-dimensional matrix whose elements are complex numbers.

Based on the resulting matrix, the elements of which can be denoted as $a(i, j)$, we can associate each signal with a certain image. This image is calculated as follows. A new matrix $B = ||b_{i,j}||$ is calculated, the elements of which are equal to:

$$b(i, j) = \sqrt{\mathrm{Re}^2 a(i, j) + \mathrm{Im}^2 a(i, j)},$$

where Re denotes the real part of the complex number, and Im denotes the imaginary part of the complex number. The contour plot of matrix B is used as the image.

The authors of the article [23] have shown that coronary heart disease can be detected by analyzing images obtained from a signal bispectrum. The method described in the above article was to allocate the area of the region within the level lines to conclude that the patient had CHD. The results obtained by the authors allow concluding that bispectrum images can be used to detect CHD.

It was proposed to use neural networks for classification of such images. The architecture of the neural network is described in Table 3.

Table 3. Neural Network Structure for the algorithm based on bispectrum

Input Layer	Shape = (3, 80, 80)
Convolution layer	Filter size = (32, 5, 5) Offset = (2,2)
Dense Layer	Units Number = 30 Activation Function = LeakyRelu
Dense Layer	Units Number = 1 Activation Function = Sigmoid

In this work, the neural network was implemented using the libraries Theano [15] and Lasagne [16].

4. Methods of constructing ensembles of algorithms

In order to increase classification performance, it was suggested to use ensembles of the algorithms. Several methods of ensembling are described in this section.

4.1. Majority voting

Given a set of algorithms $A = (A_1, A_2, ..., A_n)$ and output a vector of predictions $a = (a_1, a_2, ..., a_n)$ then the resulting answer a of the ensemble is equal to

$$a = \mathrm{mode}(a_1, a_2, ..., a_n),$$

where mode is a statistic, that is equal to the element which is most often encountered in the predictions. If there are several of them a random one of them is chosen.

4.2. EM-algorithm

The main idea of the EM-algorithm [24] is data aggregation from different people about the same event in order to get a correct evaluation. Since the goal of ensembling is the aggregation of several different algorithms, the EM-algorithm is applicable to ensemble creation.

Algorithm description:

N – size of available data, n_{il}^k – whether the k-th algorithm has the answer l ($l \in \{1, 2\}$) to the data i ($i = 1...N$),

π_{jl}^k ($j \in \{1, 2\}$) – is the probability that the k th algorithm will return j when the true answer is l.

$T_{ij} = 1$, if the true answer for the data i is j, otherwise, it equals 0.

p_j – the probability of class j in the sample.

- Step 1: Initialize the matrices π to the ideal case. T initialize the voting value for the majority.
- Step 2: Recalculate the values of matrices π and p_j:

$$\pi_{jl}^k = \frac{\sum\limits_i T_{ij} n_{il}^k}{\sum\limits_l \sum\limits_i T_{ij} n_{il}^k} \quad p_j = \frac{\sum\limits_i T_{ij}}{N}$$

- Step 3. Recalculate T_{ij}:

$$T_{ij} = \frac{p_j \prod_{k=1}^K \prod_{l=1}^2 (\pi_{jl}^k)^{n_{il}^k}}{\sum_{q=1}^2 p_q \prod_{k=1}^K \prod_{l=1}^2 (\pi_{ql}^k)^{n_{il}^k}}$$

Repeat steps 2 and 3 until the matrices π stop changing.

At the end of this algorithm, we obtain the probability of the data belonging to each class in the matrix T. As an answer, the class is taken, the probability of belonging to which is the greatest.

5. Evaluation of Algorithms

Cross-validation was used to evaluate the performance of algorithm. To avoid overfitting the ECGs of one patient did not fall simultaneously into the training and test set. The following performance criteria was introduced:

$$\frac{1}{N} \sum_{i=1}^{N} \frac{\sum_{j=1}^{n_i} \mathbf{I}_{t_{ij}=p_{ij}}}{n_i},$$

where t_{ij} is the true value of the target variable for the cardiogram j of the patient i, p_{ij} – the predicted value of the target variable for the cardiogram j of patient i, n_i – the number of cardiograms of the patient i, N – the number of patients, $\mathbf{I}_{t_{ij}=p_{ij}}$ is the indicator function which equals to 1 if t_{ij} equals to p_{ij} and equals to 0 otherwise. This criterion is called patient performance. It allows us to evaluate how well the algorithm determines a person's disease by any of his cardiograms. In addition, this performance criterion does not depend on the number of cardiograms for each patient. ROC-AUC score [25] and F-score [25] were used for models evaluation.

6. Results

The results of evaluations are shown in Table 4, where the first column shows algorithms type or ensemble type. The algorithm based on wavelet transformation is included in two variants, with word2vec and without.

7. Conclusion

In the course of this paper, the following results were obtained. When CardioQvark equipment is used, it is possible to determine CHD with an accuracy of more than 0.81 for patient performance, with an accuracy greater than 0.77 for F-score and with an accuracy greater than 0.87 for ROC-AUC score. Word2vec can increase the performance of the classification method based on the wavelet transformation. Bispectrum can be used to classify CHD. The EM algorithm is applicable for ensemble and in this case, shows the best performance of classification for all selected performance criteria.

Table 4. CHD detection results

An Algorithm	Patient Performance	ROC-AUC	F-score
Bispectrum	0.7207	0.7418	0.7244
Wavelet transformation	0.741	0.8	0.6967
Wavelet transformation +word2vec	0.7501	0.8	0.6990
R-peak's neighborhood	0.7602	0.7988	0.703
The HRV signal	0.763	0.744	0.662
3 different feature spaces	0.7632	0.8042	0.70256
majority	0.806		0.77
EM	0.8108	0.8738	0.7784

8. References

[1] Gorbachev V V 2008 *I cardiac ischemia* (Minsk: Vysh. shk) p 479 (in Russian)
[2] Dua Sumeet 2012 Novel classification of coronary artery disease using heart rate variability analysis *Journal of Mechanics in Medicine and Biology* **12(4)** 1240017-1240019
[3] Giri D 2013 Automated diagnosis of coronary artery disease affected patients using LDA, PCA, ICA and discrete wavelet transform *Knowledge-Based Systems* **37** 274-282
[4] Acharya U 2017 Rajendra et al. Application of higher-order spectra for the characterization of coronary artery disease using electrocardiogram signals *Biomedical Signal Processing and Control* **31** 31-43
[5] Kumar M R B U and Rajendra A 2017 Characterization of coronary artery disease using flexible analytic wavelet transform applied on ECG signals *Biomedical Signal Processing and Control* **31** 301-308
[6] Rangayyan R M 2015 *Biomedical Signal Analysis* (John Wiley & Sons)
[7] Rawald T M and Sips N M 2017 PyRQA – Conducting recurrence quantification analysis on very long time series efficiently *Computers & Geosciences* **104** 101-108
[8] Kantelhardt J W 2001 Detecting long-range correlations with detrended fluctuation analysis *Physica A: Statistical Mechanics and its Applications* **295(3-4)** 441-454
[9] Tianqi C and Guestrin C 2016 Xgboost: a scalable tree boosting system P*roc. of the 22nd acm sigkdd international conference on knowledge discovery and data mining ACM*
[10] Hjorth B 1970 EEG analysis based on time domain properties *Electroencephalography and Clinical Neurophysiology* **29(3)** 306-310
[11] De Cooman T, Carrette E, Boon P, Meurs A and Van Huffel S 2014 September Online seizure detection in adults with temporal lobe epilepsy using single-lead *ECG Proceedings of the 22nd European Signal Processing Conference* 1532-1536
[12] Uspensky V 2008 *Theory and practice of diagnosis of diseases of internal organs by the method of information analysis of electrocardio signals* (Moscow: Economics and Informatics) p 116 (in Russian)

[13] Pedregosa F 2011 Scikit-learn: machine learning in Python *Journal of machine learning research* 2825-2830
[14] Ripoll V J R 2016 ECG assessment based on neural networks with pretraining *Applied Soft Computing* **49** 399-406
[15] Al-Rfou R, Alain G, Almahairi A, Angermueller C, Bahdanau D and Ballas N 2016 Theano: a Python framework for fast computation of mathematical expressions *Preprint arXiv:1605.02688*
[16] Dieleman S 2015 *Lasagne: First release* (Geneva, Switzerland: Zenodo)
[17] Wang J, Liu P, She M F, Nahavandi S and Kouzani A 2013 Bag-of-words representation for biomedical time series classification *Biomedical Signal Processing and Control* **8(6)** 634-644
[18] Liu C L 2010 *A tutorial of the wavelet transform* (Taiwan: NTUEE)
[19] Mikolov T 2013 *Efficient estimation of word representations in vector space* (Scottsdale, Arizona: ICLR Workshop)
[20] Rehurek R and Sojka P 2010 Software framework for topic modelling with large corpora *Proc. of the LREC Workshop on New Challenges for NLP Frameworks*
[21] Civera M L Z and Surace S 2016 Using bispectral analysis and neural networks to localise cracks in beam-like structures *Proc. of the 8th European Workshop On Structural Health Monitoring* 1542-1551
[22] Chrysostomos N, Mysore L and Raghuveer R 1987 Bispectrum estimation: A digital signal processing framework *Proc. of the IEEE* **75(7)** 869-891
[23] Al-Fahoum A, Al-Fraihat A and Al-Araida A 2014 Detection of cardiac ischaemia using bispectral analysis approach *Journal of medical engineering & technology* **38(6)** 311-316
[24] Dawid A and Skene A 1979 Maximum likelihood estimation of observer error-rates using the EM algorithm *Applied statistics* 20-28
[25] Sokolova M, Japkowicz N and Szpakowicz S 2006 Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation *Australian conference on artificial intelligence* **4304** 1015-1021

On an approach to assessing the inter-channel phase synchronization of electroencephalogram signals

Y V Obukhov[1], R A Tolmacheva[1], A F Polupanov[1] and L A Zhavoronkova[2]

[1]Kotel'nikov Institute of Radioengineering and Electronics of RAS, Mokhovaya 11-7, Moscow, Russia, 125009
[2]Institute of Higher Nervous Activity and Neurophysiologyof RAS, Butlerova 5a, Moscow, Russia, 117485

Abstract. A new approach to evaluation of the phase coherence of electroencephalogram (EEG) signals in different channels based on the calculation and comparison of definite phase characteristics of signals at the points of the ridges of their wavelet spectrograms is considered. The approach is applied to evaluation of inter-channel phase coherence of EEG signals at cognitive tests performed for a healthy subject and for a patient after a traumatic brain injury. The method makes possible to efficiently detect phase-coupled pairs of channels of EEG and distinguish them from phase-uncoupled ones.

1. Introduction

The study of EEG inter-channel coherence is a conventional method for diagnosing brain pathologies. Phase connectivity or phase synchronization of signals in two EEG channels is evaluated with the help of the coherence. Typically [1-3], the coherence of two signals is estimated using their normalized complex cross-correlation calculated by multiplying the normalized Fourier components of the signals. The coherence between the two channels of the EEG is defined as the linear dependence of two signals at a certain frequency [2]. Let $x_i(f)$ and $x_j(f)$ are the complex Fourier transformations of time series $x_i(t)$ and $x_j(t)$ of channels *i* and *j*, respectively. Then the cross-spectrum is defined as:

$$S_{ij}(f) = \langle x_i(f) x_j^*(f) \rangle \quad (1)$$

where * is the complex conjugation and <> is the mathematical expectation.

Coherence is defined as a normalized cross-spectrum [2]:

$$C_{ij}(f) = \frac{S_{ij}(f)}{(S_{ii}(f) S_{jj}(f))^{1/2}} \quad (2)$$

and the connectivity is defined as the absolute value of the coherence:

$$Coh_{ij}(f) = |C_{ij}(f)| \quad (3)$$

The phase connectivity is calculated using the phases of the signals *i* and *j*. If $x_i = r_i \exp(i\Phi_i)$ and $x_j = r_j \exp(i\Phi_j)$ are Fourier transforms of the signals, then the cross-spectrum is calculated as [2]:

$$S_{ij}(f) = \langle r_i r_j \exp(i\Delta\Phi) \rangle \quad (4)$$

where $\Delta\Phi = \Phi_i - \Phi_j$ is the phase difference of signals in channels *i* and *j* at a certain frequency.

In order to calculate the phase connectivity, the cross-spectrum is normalized to the 'global' amplitudes $\langle r_i^2 \rangle^{1/2}$ and $\langle r_j^2 \rangle^{1/2}$. If the signals in the two channels are independent, then $\Delta\Phi$ is a random number and the connectivity is equal zero. Phase connectivity or phase synchronization is defined as an unweighted average:

$$P = \langle \exp(i\Delta\Phi) \rangle \quad (5)$$

Further, the phase difference is averaged over a certain frequency range predetermined from the neurophysiological considerations. Usually these are the ranges corresponding to the EEG rhythms, such as delta, theta, alpha, etc. rhythms. Phase-coupled pairs of brain regions are received by

calculating the averaged phase difference between all pairs of signals and selecting the cut-off threshold.

The averaging of the coherence or the phase difference for different time periods and in the frequency range predetermined on the basis of neurophysiological experience is done in the coherent analysis. Shortcomings of this approach are considered in [4]. In this regard, the actual task is the development of the method for determining the phase-coupled pairs of signals from single trial data and for a more sustainable choice of the phase coherency threshold.

We consider a new approach to the evaluation of phase synchrony of non-stationary EEG signals in cognitive tests. As a criterion of phase synchronization of two signals the following condition [5] is considered:

$$\left|\varphi_{i,j}(t)\right| \le const \tag{6}$$

where $\varphi_{i,j}(t) = n\phi_i(t) - m\phi j(t)$, ϕ is the phase of the signal, n, m are integers.

We consider the case $n = m = 1$, which can easily be generalized to the case of any $n \neq m$.

2. Method description

A method of evaluation of the inter-channel phase coherence of EEG signals is based on the calculation and comparison of definite phase characteristics of signals in different channels at the points of the ridges of their wavelet spectrograms. In case the signal satisfies asymptotic properties, the wavelet transform can be approximated in the stationary phase approximation, the points of ridges are the points of the stationary phase where the instantaneous frequency of the signal is equal to the wavelet frequency [6, 7].

At first we find a ridge with the maximum value of $|W|$ at each reference point τ_i of the Morlet wavelet spectrogram:

$$W(\tau,T) = \frac{1}{\sqrt{T}} \int x(t)\psi\left(\frac{t-\tau}{T}\right)dt, \tag{7}$$

$$\psi(\eta) = \frac{1}{\sqrt{\pi F_b}} e^{2i\pi F_c \eta} e^{-\frac{\eta^2}{F_b}}, \tag{8}$$

where we accept $Fb = Fc = 1$ [8].

Further, at the points of the ridge $R_i = \max_{f(t_i)} W(f(t_i), t_i)$ we calculatethe phase characteristic of the signal which is defined as a product of the instantaneous frequency of the signal at the time t_i and time: $\varphi_i = 2\pi f(t_i) t_i$ and the difference $\Delta\varphi_{i,j} = \varphi_i - \varphi_j$ for two signals.

In this case points beyond the ridge of the wavelet spectrogram are not taken into account in the evaluation of the phase coherency. Some points of the ridges may not satisfy the asymptotic conditions, which will lead to errors in the calculation of the phase. However, this error seems to be substantially less than errors associated with averaging the phase difference in a wide frequency range.

By removing the frequency range of the processed ridge in the wavelet-spectrogram, we can further apply the described algorithm to distinguish another ridge (in another frequency range, etc.).

3. Results

Histograms of the values of portions $\rho_{i,j} = n_{i,j}/N$, where $n_{i,j}$ is number of reference points of ridges with $|\Delta\varphi_{i,j}| < 0.05\pi$, and N is a summary number of EEG signal reference points during the test, are represented in Figure 1. The first pair of leads (Figure 1a) can be referred to a phase-coupled pair. Another pair (Figure 1b) can be referred to a phase-unconnected pair. Figure 1a shows, that $\rho_{i,j}$, less than 0.1, can be considered as a background. We consider the threshold $\rho_{i,j}^{thr}$ equal to 0.15 and we will assume that above this value of the points portions the ridge correspond to the phase-coupled pairs of leads.

EEG of healthy subjects were analyzed, which performed cognitive tasks in isolation. Below, for example, the results of the phase connectivity analysis are presented for two cognitive tests. Some items that belong to the category "clothes" or "food" were randomly listed to the subject during the cognitive test (CT1). During the test, he counts in his mind the quantity of items belonging to one of these categories, and at the end of the test declares the result. When performing a cognitive test (CT2),

the doctor randomly called the time.The test subject must imagine in his mind the dial of the clock and the position of the clock hands on it in accordance with the time mentioned. If both clock hands are in the same half of the dial, he says "yes," and if they are in different halves, he keeps silent. All tests were performed for 60 seconds.

Distribution graphs of the portions of the reference points for pairs of EEG leads, based on the histograms obtained, were built and they were sorted in order of increasing $\rho_{i,j}$ with EEG records without tests, with cognitive tests. These distributions are shown in Figure 2 for a healthy subject and for a patient after a traumatic brain injury.

The distribution of the portions of the reference points by pairs of EEG leads, sorted in ascending order $\rho_{i,j}$ for certain tests correspond to lines of a certain color (blue line: recording EEG without test; red line: record EEG with cognitive test CT1; lilac line: recording of EEG with cognitive test CT2). The abscissa axis shows the number of pairs of EEG leads. In the article records of 19-channel EEG were analyzed, therefore the number of pairs of leads is 171. Figure 2 represent that for a certain value on the abscissa axis for each test corresponds, possibly, a different pair of EEG leads. Based on the obtained pairs of EEG leads, it is calculate the number of pairs of leads for each test above the threshold $\rho_{i,j}^{thr}$ and it can be concluded that for a healthy subject the number of such pairs is higher with cognitive tests than when record is without a test (CT1: $n_{i,j}$= 131, number of coupled pairs of EEG leads: 171-131=40; CT2: $n_{i,j}$=130, number of coupled pairs of EEG leads: 171-130=41; record is without test: $n_{i,j}$=150, number of coupled pairs of EEG leads: 171-150=21 (figure 4). As well as, according to figure 3, it is possible to $n_{i,j}$=144, number of coupled pairs of EEG leads: 171-144=27).

Pairs of EEG leads in a healthy subject with a cognitive test CT1 are given as an example on the figure 3a. These pairs of leads are absent when EEG record is without test. Pairs of EEG leads in a healthy subject with a cognitive test CT2 are given as an example on the figure 3b. These pairs of leads are absent when EEG record is without test. Pairs of EEG leads in a patient with craniocerebral trauma with a cognitive test CT1 are given as an example on the figure 3c.

a) b)

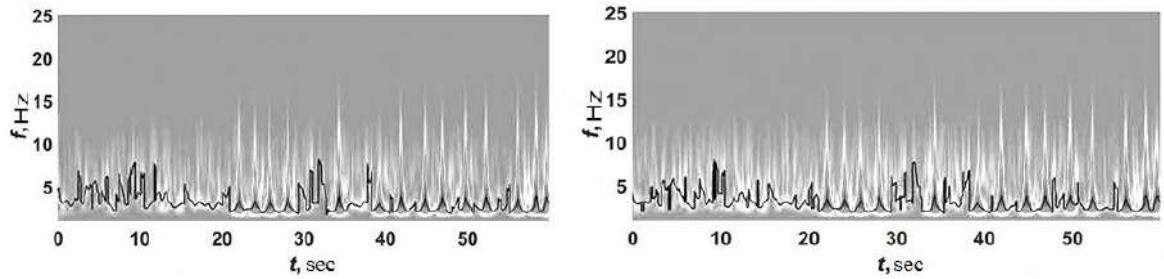


Figure 1. Wavelet spectrogram of the Morlet signal in a pair of EEG leads with a cognitive test with aselected ridge: a)For one lead. b) For the second lead. The abscissa axis is the time in seconds, the ordinate axis is the frequency in Hz, the color isproportinal to the power spectral density (red - max).

a) b)

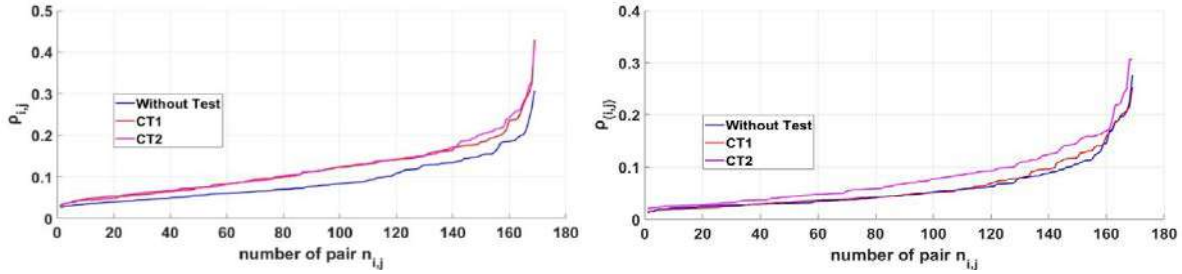


Figure 2. The distribution of the portions of the reference points for pairs of EEG leads, sorted in ascending order $\rho_{i,j}$. The abscissa is the number of the pair of EEG leads that correspond to different pairs of EEG leads. a) healthy subject; б) patient with craniocerebral trauma.

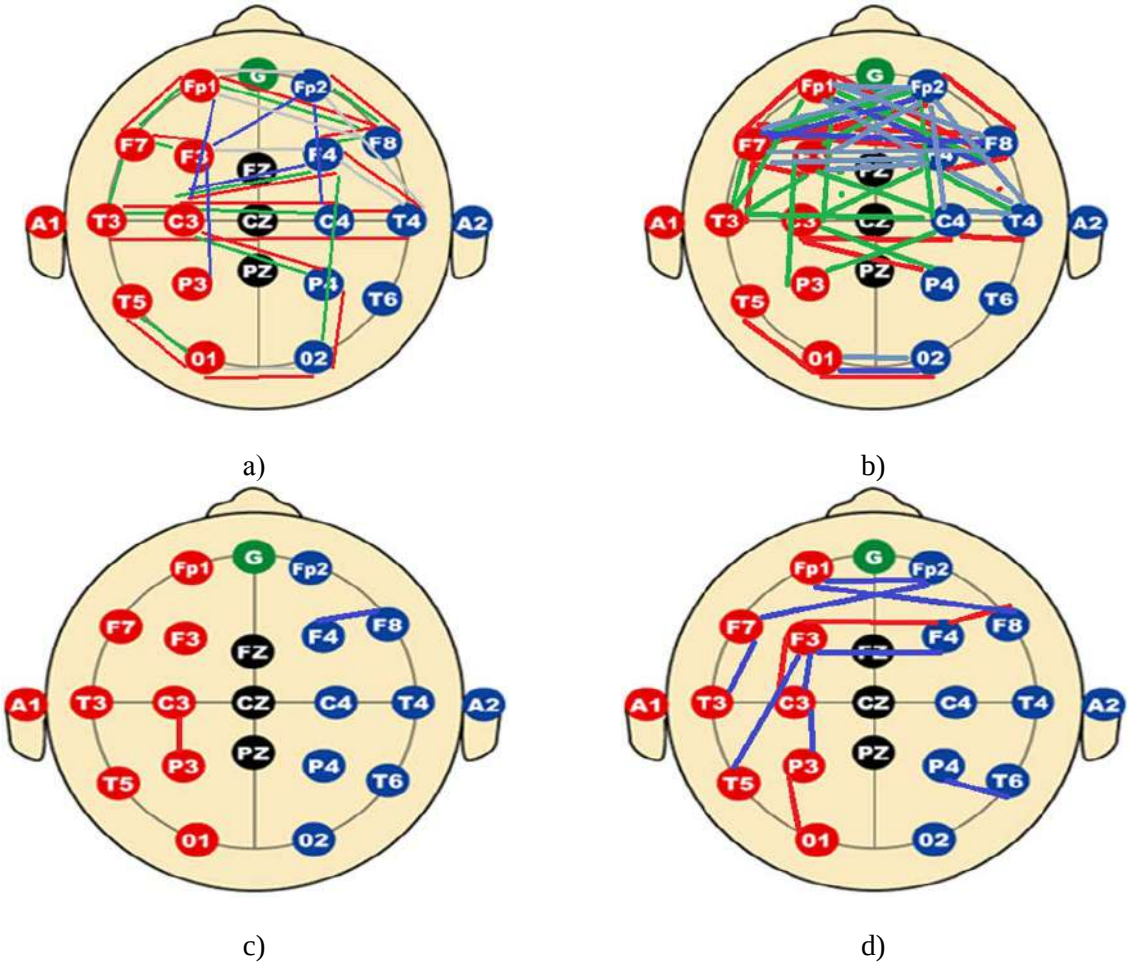


Figure 3. Distinctive pairs of EEG leads in cognitive tests CT1 and CT2 in a 4 healthy subjects: (a) – CT1, (b) – CT2; 2 patients with craniocerebral trauma: (c) – CT1, (d) – CT2.

These pairs of leads are absent when EEG record is without test. Pairs of EEG leads in a patients with craniocerebral trauma with a cognitive test CT2 are given as an example on the figure 3d. These pairs of leads are absent when EEG record is without test.

According to [9] the test CT1 is accompanied by a predominant activation of the prefrontal areas of the left hemisphere and the test CT2 is accompanied by a predominant activation of the prefrontal areas of the right hemisphere.

The distinctive phase-coupled pairs of leads demonstrate this (figure 3a and figure 3b).

4. Conclusion

The proposed approach improves the calculation of phase synchronization of EEG signals in order to get rid of the shortcomings of the existing estimates of coherence, related to the necessity of averaging the estimates over a wide range of frequencies. The method is tested in the problem of analysis of inter-channel phase synchronization in cognitive tests by healthy subjects and patients after craniocerebral injuries. The number of phase-coupled of the EEG leads is about the same as with cognitive tests for patients after craniocerebral injuries than when record is without a test.

5. References

[1] Wending F, Ansari-Asl K, Bartolomei F and Senhadji L 2009 From EEG signals to brain connectivity: a model-based evaluation of interdependence measures *J of Neuroscience Methods* **183(1)** 9-18

[2] Nolte G, Bai O, Wheaton L, Mari Z, Vorbach S and Hallet M 2004 Identifying true brain interaction from EEG data using the imaginary part of coherency *Clin. Neurophysiol.* **115** 2292-2307
[3] Zhana Y, Halliday D, Jiang P, Liu X and Feng J 2006 Detecting time-dependent coherence between non-stationary electrophysiological signals. A combined statistical and time-frequency approach *Journal of Neuroscience Methods* **156** 322-332
[4] Kulaichev A P 2009 On the informative value of coherence analysis in EEG studies *Journal of Higher Nervous Activity* **59(6)** 757-775
[5] Le Van Quyen M, Foucher J, Lachaux J P, Rodriguez E, Lutz A, Martinerie J and Varela F J 2001 Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony *Journal of Neuroscience Methods* **111** 83-98
[6] Mallat S A 2009 *Wavelet Tour of Signal Processing* (Elsevier Inc.)
[7] Guilleemain P and Kronland-Martinet R 1996 Characterization of Acoustic Signals Through Continuous Linear Time-Frequency Representations *Proceedings of the IEEE* **84(4)** 561-585
[8] Daubechies I 1992 *Ten lectures on Wavelets* (SIAM) p 194
[9] Zhavoronkova L, Shevtsova T and Maksakova O 2017 *How the human brain simultaneously solves two problems* (Saarbrücken: LAP LAMBERT Academic Publishing) p 68 (in Russian)

Acknowledgments
The reported study was funded by RFBR according to the research project № 18-07-00609.

Machine Learning in a Multi-Agent System for Distributed Computing Management

I V Bychkov[1], A G Feoktistov[1], I A Sidorov[1], A V Edelev[2], S A Gorsky[1] and R O Kostromin[1]

[1]Matrosov Institute for System Dynamics and Control Theory SB RAS, Lermontov St. 134, Irkutsk, Russia, 664033
[2]Melentiev Energy Systems Institute, Lermontov St. 130, Irkutsk, Russia, 664033

Abstract. We address the relevant problem of machine learning in a multi-agent system for distributed computing management. We propose a new approach to the agent learning in the system for managing job flows of scalable applications in a heterogeneous distributed computing environment, which includes high-performance computing clusters, as its main components. We manage parameter sweep applications that execute their jobs in a virtual machine environment. We use the specialized tools to implement such environment. In contrast to the known approaches, our approach is based on the integrated applying of methods for job classification and parameter adjustment of algorithms for functioning agents. Simulation modeling the environment allows eliciting the necessary knowledge for parameter adjustment. During the learning of agents, we use the expert knowledge of environment node administrators. An example of solving the complex practical problem that relates to studying energy development directions of Russia demonstrates advantages of the proposed approach.

1. Introduction

In the last decade, studies, connected with problems of strengthening subject orientation and intellectualizing technologies for development and use of a heterogeneous distributed computing environment (HDCE) that includes Grid-systems and cloud infrastructures, have become really relevant [1]. The progress in this direction is due to the need for the increasingly integrated use of heterogeneous environmental resources, as well as high-level support provided for end-users in the development and implementation of scalable applications.

A relevant approach to the distributed computing management in HDCE is applying multi-agent system (MAS) based on the use of market mechanisms while regulating the supply and demand of resources [2]. In such a system, an agent is a software entity that uses elements of artificial intelligence. Resource owners and their users endow it with rights and responsibilities to service and manage the computing process. Agents represent the interests of resources users and owners that often have conflicting criteria to define the computing process efficiency [3]. In processes of the executing the user jobs and coordinating their actions, they interact with each other. The agent coordination is based on their cooperation or competition. A selection of the cooperation or competition depends on their goals, roles, and mental properties.

The effectiveness of the agent work depends on the knowledge they use [4]. Stone [5] considers a wide range of basic capabilities and methods of the agent learning in systems with different

architectures. However, the agent learning remains an urgent problem in the tools of multi-agent computing management in practice, thus it reasonably requires development [6].

Various methods of machine learning have been developed [7, 8]. Usually, their purpose is an automatic improvement of decision-making quality over time in conditions of uncertainty in order to improve the efficiency of the controlled system [9]. Decision-making algorithms often depend on parameters that significantly affect the quality of management. A perspective direction of machine learning development is an integration of methods for an analysis of computational data and knowledge elicitation with the expert support of specialists in subject domains in the parametric adjustment process of these algorithms.

The paper addresses an approach to the agent learning based on a parametric adjustment of their algorithms for a job management in HDCE. In the computation management process, agents use the knowledge about both the problem specifics, which allow classifying jobs for their solving, and information about the environment, which ensure the rational distribution of the required resources.

2. Job management

In the paper, we consider HDCE that organized based on resources of the public access computer center "Irkutsk supercomputer center of the Siberian branch of the Russian Academy of Sciences" [10]. It supports two types of resources: dedicated (virtualized) and non-dedicated resources. The main components of the environment are high-performance clusters, the nodes of which differ in their computational characteristics.

MAS for jobs management in this environment includes agents for fulfillment the following operations:

- Problem formulation and problem-solving plan forming,
- Job classification,
- Creating a virtual community of agents representing environment resources,
- Parameter adjustment of agent functioning algorithms,
- Environment monitoring,
- Job dispatching in non-dedicated resources, etc.

These agents play the roles of the user agent, jobs-classification agent, agent-organizer, resource agent, parameter-adjustment agent, monitoring agent and agent-manager, respectively. Agents that represent environment resources can temporarily assume the role of agent-coordinator that regulates relationships of virtual community agents.

Representation of knowledge used by the agents is based on applying the conceptual model HDCE [11] that is a special case of a semantic network. In contrast to computational models of similar purpose (see, for example, [12]), such a model allows to carry out an interconnected description not only of algorithmic knowledge of subject areas for the solved problems, but also the knowledge about the hardware and software infrastructure of the environment and about administrative policies defined for its resources. The model includes the following knowledge components:

- Computational knowledge containing information on application modules for solving problems and system modules for computing planning, job-generating, resources allocation, monitoring of computational processes, dynamic decomposition of problems, and data preprocessing or postprocessing,
- Schematic knowledge comprising a set of objects (for example, parameters and operations) for describing the modular structure of the models and algorithms for the subject domain study,
- Production knowledge that defines the rules for applying operations and allows the applications end-users to select the best algorithms in the current computing situation,
- Infrastructure knowledge presented by the characteristics of hardware and software objects – nodes, communication channels, network devices, network topology and other structural elements, as well as information about their reliability,
- Administrative knowledge of policies against resources and users, including rules for the use of resources, rights and quotas for users and their jobs, and information about job management systems.

Figure 1 shows the job management scheme.

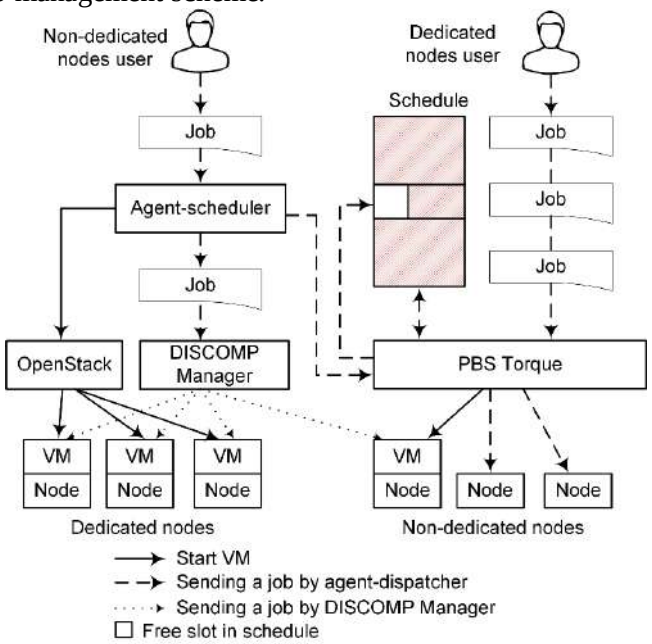


Figure 1. A scheme for multi-agent management of user jobs in a heterogeneous distributed computing environment using virtual machines.

Based on the problem formulation by the end-user of the application on the HDCE model, the user agent builds a set $P = \{p_1, p_2, ..., p_k\}$ of problem-solving plans. Then the agent-organizer integrates resource agents into a virtual community through the knowledge about a conformity of the assigned module classes and available environment resources. The virtual community includes agents that represent the resources in which modules of a problem-solving plan can be run. Virtual community participants elect the agent-coordinator in the process of their local interactions based on the modified tree algorithm taking into account communication topology of an agent network.

We apply the tender of computational work to distribute modules of problem-solving plans between agents. It is based on the one-round model second-price Vickrey auction [13]. Within the tender process, each agent makes the offers for executing modules. The agent-coordinator determines the bidders-winners in the tender. Applying the computational work tender allows using additional criteria for the job execution quality in addition to their cost that is the single condition in auctions. These are such criteria as the problem-solving time, computing reliability, information safety and other restrictions. In the case of parameter sweep applications in which each module is executed multiple times with different values of their input parameters, the bidding is conducted for the right to process data variants using this module.

The agent-coordinator conducts the tender and determines the optimal problem-solving plan $p_{opt} \in P$. It also selects resource agents (bidders-winners) participating in the plan execution. Determining the plan and its executors is implemented using the multicriteria lexicographic method of a selection taking into account the given problem-solving efficiency criteria ordered on the degree of their importance.

Executing the modules of p_{opt} in the allocated resources appointed by the agents is carried out by tools of the DISCOMP toolkit in the asynchronous mode upon the data is ready [14]. The agent-manager runs the required number of virtual machines (VMs) using the OpenStack platform tools [15] and transfers the job to DISCOMP manager, which then sends the task to run the modules to the DICCOMP clients hosted in the VMs. In case a job queue occurs in the allocated resources, the agent-manager directs the task to run the VMs in the non-dedicated resources if there are free slots in the job execution schedule of PBS Torque [16].

3. Machine learning of agents

The process of the agent learning is based on the complex use of methods of conceptual modeling, job classification and parameter adjustment of the management system. Table 1 shows the agents with methods, tools, and subjects of their learning.

Table 1. Methods, means, and subjects of machine learning of agents.

Agent	Method	Tool	Subject
Agent for problem-solving planning	Conceptual modeling of the subject domain	Toolkit DISCOMP, the XML language extension	Application developer
	Formulating problems and criteria of the efficiency of their solving		Application end-user
Job classification agent	Attributive description of job classes	Job classification System	Environment administrator
Agent-organizer of agent virtual community	Matching job classes and resources	Job classification system	Environment node administrators
Resource agent	Parameter adjustment	Simulation modeling system	Parameter adjustment agent
Parameter adjustment agent	Configuration adjustment	MAS configuration adjustment system	Environment administrator
	Environment monitoring	Meta-monitoring system	Meta-monitoring agent
Meta-monitoring agent	Configuration adjustment	MAS configuration adjustment system	Environment administrator
Job management agent in non-dedicated resource	Matching job classes and resources	Job classification system	Environment administrator

The subject domain model, problem formulations, and criteria for their solving are described through the DISCOMP tools by the application developer and its users in XML. Figure 2 and Figure 3 show fragments of such a description.

```
<parameters>
  <param name='model' type='file' filename='model.txt'>
  <param name='model_list' type='filelist'
         pattern='model_element_%1.txt'/>
  <param name='result_list' type='filelist'
         pattern='result_element_%1.txt'/>
</parameters>
<modules>
  <module name='decompose'>
    <commands os='Linux'>
      <start>decompose.exe</start>
    </commands>
    <parameters>
      <input><param name='model'/></input>
      <output><param name='model_list'/></output>
    </parameters>
  </module>
  …
</modules>
```

Figure 2. Subject domain model.

```
<process>
  <stage>
    <module name=' korrectiva_decompose'/>
  </stage>
  <stage>
    <listmodule name='korrectiva_solver'/>
  </stage>
  <stage>
    <module name=' korrectiva_analyse'/>
  </stage>
</process>
```

Figure 3. Problem formulation.

In the job classification system [17], node administrators, based on their practical skills and experience, define the set $H = \{h_1, h_2, ..., h_m\}$ of the possible job characteristics (problem-solving time, sizes of RAM and disk memory, number of nodes, processors and cores, module execution modes, etc.), and their domains (Figure 4). Next, they form the set $C = \{c_1, c_2, ..., c_n\}$ of job classes that have characteristics from H. In case the characteristic domain is included in the concrete class c_i, it can be specialized (restricted). The formed classes are mapped to the most appropriate resources for executing jobs that belong to those classes.

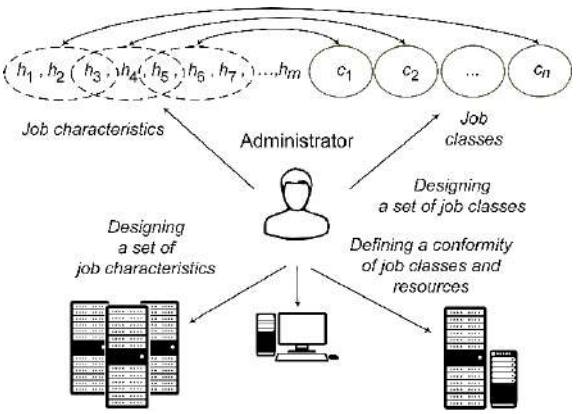


Figure 4. Designing the job classification system.

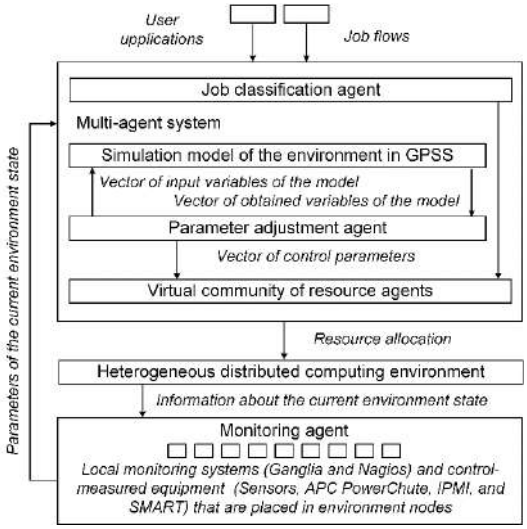


Figure 5. Parameter adjustment scheme.

The environment administrator sets the following configuration parameters:

- Parameter values that determine the intentions of resource agents to execute jobs of different classes,
- Lower and upper limits of the allowable deviation from the average resource load for agents of virtual communities,
- Amount of fines for the deviation from the average load (for resource agents),
- Composition and frequency of information collection, and data formats,
- Controlling and measuring means, and monitoring systems that will be used,
- Change limits of measured values and control actions applied when they are reached (for the monitoring agent);
- Permissible quotas on the number of jobs, their execution time and the number of nodes used,
- Characteristics of the slots in the PBS Torque system schedule in non-dedicated resources (for the jobs scheduling agent).

Figure 5 represents the parametric adjustment scheme of functioning algorithms of resource allocation agents. The agent-classifier identifies the job classes. The virtual community of resource agents is formed on the basis the matching classes to resources. These agents allocate resources. They use the tender of computational works and algorithms of their work.

The algorithms are determined by the vector of control parameters that provide agents with the selection of the optimal strategy of behavior. The parameter adjustment agent controls the parameters that reflect the values of the vector of input sweep variables of the HDCE simulation model. This vector corresponds to the optimal observed variables of the model that are calculated based on the parameter sweep computing. Multicriteria rules of the discrete selection ensure finding the optimal values [18].

The monitoring agent is designed to provide the HDCE subjects with up-to-date information on the loading of its resources, physical state of the equipment and engineering infrastructure devices [19]. In

contrast to other monitoring systems, an important feature of the monitoring agent use is the agent ability to analyze and apply control actions directly on the computing node where the agent operates. The monitoring agent collects, unifies, aggregates, and transmits data to the expert system for their analysis. In case critical events are detected, the necessary functions of the executive system of the agent are performed in order to apply control actions for automatic troubleshooting. At the same time, the administrator can pre-train monitoring agents taking into account the purpose of computing nodes and the jobs executed on them.

Evaluation of the successful agent learning is the class determination correctness by the agent-classifier and the resource use efficiency by their agents.

We apply an attributive description based on mandatory and optional sets of characteristics for a job specification [17]. If one or more optional characteristics are absented then an uncertainty can arise in the job classification. This uncertainty leads to the ambiguous in the class determination. The use of additional knowledge about the ranks, weights and computational history of job characteristics can significantly mitigate this uncertainty. The computational history is also used by the classifier-agent to evaluate its decisions. A set of characteristic functions is developed for recognition of job classes using different components of knowledge.

Figure 6 shows the results of classifying more than 80000 jobs of the real flow that were running on three clusters with different nodes. We compare the primary job class identification to the classification with using the additional knowledge. It is obvious that the class determination error in a percentage of the number of jobs with the uncertainty is significantly lower in the second case.

Evaluation of resource agent actions is regulated by the system of fines for deviation from the average resource load in their virtual community. Based on the analysis of the resource allocation results, agents can change the intent of executing different job classes.

Figure 7 shows the results of the average CPU load. These results show that the parameter adjustment of agents taking into account the specified job execution criteria significantly have been improved the processor load balancing.

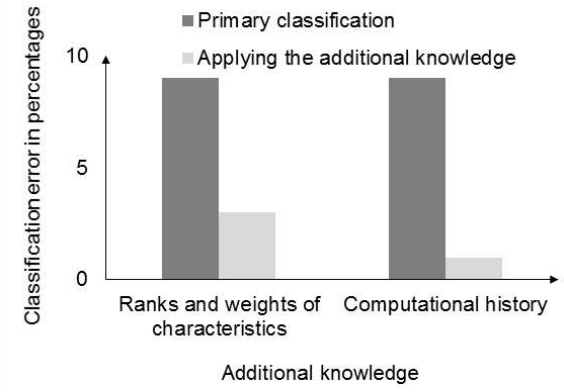


Figure 6. Job classification.

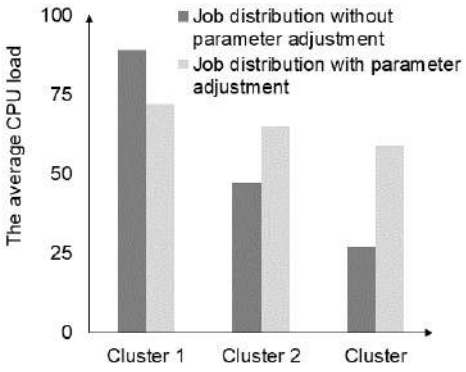


Figure 7. Load balancing.

4. Computational experiment

The example of solving a complex practical problem of determining the critical elements in technical infrastructure networks demonstrates the features and advantages of the proposed approach [20]. It consists of the study of failure sets, each of which represents a set of failed elements and has only one negative consequence of the impact on the system. The number n of simultaneously failed elements characterizes the failure set. The researcher selects the number n depending on the total number m of the system elements. For practical reasons, n has not exceeded 3 or 4 owing to the number of possible failures sets equal to $\frac{m!}{(m-n)!n!}$ is growing rapidly together with an increase of n.

To solve the problem, we developed a scalable application that supports the parameter sweep computing. The object of the study is the unified gas supply system of Russia. Its infrastructure contains 382 nodes, including 28 natural gas sources, 64 consumers, 24 underground gas storages and

266 key compressor stations, 486 arcs representing the main gas pipelines and outlets to distribution networks. We selected 415 arcs and 291 nodes (natural gas sources, underground gas storages, and key compressor stations) in this infrastructure. The selected 706 elements were calculated with $n=3$ and $n=4$. Thera is 58400320 and 10263856240 sets of failures for $n=3$ and $n=4$ correspondingly. The evaluated time to study all failure sets on one core of the Opteron 6276 Interlagos processor is 50 days if $n=3$ and more than 81 years if $n=4$. The evaluated time to study all failure sets on one core of the Intel Xeon E5-2695 processor is 14 days if $n=3$ and more than 32 years if $n=4$. These evaluations necessitate the use of high-performance computing.

To carry out an experiment, we create HDCE that includes the nodes of two segments of the HPC-cluster "Academician V. M. Matrosov", which is a part of the Irkutsk supercomputer center. The environment nodes have the following characteristics:

- Two processors AMD Opteron 6276 Interlagos (16 cores, 2.3 GHz, 64 GB RAM) in the first segment,
- Two processors Intel Xeon E5-2695 v4 Broadwell (18 cores, 2.1 GHz, 128 GB RAM) in the second segment.

Existing quotas of the cluster resource allocation do not allow to carry out a fully computational experiment for $n=4$ owing to restrictions of the maximum number of the resources allocated to the user in one segment. A user cannot use more than 15 and 20 nodes with the maximum of a job service time that equal 20 and 10 days in the first and second segments respectively. If a user uses the resources of one segment, it can run jobs on the second segment only if there are free slots in the job schedule. Owing to the aforementioned restrictions, we select the following scheme of the computational experiment:

- Allocating 20 nodes for a period of 15 days in the schedule of the second segment (the maximum allowed time of the resources use within the existed quota),
- Using free slots in the schedule of the first segment (within the existed quota).

In the first segment, the schedule slots are determined in the presence of jobs that are waiting for its launch for the execution. The agent-manager interacts with the PBS Torque job queue manager to identify nodes that currently have free cores. The number of slots corresponds to the number of such nodes. The number of free node cores is the width of the slot. The period until the end of the nearest job is the slot duration. It cannot exceed 1 day.

The agent-manager predicts the time when the nearest job will be completed. Predicting the time of the resources release is based on the maximum requested job run-time using the coefficient reflecting the real execution time of user jobs taking into account the computational history.

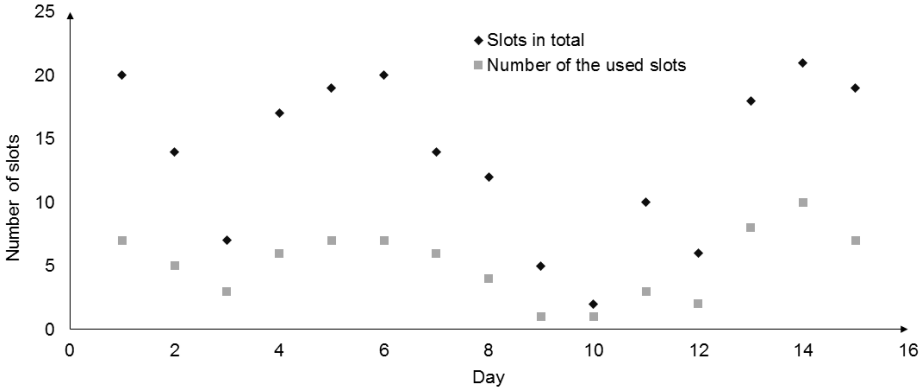


Figure 8. Free slots in the job schedule in the first segment.

Figure 8 shows the total number of slots in the job schedule in the first segment during the experiment and the number of slots used. We can see, that 62% of the slots were not used because of the possible computing inefficiency in them owing to high overheads for running and terminating VMs compared to the slot duration. Another reason not to use slots is the negative impact on the problem-solving processes of other users within the same node where the slot is.

Knowledge about the computing inefficiency in slots with specific characteristics is reflected in the job class description and applied by agents in the resource allocation process. The agent learning in the resource allocation process has increased the number of resources available for the experiment by 27 percent and completed the experiment in 15 days for $n = 4$.

5. Conclusions

We have considered a multi-agent system for distributed computing management in a heterogeneous distributed computing environment with virtualized resources. In contrast to the known multi-agent system, its functioning is based on the agents' complex use of the following knowledge:

- Computational knowledge of software modules for both the problem-solving in the subject problems and operating with system objects,
- Schematic knowledge of the modular structure of the model and algorithms,
- Productive knowledge to support decision-making on the selection of optimal algorithms depending on the environment state,
- Knowledge of the hardware and software infrastructure and administrative policies in its nodes.

This knowledge is represented in the form of a conceptual model that is a special case of the semantic network.

To support the agent learning, we have developed a new technology of the parameter adjustment of multi-agent algorithms for managing the heterogeneous distributed computing environment. An agent applies it to optimize the resource allocation when application user jobs are executed.

The proposed learning uses both the practical experience and skills of specialists in their subject domains (environment administrators, developers, and end-users of applications) and the knowledge elicited by agents. In contrast to the known approaches, selecting the control parameters of agent functioning algorithms within framework of the proposed agent learning are based on the integrated applying the job classification, matching classes to resources, meta-monitoring, and simulation modeling.

Thus, the developed technology allows performing detailed accounting the properties of distributed resources and characteristics of executed jobs, evaluating the current environment state, and predicting its evolution. Thereby, it provides a high degree of efficiency, reliability, and scalability of the computational process of solving large problems.

We have developed a scalable application to solve the important large-scale problem of studying development directions of the Russia energy sector from the standpoint of energy security in practice. We have also provided the intensive experiments to solve this problem based on the parameter sweep computing in HDCE. The experimental analysis confirms the effectiveness of multi-agent computing management and agent learning.

6. References

[1] Talia D 2012 Clouds meet agents: Toward intelligent cloud services *IEEE Internet Comput.* **16** 78-81
[2] Singh A, Juneja D and Malhotra M 2017 A novel agent based autonomous and service composition framework for cost optimization of resource provisioning in cloud computing *J. King Saud University Comput. Info. Sci.* **29** 19-28
[3] Shyam G K and Manvi S S 2015 *Proc. of the 2015 IEEE Int. Advance Computing Conf.* 458-463
[4] Talia D 2011 *Proc. of the 12th Workshop on Objects and Agents* **741** 2-6
[5] Stone P and Veloso M 2000 Multiagent systems: A survey from a machine learning perspective *Auton. Robots.* **8** 345-383
[6] Madni S H H, Latiff M S A and Coulibaly Y 2017 Recent advancements in resource allocation techniques for cloud computing environment: a systematic review *Cluster Comput.* **20** 2489-2533
[7] Hastie T, Tibshirani R and Friedman J 2009 *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Berlin, Heidelberg: Springer)

[8] Murphy K P 2012 *Machine Learning: a Probabilistic Perspective* (Cambridge: MIT Press)
[9] Jordan M I and Mitchell T M 2015 Machine learning: Trends, perspectives, and prospects *science* **349** 255-260
[10] Access mode: http://hpc.icc.ru
[11] Bychkov I, Oparin G, Tchernykh A, Feoktistov A, Bogdanova V and Gorsky S 2017 Conceptual model of problem-oriented heterogeneous distributed computing environment with multi-agent management *Procedia Comput. Sci.* **103** 162-167
[12] Oparin G A, Feoktistov A G and Feoktistov D G 1996 Combined abstract-program execution in the Saturn instrumental complex *Autom. Control Comp. S.* **30** 57-61
[13] Vickrey W 1961 Counterspeculation, auctions, and competitive sealed tenders *J. Finance* **16** 8-37
[14] Edelev A V and Sidorov I A 2017 Combinatorial modeling approach to find rational ways of energy development with regard to energy security requirements *Lecture Notes Comp. Sci.* **10187** 310-317
[15] Bumgardner V K 2016 *OpenStack in Action* (Manning Publications)
[16] Access mode: http://www.adaptivecomputing.com/products/open-source/torque/
[17] Feoktistov A, Tchernych A, Kostromin R and Gorsky S 2017 Knowledge Elicitation in Multi-Agent System for Distributed Computing Management *Proc. of the 40th Int. Convention on information and communication technology, electronics and microelectronics* 1350-1355
[18] Bychkov I V, Oparin G A, Feoktistov A G, Sidorov I A, Bogdanova V G and Gorsky S A 2016 Multiagent control of computational systems on the basis of meta-monitoring and imitational simulation *Optoelectron. Instrum. Data Process.* **52** 107-112
[19] Bychkov I, Oparin G, Novopashin A and Sidorov I 2015 Agent-based approach to monitoring and control of distributed computing environment *Lecture Notes Comp. Sci.* **9251** 253-257
[20] Jonsson H, Johansson J and Johansson H 2008 Identifying critical components in technical infrastructure networks *Proc. Inst. Mech. Eng. O J. Risk. Reliab.* **222** 235-243

Acknowledgments
The study is partially supported by Russian Foundation of Basic Research, projects no. 16-07-00931-a and no. 18-07-01224-a, and Presidium RAS, program no. 30, project “Methods, algorithms and tools for the decentralized group solving of problems in computing and control systems”.

The development of a virtual laboratory based on Unreal Engine 4

D A Sheverev[1] and I N Kozlova[1]

[1]Samara National Research University, Moskovskoye shosse 34A, Samara, Russia, 443086

Abstract. In our days we have the problem in the system of training highly skilled engineering staff– the impossibility of teaching students high-tech equipment because of its fragility and high cost. The solution of this problem is the creation of a virtual laboratory works, which represent a computer simulation of the real hardware – the simulator. The main advantages of virtual laboratories is the lack of significant costs for the buy, repair and maintenance of expensive equipment, increased visibility of the processes of the explored process, unlimited up building the virtual laboratory's functional and etc. At the moment the laboratory work is developed and its purpose is to research the hysteresis of linear deformations of piezoelectric elements and it is tested in real educational process. Further development of this work is to increase the number of training models, and of course as support for mobile devices.

1. Introduction

The main task of education was and always is the development of the student's personality in conjunction with the mastery of knowledge, skills and abilities that can be applied in practice [1-5]. In the system of higher professional education, especially for the training of engineering staff, it is extremely important to obtain practical skills in working with equipment [6,7]. And the main opportunity is to receive them in the process of training is the performance of laboratory work.

We have the first problem of the existing system of education. For training highly qualified staff for advanced areas of the economy, it is necessary to develop skills in high-tech equipment, for example, equipment is used in the production of integrated circuits. In view of the extremely high cost of such equipment and consumables to it, as well as the complexity of operation (requirement of special skills for its use, the possibility of failure due to incorrect use), doing laboratory work on a regular basis is extremely difficult.

Also, in our days, the activity of higher educational institutions is in priority of the organization of the scientific research, while the provision of the educational process by the laboratory fund is financed in a limited way. Timely updating of the laboratory base is a costly and extremely inertial measure, the effectiveness of which is not obvious. However, employers as one of the criteria put forward exactly the skills of working on modern equipment.

The solution of this problem may be the simulation of laboratory equipment. Simulation which is built on well-known mathematical models of various physical and chemical phenomena can be approximated to the actual experiment. And such approach will make it possible, without significant expenditures and significantly expand the range of practical issues to improve the knowledge of material on the topic of the discipline.

Another important aspect is the fact that virtual laboratory's works can help to solve another problem. We are talking about a rapid increase in the abstractness of the processes under investigation, which further increases the gap between theory and practice.

The use of virtual laboratories can help to relieve tension in this contradiction "theory-practice", it allows forming primary (basic) skills of working with research high-tech equipment without wear and tear of the latter, establishes and / or maintains the connection of theoretical knowledge with practical activity develops and consolidates interest to the chosen professional sphere.

The student's interest in professional activities, in turn, allows:

- to reveal the potential of students, because its promotes the development of the ability to set, formalize and solve problems, generate new ideas and create new knowledge;
- to improve the efficiency of the research work of the unit, because an enthusiastic student wants to expand the range of tasks from standard training to unique research.
- to show students a culture of production, because it opens the prospects for a systematic approach to solving the tasks.

2. What is a virtual laboratory and what is needed to create it?

The virtual laboratory is a simulation of working with the real equipment based on already known experience and mathematical models. It includes three-dimensional models of devices and appliances, environments (walls, doors, tables, chairs, etc.), which are recreated with high accuracy and realism. Also to enhance the realism all models are provided with detailed textures. Examples of such models you can see in Figure 1.

Figure 1. Voltmeter model in the left and real reference on right.

In the virtual laboratory, we use the first person's view of the abstract character to control the situation in the virtual laboratory, using the keyboard and mouse, similar to computer games.

In order to "revive" the virtual laboratory is necessary to write the interaction logic of user with it, as well as the objects of the laboratory with each other. For example, when approaching the door, it should open, etc.

Also in the virtual laboratory included the visualization of various effects which are studied in the work. For example, in Figure 2 we can see an example of a crystal lattice inside, by means of which we can see the crystal structures.

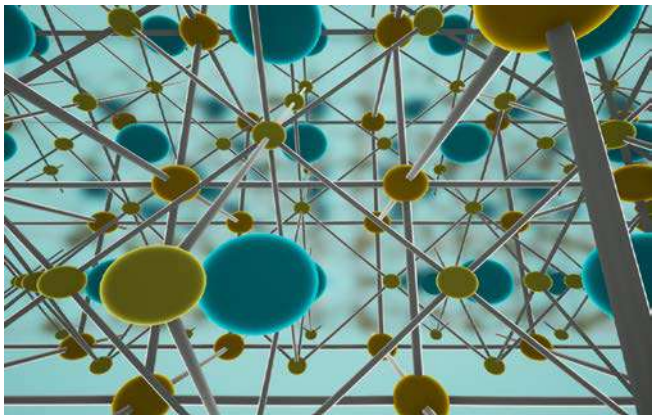


Figure 2. Crystal lattice visualization.

The important part is the methodological support is presented in the form of an interactive methodology that can be called at any time. The example of methodological guidance is shown in Figure 3.

Figure 3. Methodical instructions in the virtual laboratory.

3. Advantages of virtual laboratories

What is the advantage of virtual laboratory work in compare with the real ones? I just want to say that their use does not exclude the work with real equipment, but is in addition to it.

For the training of highly qualified staff for the advanced fields of the economy (micro and nanotechnology), it is necessary to develop skills in high-tech equipment (such as plasma-chemical etching facilities, electronic lithography, precise analytical equipment, etc.). In view of the extremely high cost of such equipment and consumables to it, and the complexity of operation (requiring special skills for its use, the possibility of failure due to incorrect use), laboratory work on a regular basis is extremely difficult. Virtual laboratory work can solve the problem. Virtual equipment does not break down, does not wear out and does not require materials for the work.

Another important advantage is an increase in the visibility of processes and phenomena. Progress now dictates the high speed of development of technologies in the direction of their complication and increase of abstractness of the studied processes. High abstractness can reduce the effectiveness of understanding the material in the preparation of highly skilled engineering staff. The use of virtual reality solves this problem, allowing visualizing the physical and chemical processes and visualizing them. The motivation of students to the learning process is increased, making it interesting, and also adding game moments.

Also, the advantages of virtual laboratory work is unlimited increase in functionality, the addition of new laboratory works and other methodological developments that require visualization without changing the hardware complex or with irrelevant additions.

4. Why Unreal Engine 4?

For realizing the software component of the project is used the game engine "Unreal Engine 4" from Epic Games. It is used for arranging objects at the level, writing logical and mathematical operations, adjusting lighting and materials (textures), and final assembly of the application. Advantages of its using are:

- providing an accurate lighting model makes it possible to achieve a realistic picture;
- visual programming system "Blueprint" accelerating the development process and eliminating necessity for knowledge of programming languages. We can see it in Figure 4;
- extensive set of the development tools.

5. What is finished?

At the moment, the development of the laboratory work "Investigation of the hysteresis of linear deformations of piezoceramic elements" in the discipline "Physical and chemical bases of micro- and nanotechnology" is being actively carried out. The basic principles of developing logical and mathematical algorithms, applications based on them and 3D visualization are mastered. By now, the performance of laboratory work is carried out using a keyboard and mouse and it is the management of an abstract character in the first person.

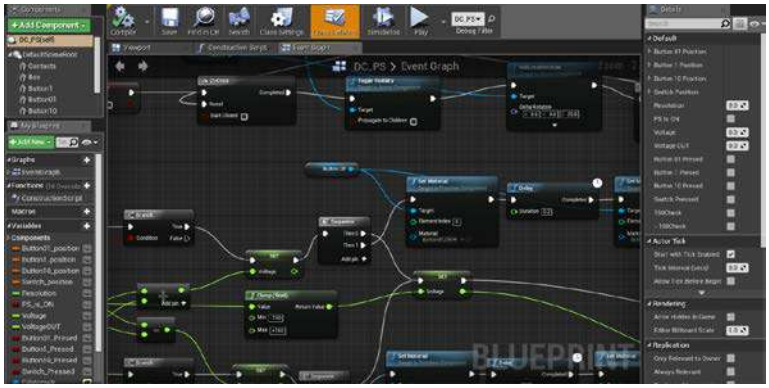


Figure 4. Visual programming system "Blueprint".

The room in which the laboratory work is divided into 3 sections:

1. The training rooms. At the beginning of the virtual laboratory work is necessary undergo training in the basics of control and interaction with objects;

2. Room of demonstrations. It contains extensive displays showing the monitoring process (and in particular the hysteresis of piezoceramics) in different levels, starting from a general view of the piezoelectric element, ending with the crystal lattice;

3. Room with laboratory equipment. It contains laboratory bench and rack with the samples is shown in Figure 5.

Figure 5. Laboratory bench and rack with the samples.

In any time the student can access the electronic methodical instructions and to study the theory or to ask a drone - assistant is presented in Figure 6.

Figure 6. Drone – assistant.

The application is at the testing stage in the conditions of a real educational process. Also work is being done on the sound accompaniment of laboratory work, which is an important factor that allows using an additional channel of human perception.

6. What is next?

The next stage in the development of this work is the introduction of virtual reality technologies, which will bring the level of perception of the material to a fundamentally new level. When using the virtual reality helmet and special controllers, the user creates a "full immersion" effect, which makes it easier to relate virtual equipment to the real one.

The virtual laboratory, created as a result of the completion of this project will consist of two parts: hardware and software.

The hardware consists of a high-performance computer, a virtual reality helmet, controllers and a motion capture sensor system. The most suitable is the VR helmet HTC Vive, because it has undeniable advantages over other models presented in the market and the capabilities that are required for the implementation of the project:

- Support for "tracking room" Lighthouse, which allows you to navigate while in virtual reality in a zone up to 4x5 meters (the other virtual reality glasses are designed for sitting). This allows you to create a greater immersion, thereby maximally approaching the sensation of being in a real laboratory;
- Using the Lighthouse system allows you to effectively track the head in six degrees of freedom. It minimizes the effect of "motion sickness" with a long time in the helmet, which is important;
- Presence in the complete set of controllers.
- Convenience in use with glasses for vision correction, because students wear them;

The general view of the laboratory is equipped with a helmet of virtual reality is presented in Figure 7 and Figure 8, where:

1. The HTC Vive virtual reality helmet, provides a highly detailed stereoscopic image for each eye;
2. The system of motion detectors Lighthouse. It projects the real movements in the movements of the managed virtual character;
3. High-performance laptop is the computational core;
4. Limiter of an accessible area for movement, designed to ensure the safety of others and the person using the helmet.
5. Controllers that allow you to interact with objects in virtual reality (buttons, toggle switches, knobs, curtains, touch panels, etc.)

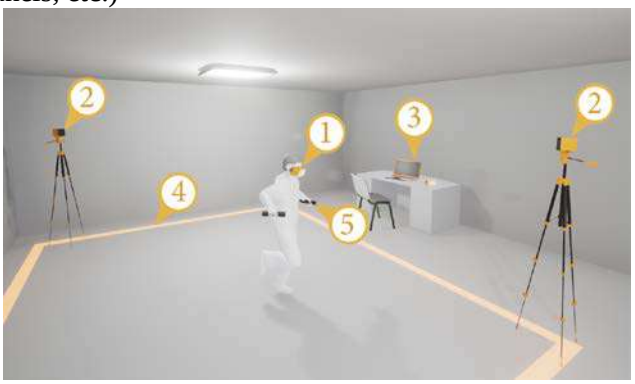


Figure 7. General view of the laboratory which is equipped with a virtual reality helmet.

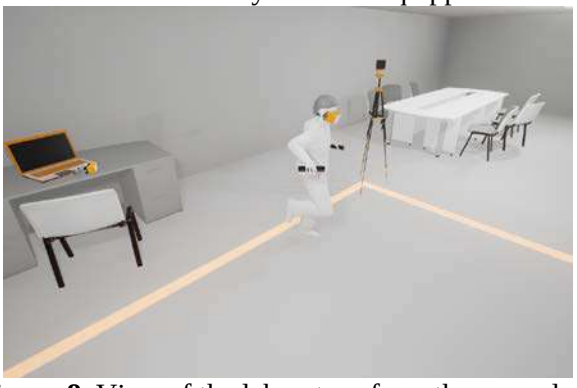


Figure 8. View of the laboratory from the second angle.

The program part is the cycle of laboratory works, supported by theoretical material, consists of 4-6 laboratory works on various topics of the discipline. The laboratory works are carried out for a small group of students from 6 to 8 people, which allows individualizing the training and controlling the assimilation of the material. Before the beginning of the cycle of laboratory work, it is necessary to provide safety training when person is in the helmet, as well as the basics of control and interaction in a special application that is part of the software package.

Each laboratory work is carried out in stages in the following order:

(If one of the labs has already been completed, then before proceeding to the next, you need to protect the previous one by answering a few test questions)

1. Equipping by a helmet of virtual reality, controllers and checking their functionality;
2. Studying of the theoretical foundations of laboratory work, which are presented in the form of an interactive presentation with sound;
3. Acquaintance with the basics of virtual equipment operation and the order of performing laboratory work;
4. Performing laboratory work;
5. Printing out the results (tables and graphs) obtained during the laboratory work and (or) sending them to the student's e-mail; All results obtained are stored in the database, which prevents the copying of the results of work;
6. Passage of the control test, which confirms the assimilation of the material studied in the course of laboratory work;
7. Drawing up a report on the work done.

During the performance of the laboratory work, the teacher has the opportunity to observe the video stream which are reproduced in the helmet of virtual reality with the aim of correcting the course of laboratory work.

7. Conclusions

The use of virtual reality opens a lot of new opportunities in learning and education that are too complex, time-consuming or expensive with traditional approaches, if not all at the same time. There are five main advantages of using VR technologies in education.

- Visibility. Using 3D-graphics, you can detail the chemical processes up to the atomic level. And nothing forbids to go deeper still further and to show how inside the atom the nuclear division is divided before the nuclear explosion. Virtual reality is capable of not only give information about the phenomenon itself, but also demonstrate it with any degree of detail.
- Safety. The work of potentially dangerous equipment and materials - you can immerse the viewer in any of these circumstances without the slightest threat to life
- Involve. Virtual reality allows you to change scenarios, influence to the course of an experiment or solve a mathematical problem in a playable and understandable form.
- Focusing. The virtual world, which will surround the viewer from all sides to all 360 degrees, will allow to concentrate entirely on the material and not be distracted by external stimuli.
- Virtual lessons. The view from the first person and sensation of the adventure In the painted world is one of the main features of virtual reality. It will allow you to conduct lessons entirely in virtual reality.

8. References

[1] Annetta L A 2008 Video Games to Education: Why They Should Be Used and How They Are Being Used *Theory Into Practice* **47(3)** 229-239

[2] Rozhkov O V, Piskunov D E, Nosov P A, Pavlov V Yu, Khorokhorov A M and Shirankov A F 2018 Bauman MSTU scientific school “Zoom lens design”: features of theory and practice *Computer Optics* **42(1)** 72-83 DOI: 10.18287/2412-6179-2018-42-1-72-83

[3] Kolomiets E I 2017 International Master's degree program "High-Performance and distributed information processing systems" *Procedia Engineering* **201** 832-848

[4] Kolomiets E I 2017 Analisys of the international Master's degree program "Mathematical Modeling and Information Technologies in Photonics" *Procedia Engineering* **201** 849-862
[5] Kazanskiy N L 2017 Efficiency of deep integration between a research university and an academic institute *Procedia Engineering* **201** 817-831
[6] *UnrealEngine 4 Documentation* (Access mode: https://docs.unrealengine.com) (28.05.2017)
[7] Sheverev D A and Sanoyan A G 2016 Laboratory stand for the research of hysteresis of linear deformations of piezoelectric elements for the purpose of precision mixing of objects in manometer range *Actual Problems of Radio Electronics and Telecommunication* 139-141

Simulation modelling of the transmission system of the telemetric information on the status of the on-board aircraft status

M B Guzairov[1], A I Frid[1], A M Vulfin[1] and V V Berkholts[1]

[1]USATU - Ufa State Aviation University, Karla Marksa street 12, Ufa, Russia, 450000

Abstract. Modern aviation systems are complex hierarchical computing systems, coupled with a powerful periphery (digital board). The flight is performed with constant communication with ground control points and requires exchange of large amounts of data. The reliability of the systems providing flight can be violated due to equipment failures, human factor and external unauthorized exposure. The paper discusses the issues of guaranteeing the reliability of telemetry information transmission systems on the state of the aircraft and its systems based on artificial intelligence methods in terms of ensuring information security. The structure of the imitation stand is proposed to test the system for transmitting telemetric information about the state of on board airborne systems to the ground. The purpose of such system is to improve the efficiency of modelling and analysis of the system's security.

1. The first section in your paper

Provision of dependability telemetric data transmission systems (TMI) is a primary goal of ensuring the efficient functioning of the aircraft.

The application of technologies for monitoring the status of critical elements of the aircraft's design in real time will allow a full (operational and intelligent) analysis of the operational capability of the aircraft systems (LA), crew conditions and control of its actions in the ground control center.

For example, the Bombardier jetliner was demonstrated at the Paris air show. It has a motor that is equipped with 5000 sensors, which generate up to 10 GB of data per second. A single twin-engine aircraft with an average 12-hour flight time can create up to 844 TB of data [1]. The arising malfunctions and failures of the on-board equipment can be diagnosed on the basis of the processed telemetric information. This allows experts of ground technical services to prepare for operational repair even before landing the aircraft. Accumulated and processed TMI will allow specialists to provide reasonable support for decision-making in the event of technical failure of aircraft units or modules.

The possibility of transmitting TMIs on the actual state of individual modules to the manufacturer of aviation equipment units will make it possible to increase the operational efficiency of the aircraft in its normal state and in the event of failures, as well as when investigating incidents. For example, a study of ground-to-board communication systems showed that the ACARS system, despite its versatility and ubiquitous use, is vulnerable, and if it is hacked with ADS-B, an attacker can gain access to the flight control system, download flight plans and detailed commands [2].

The purpose of this study is to increase the security of the TMI transmission system on the status of individual elements of on-board systems in an automatic mode based on the use of modern technologies of protection and processing of TMI using an imitation stand.

To achieve this goal, the following tasks are formulated:

1. Develop a structural scheme for the collection, transmission and reception of telemetric information on the status of individual elements of the aircraft systems of the aircraft;

2. Develop the architecture of the system for analyzing the security of the TMI transmission system on the status of individual elements of the aircraft system.

Formatting Structural scheme for the collection, transmission and reception of TMI on the status of individual elements of the aircraft system.

2. Structural scheme for the collection, transmission and reception of TMI on the status of individual elements of the aircraft system

Automated information system (AIS) of ground maintenance services is a set of software and hardware needed to receive, store and process information on technological parameters of complex technical device (CTD). A typical block diagram of the existing TMI processing system of the manufacturer is shown in Figure 1. An analysis of existing approaches to solve the problem of ensuring the reliability of such systems is considered in the work [3].

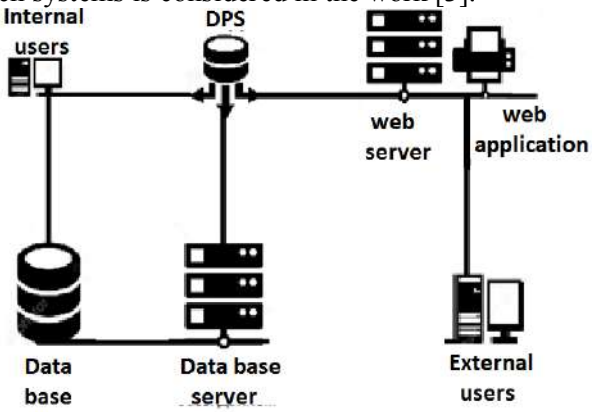


Figure 1. Structural diagram of the TMI processing system.

AIS solves the main tasks associated with receiving information about the state of a complex technical device. Data is planned to be received in three different ways (figure 2):

1. Directly from the system. STD is a component of this system;

2. By reading the event log from the CTD module sensors. The status of the modules is read and stored during the previous period of operation.

3. Entering events into the database manually. The operator processes the information and enters information through the WEB application.

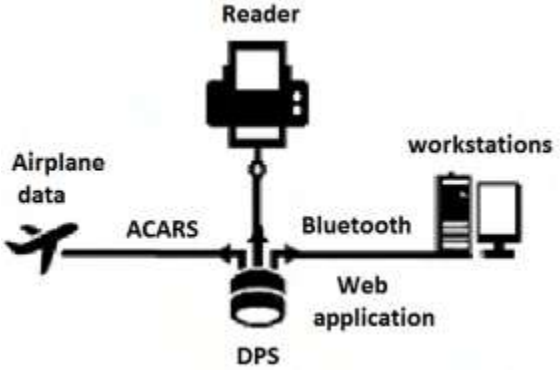
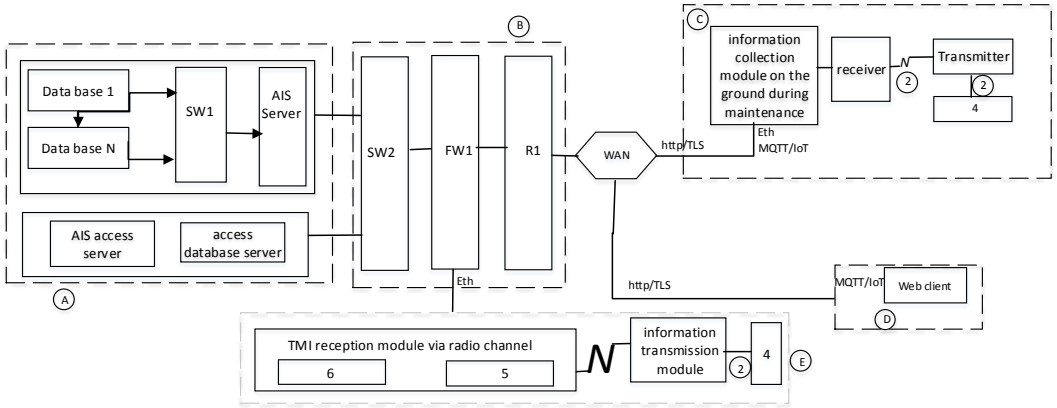


Figure 2. Methods of obtaining TMI.

Data is entered into the database via web-application in the second and third cases. WEB-application is an insulating layer between external networks and internal structure of AIS, since access from an external network is one of the most vulnerable places of the system. The structural scheme of the collection and transmission of TMI is presented in Figure 3.



A-simulation model of the server part of the processing system
B- module for simulating the network subsystem
C-module for information collection at stations. Services
D - module for simulating access via web-client
E- data transfer via radio channel

1. Database replication
SW1- communication equipment of the AIS module
SW2- communication equipment of AIS network
FW1-Firewall
R1-Router NAT
2. RS-485 / Modbus
3. NIST Technical Series Publication (Bluetooth)
4. TMI collection assembly from aircraft

Figure 3. Structural scheme for the collection, transmission and reception of telemetric information on the status of individual elements of the aircraft system.

In the work of the authors [3], the problem of providing secure access to the existing database containing critical information about the parameters of the life cycle of complex technical products (CTD) using the example of gas turbine engine control system was considered.

Thus, increasing the security of access to a database containing critical information about an exploited product is based on the development of a secure WEB application architecture. It serves as an isolating layer for external AIS clients, which allows for the provision of the ability to transmit and analyze at ground service stations and to provide remote access to the required data. Prevention of vulnerabilities in the WEB-application is carried out by implementing measures to develop secure software, established by GOST R ISO / IEC 12207 [4].

Modeling security threats and identifying vectors of possible attacks, as well as their analysis, allowed to formulate countermeasures for each of the vectors at different architectural levels of the WEB application.

The growth in the volumes of telemetric information forces the aviation industry to consider new approaches to collecting and analyzing a large amount of data on the condition of aircraft nodes and elements. The concept of an industrial Internet of things is developing actively (IIoT). This is a deployed network consisting of a large number of devices equipped with a set of sensors that exchange data with each other through low-power and short wireless connections. The first step is to collect data from the sensors. One of the most promising solutions is a protocol with low power consumption and low transmission radius IEEE 802.15.4 IEEE 802.15.4e [5]. A short range is sufficient to transmit data within the ground service station. The IEEE 802.15.4 and IEEE 802.15.4e protocols and their architecture layers are subject to IETF standards [6].

The question of analyzing the security of the system for collecting, transmitting and receiving telemetric information on the status of individual elements of aircraft systems during data transmission through the first two channels remains an open question. The violation of the information security of the TMI collection and transmission system under consideration can be caused by a variety of

different reasons: vulnerabilities in operating systems and server and client applications; incorrect configuration of hardware and software; access control settings errors, and so on.

External and internal violators can implement various strategies to attack the system. The attacker can use combinations of available vulnerabilities and network configuration and security policy (PB) deficiencies. These strategies can be designed to access a database that contains critical information about STI. Strategies can also include multi-step chains of attacking actions to implement security threats. This arises the task of verifying the provision of the necessary level of security The level of security of the system is determined by a set of parameters for the configuration of the data transmission network, security policy and protection mechanisms.

The initial data for analyzing the security of the TMI collection and transmission system are its specifications and security policies. The task of developing the architecture of the security analysis system (SAS) is due to the high complexity of the TMI collection and transmission system. This increases the number of vulnerabilities and potential errors of the software and hardware components. The SAS of the TMI collection and transmission system should detect errors in the transmission network configuration, possible routes of attacking actions of various categories of violators (for implementing various security threats), identify critical network resources and ensure the selection of adequate security policy threats. The core of the CAS are algorithms for assessing the level of security of the system, it is based on the construction of possible actions of violators in the form of attack graphs. The properties of the attack graph are checked and the system security metrics are calculated [7].

The system for analyzing the security of the transmission of telemetric information includes the implementation of a set of models:

• behavior model of an attacker;

• simulation of the graph of attacking actions aimed at the implementation of information security threats taking into account attacker's qualification;

• vulnerability model of the system;

• calculation of security metrics system and assessment of the overall level of security.

The security analysis system is a software package implemented as a hierarchical set of client-server applications. AIS enterprise interaction is carried out by means of a Web application with remote TMI transmission modules: a secure channel, single-board microcomputers with a set of required wired and wireless network interfaces, intermediate switching equipment, a client and server computer. Input data for security analysis are:

• specification of the configuration of the information system for data collection and transmission of TMI;

• specification of the planned security policy;

• hardware and software vulnerabilities;

• model of the intruder;

• requirements to the security of the information system.

The output of the CAS are:

• identified vulnerabilities of the system;

• graphs of possible attacks; • security metrics;

• evaluation of the overall level of security of the system and its components. The obtained results provide the development of well-founded recommendations on elimination of identified bottlenecks and strengthening of the system's security.

The simulation stand is a software implementation of the system for analyzing the protection of the transmission of telemetric information on the state of airborne systems. The generalized architecture of the proposed CAS is shown in Figure 4.

The module for configuring the security analysis model allows the security specialist to manage the operation of all system components, specify the input parameters of the analysis process, the requirements for the level of security, and review security analysis reports and receive recommendations for enhancing the security of the TMI transmission system. The management and administration module provides access to external vulnerability databases and configuration settings

for other modules (modeling of IS nodes, simulating an attacker and its actions, as well as updating the DB and the Knowledge Base (KB)).

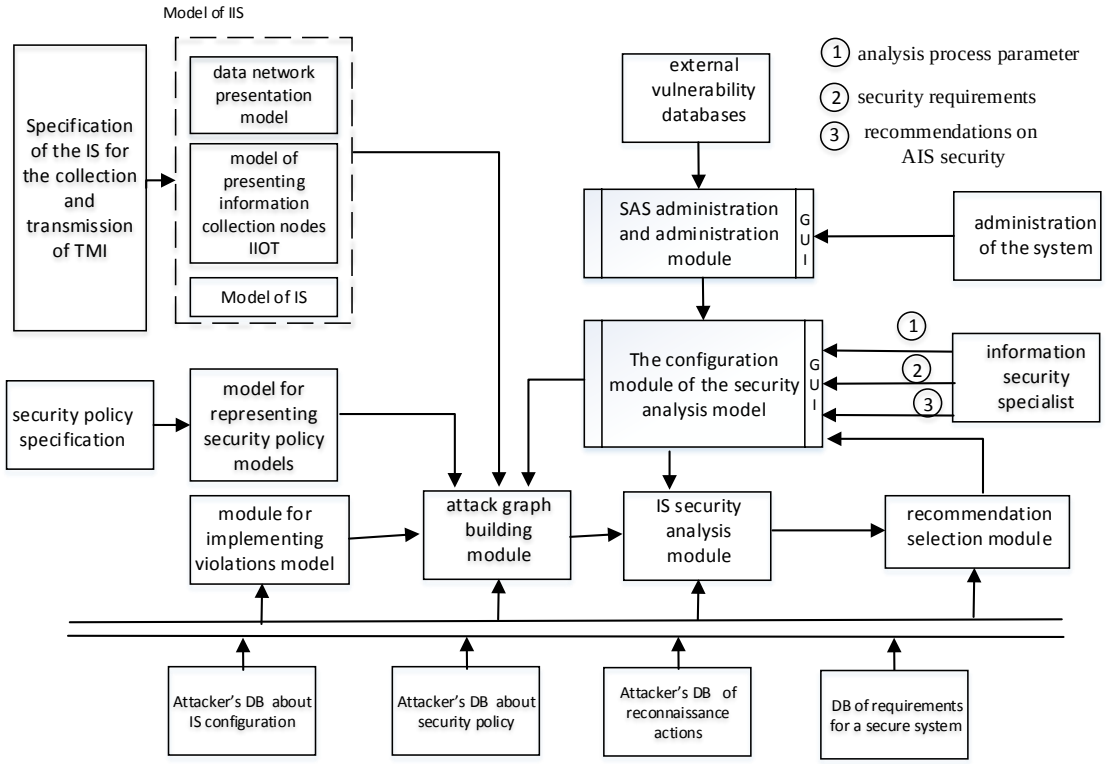


Figure 4. Architectures of the TMI transmission security analysis system on the status of individual elements of the aircraft systems.

The module for forming the internal representation of the analyzed network and the security policy converts the data about the analyzed network and the implemented security policy into an internal representation.

The security systems and security policies introduced in the SAS should describe the components of the protected system (network) with the necessary degree of detail - the software used (in the form of software product names and versions) must be specified The data warehouse consists of the following databases and knowledge bases:

(1) knowledge base on the structure and configuration of the TMI transmission network and the implemented security policy;

(2) an attack database;

The knowledge base group consists of four bases:

(1) knowledge base on the configuration of the analyzed system;

(2) knowledge base on the security policy (SP) implemented in the information system;

(3) knowledge base of the intruder about the configuration of the analyzed network;

(4) knowledge base of the offender on the network security policy implemented in the network.

The network configuration database and the security policy database contain information about the architecture and specific parameters of the network and the rules that describe its operation. The action database group consists of the following bases:

(1) DB of actions using vulnerabilities. It is built on the basis of an external database of vulnerabilities.

(2) The intelligence database contains actions aimed at remote retrieval of information about the host or network.

The security requirements database contains predefined sets of values of security metrics, each of which corresponds to the requirements for systems of a certain security class [8].

The database update module and the knowledge base downloads open vulnerability databases [9,10] and updates the database of attacking actions.

The attack graph module simulates possible actions of the intruder in the analyzed network, using information about available actions of various types (attacking, reconnaissance), network onfiguration and the used security policy.

The module for the implementation of the intruder's model provides a definition of the initial position of the offender, the level of knowledge and skills, the primary knowledge of the analyzed computer network. The level of knowledge and skills determines the set of actions used by the offender.

The security analysis module forms a set of composite objects of the attack graph. It also calculates security metrics related to these objects, estimates the overall level of security of the computer network. Algorithms for the intellectual analysis of large volumes of transmitted TMI using Big Data technologies use preprocessing methods with highlighting useful information about the hidden actions of the attacker. The selection in the sequence of values of the observed parameters of TMI, features, repetitive processes, hidden anomalies and patterns in automatic mode underlies the procedure for detecting an attack of an attacker and replenishing the corresponding BR. Technologies of intellectual analysis of large volumes of accumulated data have proved themselves in the tasks of technical diagnostics of complex control objects [11,12] and detection of fraudulent activities in information systems. The results of the analysis are processed in the advisory module based on the application of methods and algorithms for data mining. The analysis results are compared with the requirements determined by the information security specialist, and recommendations are made to improve the overall level of security of computer networks.

Ensuring the reliability of systems for transmitting telemetric information about the state of aircraft using artificial intelligence methods.

Assurance (dependability) allows to solve complex problems of ensuring trouble-free operation, fault tolerance, availability, security, serviceability, observability of the TMI transmission system from the aircraft. The SAS being developed allows one to estimate one of the parameters of the overall integral index of the system's reliability. It is showed in figure 5.

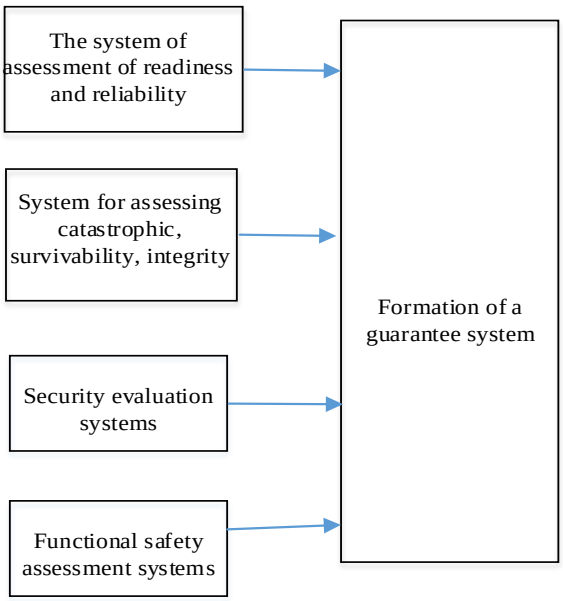


Figure 5. Architectures of the TMI transmission security system on the status of the individual elements of the aircraft systems.

Therefore, it is necessary to build a hierarchy of models that allow to comprehensively assess various aspects of the TMI transmission system and develop a methodology for assessing the overall integral index of the system's overall security.

3. Conclusion

The structure of an imitation stand for testing a system for transmitting telemetric information about the state of airborne systems to a ground station for processing telemetric information and control is proposed to improve the efficiency of modeling and analysis of system security. A block diagram is proposed for the collection, transmission and reception of telemetric information on the status of individual elements of the aircraft systems of the aircraft. The architecture of the security analysis system of the TMI transmission system on the status of individual elements of the aircraft system is developed. To assess the overall integral index of the system's overall security, it is necessary to build a hierarchy of models that allow for a comprehensive assessment of various aspects of the TMI transmission system and develop a methodology

4. References

[1] *Internet of Aircraft Things: An Industry Set To Be Transformed* (Access mode: http://aviationweek.com/connected-aerospace/internet-aircraft-things-industry-set-betransformed) (02.01.2016)
[2] *Aircraft Hacking Practical Aero Series* (Access mode: https://conference.hitb.org/hitbsecconf2013ams/materials/D1T1%20-20Hugo%20Teso%20-20Aircraft%20Hacking%20-%20Practical%20Aero%20Series.pdf) (23.02.2016)
[3] Guzairov M B, Frid A I, Vulfin A M, Berkholts V V, Zakharov D Ju and Mironov K V 2017 The architecture of the web application for protected access to the informational system of processing critically important information *Proceedings of the 17st Computer Science and Information Technologies Conference* **1** 16-19
[4] GOST R ISO / IEC 12207-2010 *Information technology. System and software engineering* (Software life cycle processes)
[5] Palattella M, Accettura N, Vilajosana X, Watteyne T, Grieco L, Boggia G and Dohler M 2013 Standardized Protocol Stack for the Internet of (Important) Things *IEEE Communications Surveys & Tutorials* **15(3)** 1389-1406
[6] *IEEE Standard for Local and Metropolitan Area Networks: Overview and Architecture* (Access mode: http://ieeexplore.ieee.org/document/6847097) (23.04.2012)
[7] Kotenko I V, Stepashkin M V and Bogdanov V S 2006 Analysis of the security of computer networks on various stages of their life cycle *Priborostroenie* **49(5)** 3-8
[8] The Russia FSTEC order dated March 14 2013 N 31 *On the requirements approval for the information security that is not a state secret contained in government information systems* (Access mode: http://fstec.ru) (14.04.2013)
[9] *NVD. National Vulnerability Database* (Access mode: http://nvd.nist.gov)
[10] *OSVDB* (Access mode: https://blog.osvdb.org/)
[11] Vulfin A M and Frid A I 2011 Neural network model for the analysis of time series process in the methodology Data Mining *Information-Control Systems* **5** 31-38

Acknowledgments
This article is supported by RFBR grant № 17-07-00351.

Cloud technologies in the problems of mathematical analysis of cardiological information

E Zimina[1,2], M Novopashin[2] and A Shmid[1, 2]

[1]National Research University Higher School of Economics, Myasnitskaya str. 20, Moscow, Russia, 101000
[2]EC-Leasing, Varshavskoe shosse 125/1, Moscow, Russia, 117587

Abstract. The article includes the observation of the cloud services and technologies usage. The article contains a review of mathematical analysis of cardiac information using cloud technology, which produces storage, analysis and forecasting on the basis of owned data. In addition, the authors consider the possibility of integrating cloud technologies with external systems. The massive use of mobile devices for the removal of the electrocardiogram (ECG) leads to a quantitative increase of the patients number available for ECG investigation. Thus, there are new opportunities to research the oscillatory processes of long-term dynamics of the individual state of the cardiovascular system (CVS) of any patient. The article demonstrates new opportunities the long-term continuous monitoring of the patients CVS, which allows identifying regularities of the dynamics of the CVS. Also this article comprises the observation of the existence of an adequate model of CVS as a distributed nonlinear self-oscillating system of the class model returns the Fermi-Pasta-Ulam (FPU).

1. Introduction

According to the forecast of the McKinsey Institute for the 2025 year in the field of digital economy the most significant areas of technologies application are Mobile Internet, Automation of knowledge work, Internet of Things and Cloud. The total market volume will be around 30 trillion dollars. For medical use of these technologies the estimated volume is approximately 30%. For comparison, the share of oil and gas sector of the world economy is estimated at only $ 1.5 trillion.

Also, according to the Google Academy (Google Scholar) there is a sharp increase in the number of publications with reference to Big Data among all publications on research in the field of health [1].

The aim of the project is to create a centralized system of cardiological information processing of the population outside the medical institutions, providing personalized monitoring, analysis and forecasting of the development of the patient's condition in real time before his first appearance in the hospital.

This project is one possible response to the global health challenge: on the one hand, the growing global contradiction between the existing medical capabilities of the world health system and, on the other hand, the shortcomings in the concept and technologies of its organization, leading to a rapid increase in the lack of all types of resources in its application [2].

The primary component of the system of cardiological information processing in this project is the cloud, which combines in a single automated system and mobile devices, a group of trained computers and medical personnel who make decisions on scenarios of interaction with patients.

A system of processing of cardiac information is being developed, which provides personalized monitoring, analysis and forecasting of the patient's condition in real time.

2. The concept of FPU auto-return

We introduce the notion of FPU auto-return at the mathematical level. The phenomenon of canonical Fermi-Pasta-Ulam (FPU) auto-return in passive systems was the first described as a result of numerical study of differential solutions – difference equations describing a chain of nonlinearly coupled vibrators [3].

In these model chains the phenomenon of attenuation has not been submitted. To describe real dynamic processes, such as cardiac activity, a more appropriate model is needed that can describe returns in autonomously functioning or self-oscillating systems.

A significant contribution in this direction was the work of American researchers Zabuski and Kruskal [4], who proposed to describe the FPU auto-return in the framework of the Korteweg de Vries equation (KdV) with periodically changing boundary and initial conditions.

Using the results of this work it is possible to simplify the solution of the KdV equations in the form knoid waves, replacing them with the solutions of the equation of Van der Pol, close to harmonic, periodic and low frequency changes in boundary and initial conditions on the relaxation solution of the equation of Van der Pol [5].

Meanwhile, if there is the theorem application of the replacing wave links with lagging ones possibility, it is possible to present a mathematical model of the electrical activity of the heart (ECG) in the form of FPU auto-return, described in the framework of Van der Pol equations with delay in the following form [6]:

$$\frac{d^2M_1}{dt^2} - a_1(1 - Y_1)\frac{dM_1}{dt} + \omega^2(1 + a_1M_2)M_1 = c_1\frac{d^2M_2}{dt^2} + d_1F_1 + d_2F_2 \quad (1)$$

$$b_1Y_1 + T_1\frac{dY_1}{dt} = M_1^2 \quad (2)$$

$$\frac{d^2M_2}{dt^2} - a_2(1 - Y_2)\frac{dM_2}{dt} + {\omega_2}^2(1 + a_2M_1)M_2 = c_2\frac{d^2M_1}{dt^2} + d_1F_1 + d_2F_2 \quad (3)$$

$$b_2Y_2 + T_2\frac{dY_2}{dt} = M_2^2 \quad (4)$$

where M_1 - the value proportional to a dynamic electrical potential of the myocardium, M_2 - the value proportional to the dynamic electric potential is a small locality of the myocytes at the surface of the myocardium, Y_1 - the value proportional to the delay in the propagation of electrical impulse in the myocardium, Y_2 - the value proportional to the delay in the propagation of electrical impulse in the locality of myocytes, b_1 - the value proportional to the surface areas of the myocardium, b_2 - the value proportional to the surface area of the locality of myocytes, T_1 - the value proportional to the time of myocardial contraction, T_2 - the value proportional to the period of oscillation in the locality of myocytes, F_1 - the function of the resonant effect of the external environment on the heart rate (about 1 Hz), F_2 - the function of the resonant effect of the external environment on the heart at a high frequency of about 20 Hz, c_1, c_2, d_1, d_2 - constants.

The properties of ECG spectrum of healthy people and patients with ischemic heart disease found in the course of the study will allow to formulate the problem of mathematical modeling of the heart electrical dynamics with the description of the phenomena observed in the experiment [7] so that the heart can be represented as an open dynamic structure with many self-oscillating degrees of freedom, as shown in formulas 1-4.

The approaches proposed by Norbert Wiener, Ginsburg and Landau have been successful in modelling the dynamic states of cardiac electrical activity such as normal functioning, ischemic disease and myocardial infarction.

3. The Cloud Technologies

The cloud allows us to receive, store and process the electrocardiogram, photoplethysmogram (pulse wave), echocardiograms, seismocardiogram for further calculations. A first lead is required for in-depth analysis, but any additional leads only clarify further analysis.

The cloud is replenished by means of application programming interface (API) and adapters of two types: intended for self-unloading of data and intended for reading ECG formats received from specialized devices.

The cloud is designed for storage and processing using Big Data technologies information obtained from specialized agencies and from the mobile user's device.

The results of the analysis are also available via the API and special adapters to integrate with external systems.

Calculations of indicators are made on the basis of frequency and energy characteristics of the received signals. There is a selection of typical and transient frequency states for each patient individually, as well as for target samples of signals, divided by age, sex and the presence of pathologies in the patient.

The analysis determines the proximity of the calculated characteristics of the patient to one of the certain frequency states according to the data of one examination. Tracks the progress indicators of the patient status between the model states of the patient and the target samples to determine trends in the development of pathologies, effects of medications, post-operative rehabilitation.

The biological age of the patient's cardiovascular system is determined as a separate type of trajectory. There is an opportunity to detect acute the patient's condition that requires professional help.

By lengthening the trajectory of the calculated characteristics of the patient's condition between the typical patient state and the target samples, prediction of the future states of the patient for the next few days is constructed.

When a new cardiogram is entered into the database, two types of treatment are initiated: cardiogram treatment for the current patient and processing of an updated set of all cardiograms from the database.

4. Work with cardiograms

4.1. Cardiograms handling

The investigation of cardiograms is necessary to obtain calculations on the basis of which the analysis of the patient's condition and prognosis of his condition will be made. Processing of a new incoming cardiogram is performed in the sequential execution of several types of tasks:

- preparation of cardiogram for further calculations;
- calculation of the Fourier spectrum and its characteristics (harmonic and energy) for the unfiltered signal, signal with the corrected contour and signal with the corrected contour and free from defects;
- calculation of spectral shapes for the signal with the corrected contour and free from defects;
- recalculation of cardiogram clusters in the entire series of the patient cardiogram with different clustering parameters;
- calculation of the patient trend indicators (biological age).

The updated set of all cardiograms is processed according to the schedule and for different target groups by age, sex, pathology information for various clustering parameters.

Age groups are as follows: all ages and 5 separate age groups from young to old. Gender groups are as follows: men and women.

The following groups are distinguished by pathologies:

- spaced sampling cardiogram diagnosis of ischemia (confirmed and not confirmed);
- marked sample of cardiogram diagnosis of tuberculosis (confirmed and not confirmed);
- marked selection of cardiograms diagnosed diabetes (confirmed and not confirmed).

If there is not new ECG since the last clustering, the launch of the new clustering is not performed.

In the processing of individual prior ECG to performing any calculations and research, it is necessary to correct the contour, i.e. to remove the drift contours, and clear it from the strong defects, because it spoils the frequency response of the required ECG is needed for further analysis [8].

Drift contours removed with the help of the two median method, from original signal elementwise subtracted calculated by the two median contour.

4.2. Cardiogram preparation for calculation of indicators

Before processing a separate cardiogram, it is necessary to remove the drift of the contour, and clean it from strong defects.

Drift contours removed two median method, from original signal elementwise subtracted calculated by the two median contour. Defect filter is used. RR intervals are cut out from the signal, in which the signal amplitude spread is much higher than the value of the median amplitude spread over a sliding window [9]. Parts of the signal are glued together at the vertices of R-peaks, between which the RR-interval was cut. The edge parts before the first detected R-peak and after the last detected R-peak are also discarded so that all filtered signals start at R-peak and end at R-peak.

The percentage of the dropped part of the signal in relation to the entire length of the signal is stored separately.

5. Calculation of Fourier spectrum and its characteristics

5.1. Methods of research

Patients with permanent atrial fibrillation were excluded from the study, as well as the following patients were excluded: with complete atrioventricular blockade, complete blockade of the left leg of the bundle branch block, the technical causes of false negative or false positive detection in the identification of myocardial ischemia (recommendations for Holter), life-threatening arrhythmia, acute cerebrovascular accident, tumors, severe anemia, thyroid disease in the stage decompensation, lack of informed consent of the patient, failure to conduct stress tests and with questionable results of stress tests and Holter monitoring (HM) ECG [10].

The diagnosis was based on clinical, anamnetic, instrumental methods of research. Verification of the diagnosis was carried out using samples with physical activity (veloergometry or stress echocardiography). All patients underwent HM ECG.

5.2. Calculation of Fourier spectrum

For an unfiltered signal, a signal with a corrected isolation, and a signal with a corrected isolation and filtered defects, a fast Fourier transform algorithm is used. The amplitude-frequency spectrum is calculated from the complex Fourier transform coefficients. On the spectrum, harmonics are defined – characteristic “humps” in the region of certain frequencies. The amplitude and frequency of each harmonic are determined inside the window calculated based on the base frequency. The base frequency is the heart rate displayed in Hertz, i.e. beats per second. It is calculated as follows:

f_{base}=1/median({RRs}), where {RRs} is the set of RR interval lengths.

For each harmonic, the frequency f_i, amplitude A_i and energy index – the spectral density E_i are calculated.

6. Clustering of cardiograms

6.1. Clustering of cardiogram samples

Clustering is performed for a variety of cardiograms, isolated by certain characteristics. For the tasks clustering is performed for all individual patient's cardiograms with different clustering parameters according to the selected forms of the spectrum.

The distance value from the shape of the spectrum to the cluster is considered as the distance from the shape of the spectrum to the cluster middle line, which is always stored as an intermediate result. Merging of clusters is accompanied by a recalculation of the values of the average lines in accordance with the number of cardiogram in each of the merged clusters. Also clustering can be carried out by other methods: on the basis of the form obtained by another algorithm, based on the average cardiocycle and on the basis of amplitudes or spectral densities of harmonics.

6.2. The calculation of indicators of trends in the patient

To determine the trend of changes in indicators of the patient's condition between the two states, the most popular clusters are the largest number of cardiograms.

At this stage, a cluster of patients cardiograms of the younger age group and a cluster of patients cardiograms of the older age group, regardless of gender, are distinguished. Next, we will call it the “young” cluster and the “old” cluster, respectively.

The trend index t_(ECG-M) is the ratio of the distance from the current form of the patient's spectrum to the middle line of the “young” cluster to the distance to the middle line of the “old” cluster.

For each trend indicator, the confidence factor w_(ECG-M) calculated from the triangle rule is considered. The most indicative are the patient's conditions located close to the line between the "young" cluster and the "old" cluster.

The trust factor is set to 0 to 1. The higher it is, the more significant the indicator is for the research.

Based on a set of pairs {t_(ECG-M),w_(ECG-M)}, the trend indicators for the last day, last 4 days, week (last 7 days), month (last 30 days), and all time of patient observation are determined. They are calculated by the formula of mathematical expectation. On request, the trend indicator can be calculated for an arbitrary period of time.

6 values are calculated for each cardiogram, which are available for external systems.

6.3. Building ECG clusters principles

Cluster refers to a set of one or more ECG and associated with this set, the arithmetic mean of the spectrum, calculated automatically as a result of the ECG ordered classification process.

Comparison of ECG Fourier spectrum was implemented on the basis of the FPU auto-return phenomenon with the visual perception of the spectrum. For this purpose, the concept of spectrum shape was introduced. Spectrum shape-smoothed curve based on the median of the upper 30% of the values in the window 2 Hz, taken in the range from 0 to 40 Hz. Window width 2 Hz is taken specially to eliminate the influence of pronounced harmonics for ECG with a heart rate less than 120 beats/min. Form of the spectrum is used further in the algorithm for determining regions of ischemia and clustering of ECG in the form of the spectrum.

Combining similar self-oscillating processes of the heart into clusters on the basis of belonging to the diagnosis as a result give a finite number of subsets. However, it follows from the FPU model that there may be such variants of behavior of the heart that they will not “fall” into any of the clusters, which correspond to the known diagnoses. Such “outliers” reflect the process of the heart model transition from one state (cluster) to another [11].

For clustering, the form is approximated in steps of 0.1 Hz to 400 points and normalized to the arithmetic mean for comparison universality.

ECG clustering in the form of spectrum is performed by the method of extraction of connected components. The distance between the forms of the compared ECG spectrum is taken Euclidean. In the process of clustering the original fully connected graph, in case when it links of greater length than the assigned value are dropped, is scattered into a group of subgraphs, which form the desired clusters as a result.

In order to simplify the algorithm clusters are formed automatically based on ECG classification as they are received (the so-called online method). In this case, operations are meant to supplement the cluster with a new form of ECG, it creates a new cluster and merges clusters based on the distance between them.

It is practically obtained that there is always a cluster with a predominant number of ECG trapped in it. Presumably, the ECG of this cluster characterize the normal state of the patient. ECG, trapped in the other clusters, characterize to some extent the abnormal condition of the patient, caused by both external stress and pathologies. Often there are clusters consisting of a single ECG, their appearance gives an indication of the intermittency of complex FPU auto-return patient (when between occurrences of a certain state there is a change of more states than from repeated).

For such reference points several clusters of healthy people, patients with ischemia were selected. In the Figure 1 there is a typical Fourier spectrum of the ECG of a healthy person. In the Figure 2 the typical Fourier spectrum of the ECG of a patient with a developed ischemic process is given.

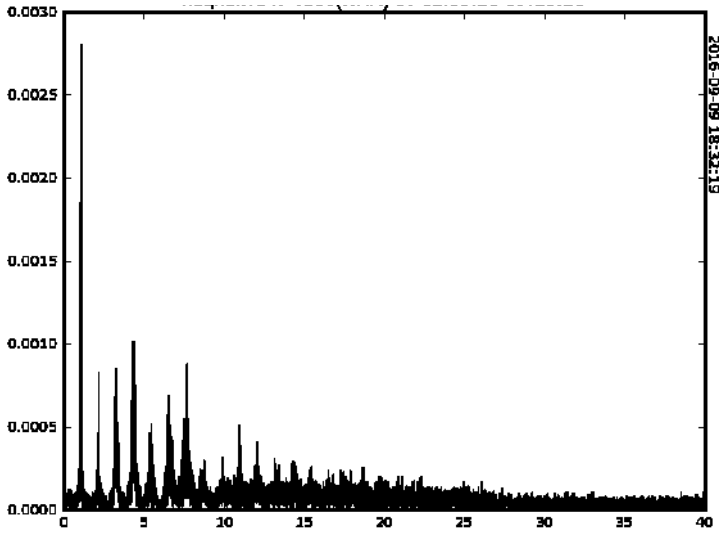


Figure 1. Typical Fourier spectrum of an ECG of a healthy heart. Horiz. Axis-frequency in Hz, vertical. Axis - potential in mV.

The cluster with the largest set of ECG is called the base cluster. For all other clusters, the measure of cluster deviation from the base as Euclidean distance between the corresponding averaged forms is introduced.

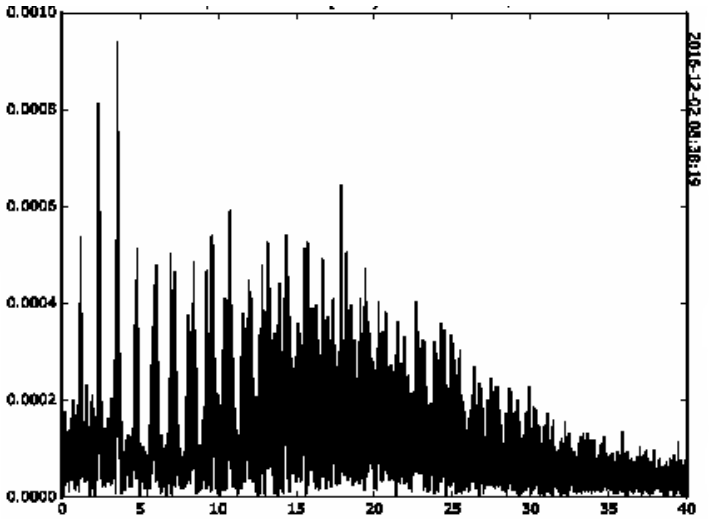


Figure 2. Typical Fourier spectrum of an ECG of a patient with advanced ischemic process. Horiz. Axis-frequency in Hz, vertical. Axis - potential in mV.

The developed ischemic process in accordance with the hypothesis is the formation of an autonomous dynamic system in the myocardium, which has oscillatory parameters different from the parameters of oscillations in the normal tissue of the myocardium. In particular, the main frequencies of the ECG spectrum of a normally functioning heart are in the range of 1-12 Hz, while the spectrum of the ischemic process is in the range of 14-20 Hz and can be dynamically changed.

On the basis of the characteristic features of the individual clusters obtained and the data on patient's pathologies, clusters similar in characteristics to the shape of the spectrum are identified and displayed in separate groups on pathologies. Thus, the clustering method can be used to diagnose ischemia, determine the cardiological status of healthy people.

7. Conclusion

The article discusses the possibility of storage and processing of cardiac information for mathematical analysis using cloud technologies.

During the study, the possibilities of analysis and monitoring of medical data, as well as forecasting in the tasks of mathematical analysis of medical data using cloud technologies were reviewed.

The mathematical modeling of the heart by using the decomposition of Fermi-Pasta-Ulan is carried out. To describe the mathematical model of the heart, we present a system of connected cells of myocytes, representing separate self-oscillating degrees of freedom, described by a system of coupled nonlinear differential equations of Van der Pol.

Cluster analysis is carried out on the basis of search of similar forms of Fourier spectrum obtained by modeling the heart FPU.

The main result of this work is the implementation of processing and storage of cardiograms (any number of leads), photoplethysmograms (pulse wave), echocardiograms, seismocardiograms for further calculations in the cloud and the organization of access to the results.

Thus, the use of the cluster approach can be useful for the diagnosis of various degrees of ischemia, determining the cardiological status of healthy people with an assessment of its age-related changes. In addition, the proposed cluster method will assess the effect of the therapy in dynamics, and change it in case of inefficiency.

8. References

[1] Andreu J, Poon C, Merrifield R D and Yang G-Z 2015 Big data for health care *IEEE Jour. of Biomedical and Health Inform.* **19** 1193-208
[2] Braunwald E 1997 *Heart disease: textbook of cardiovascular medicine* (Philadelphia: W. B. Saunders Co.) p 108
[3] Fermi E, Pasta J and Ulam S 1955 *Studies of nonlinear problems* (Collected Papers of E. Fermi University of Chicago Press) p 978
[4] Zabusky N and Kruskal M 1965 Interaction of solitons in a collisionless plasma and the recurrence of initial states *Phys. Rev. Lett.* **15** 240-243
[5] Van der Pol B 1926 On relaxation-oscillations *The London, Edinburgh and Dublin Phil. Mag. & J. of Sci.* **2** 978-992
[6] Novopashin M, Shmid A and Berezin A 2017 Fermi-Pasta-Ulam auto recurrence in the description of the electrical activity of the heart *J. Medical Hypothesis* **101** 12-16
[7] Berezin A 2004 Resonant interaction between relapses Fermi-pasta-Ulam *Bulletin of the Physics Institute FIAN* **3** 13
[8] Ezhkov A, Isakov R, Sountsova O and Sadovskiy S 2016 Connection between heart control and lungs pathology based on the analysis of an electrographic signal recorded by the cardiac monitor cardioqvark *Research Journal of Pharmaceutical, Biological and Chemical Sciences* **7** 3137-3142
[9] Rubanik V 1969 *Oscillations of quasilinear systems with a time lag* (Moscow: Science) p 288
[10] Moyer V 2002 Screening for coronary heart disease with electrocardiography: recommendation of the us preventive services task force *Annals of Internal Medicine* 157
[11] Novopashin M, Schmid A and Berezin A 2017 Forrester's concept in modeling heart dynamics in IOSR *Journal of Computer Engineering IOSR* **3** 113-121
[12] Ikezi H 1978 Solutions in plazma *Solitons in Action* 163-168
[13] Lichtenberg A and Lieberman M 1992 Regular and chaotic dynamics *Applied Mathematical Sciences* **38**
[14] Longren K 1978 Solitons in nonlinear electric lines *Solitons in Action* 138-162

Application of time series analysis for structural and parametric identification of fuzzy cognitive models

R A Isaev[1] and A G Podvesovskii[1]

[1]Bryansk State Technical University, 50 let Oktyabrya 7, Bryansk, Russia, 241035

Abstract. The article deals with problems of structural and parametric identification of fuzzy cognitive models on the basis of statistical data analysis. The feasibility of application of time series analysis for solving these problems is justified. The Granger causality test is proposed for structural identification. An approach for parametric identification based on distributed-lag time series model is also proposed. The results of experimental verification of the described approaches are presented.

1. Introduction

A cognitive approach is one of approaches to the study of semi-structured systems, which is widely used at the present time. According to the definition given in [1], this approach focuses on the development of formal models and methods supporting the intelligent problem-solving process as they include human cognitive capabilities (perception, conception, cognition, understanding, explanation) in solving management problems. Structure and target modeling and simulation modeling methods based on cognitive approach are commonly subsumed under the umbrella term “cognitive modeling”. In general terms, cognitive modeling refers to the study of structure, functioning and development of a system by analyzing its cognitive model. The cognitive model is based on a cognitive map, which reflects researcher's subjective notion (individual or collective) of the system as a number of semantic categories (known as factors or concepts) and a set of cause-and-effect relationships between them.

A cognitive model is an effective tool for exploratory and estimative analysis of the situation. It does not give an opportunity to obtain accurate quantitative characteristics of the system under study, but it allows to assess trends related to its functioning and development, and to identify the key factors influencing these processes. Thus, we can search, generate and develop effective solutions for system management, as well as identify risks and develop strategies to reduce them.

Cognitive modeling starts with creating a cognitive map of the system under study on the basis of information received from experts. The next step includes direct simulation. Its main objectives are forming and testing hypotheses for the structure of the system under study, which can explain its behavior, also developing strategies for various situations in order to reach the specified target states.

Tasks solved by means of cognitive modeling can be divided into two groups:

1. Tasks of structure and target analysis:
   - finding the key factors influencing the targets;
   - identification of contradictions between the targets;
   - identification of feedback loops.
2. Tasks of dynamic analysis (scenario simulation):
   - self-development (“what if we do nothing”);

- managed development:
  - direct task (“what if”);
  - inverse task (“how to”).

Thus, the scenario simulation allows prediction of the simulated system states under different control actions and search for alternative control solutions bringing the system to the target state.

Mathematical apparatus most commonly used to represent cognitive models and underlying methods for their analysis is fuzzy logic. As a result, there appeared a whole class of cognitive models based on different types of fuzzy cognitive maps (FCM). A detailed overview of such models can be found, for instance, in monograph [2]. One of FCM varieties, well-proven in practical analyzing and modeling of semi-structured organizational, social and economic systems are Sylov’s FCMs. They were firstly proposed in [3] and represent the development of signed cognitive maps [4].

2. Formal definition and structure of Sylov’s fuzzy cognitive map

As previously mentioned, a cognitive model is based on formalization of cause-and-effect relationships which occur between factors characterizing the system under study. The result of the formalization represents the system in the form of a cause-and-effect network, termed a cognitive map and having the following form:

$$G = < E, W >,$$

where $E = \{e_1, e_2, \ldots, e_K\}$ is a set of factors (also called concepts), W is a binary relation on the set E, which specifies a set of cause-and-effect relationships between its elements.

Concepts can specify both relative (qualitative) characteristics of the system under study, such as popularity, social tension, and absolute, measurable values: population size, cost, etc. Moreover, every concept e_i is connected with a state variable v_i, which specifies the value of the corresponding index at a particular instant. State variables can possess values expressed on a certain scale, within the established limits. Value $v_i(t)$ of state variable at instant t is called the state of concept e_i at the given instant. Thus, the state of the simulated system at any given instant is described by the state of all concepts included in its cognitive map.

Concepts e_i and e_j are considered to be connected by relation W (designated as $(e_i , e_j) \in W$ or $e_i W e_j$) if changing the state of concept e_i (cause) results in changing the state of concept e_j (effect). In this case, we say that concept e_i influences concept e_j. Besides, if the value increase of the concept-cause state variable leads to the value increase of the concept-effect state variable, then the influence is considered positive (“strengthening”); if to the decrease – then negative (“inhibition”). Therefore, the relation W can be represented as a union of two disjoint subsets $W = W^+ \bigcup W^-$, where W^+ is a set of positive relationships and W^- is a set of negative relationships.

Fuzzy cognitive model is based on the assumption that the influence between concepts may vary in intensity; whereas, intensity may be constant or variable in time. Taking into account this assumption, W is set as a fuzzy relation, however, its setting depends on the adopted approach to formalization of cause-and-effect relationships. A cognitive map with fuzzy relation W is termed a fuzzy cognitive map.

Sylov’s fuzzy cognitive map represents FCM, characterized by the following features:

1. State variables of concepts can possess values on the interval [0, 1].
2. Influence intensity is considered constant, so relation W is specified as a set of numbers w_{ij}, characterizing the direction and degree of influence intensity (weight) between concepts e_i and e_j:

$$w_{ij} = w(e_i , e_j),$$

where w is a normalized index of influence intensity (characteristic function of the relation W) with the following properties:

a) $-1 \le w_{ij} \le 1$;
b) $w_{ij} = 0$, if e_j does not depend on e_i (no influence);
c) $w_{ij} = 1$ if positive influence of e_i on e_j is maximum, i.e. when any changes in the system related to concept e_j are univocally determined by the actions associated with concept e_i;

d) $w_{ij} = -1$ if negative influence is maximum, i.e. when any changes related to concept e_j are uniquely constrained by the actions associated with concept e_i;
e) w_{ij} possesses the value from the interval (–1, 1), when there is an intermediate degree of positive or negative influence.

Clearly, FCM of this structure can be graphically represented as a weighted directed graph, which points correspond to elements of set E (concepts) and arcs correspond to nonzero elements of relation W (cause-and-effect relationships). Each arc has a weight which is specified by the corresponding value w_{ij}. In this case, relation W can be represented as a matrix of dimension $n \times n$ (where n is the number of concepts in the system), which can be considered as the graph adjacency matrix and is termed a cognitive matrix.

3. The present state of research in the field of fuzzy cognitive models identification

In the course of building a FCM we can distinguish two stages:

- structural identification, which implies determining a set of concepts E and a crisp relation W over this set, i.e. verification of connections between concepts;
- parametric identification, which implies transition from the crisp relation W to a fuzzy one, i.e. determination of connection weights (influence intensity) between concepts.

Experts are the key source of information at both stages of building a map. In particular, at the structural identification stage, a list of concepts is formed by an expert (or a group of experts). Then, connections between concepts are added to the cognitive model on the basis of the expert notion of the simulated situation.

Expert methods are also most commonly used at the parametric identification stage. They can be direct and indirect. The direct methods imply immediate (explicit) weighing by an expert. The indirect methods are used to minimize the impact of subjectivity in the process of weighing, and they are based on breaking the general task of determining weights into a number of simpler sub-tasks. Saati's pairwise comparison method, Yager's level set method and Churchman-Ackoff method are examples of indirect methods. Description of these methods, as applicable to defining FCM weights, can be found in monograph [5] (section 3.2).

As previously noted, some concepts can set quantitative parameters of the system under study and consequently have numerical state variables. Provided that there is statistical information about the values of these variables, it can be used to identify connection weights between such concepts instead of expert assessment. Thus, statistical methods can be used to identify FCM parameters alongside with the expert methods. The possibility of using this or that method is determined by the nature of the available statistical information [6]. For instance, if statistical data about concepts are represented in the form of spatial sampling, a linear regression model can be applied to identify the sign and intensity of influence between the concepts.

Method based on a pair linear regression model was proposed in monograph [7] (sections 4.2-4.3). With its help, it's possible to identify the sign and intensity of influence between two concepts. Generalization variations of this method based on multiple regression analysis are of special interest and allow identifying parameters of influence of several concepts on a concept.

Attempts to apply correlation and multiple regression analysis to build fuzzy cognitive models of social and economic systems were undertaken in [8, 9]. Nevertheless, the results given in these papers can't be regarded satisfactory for several reasons. First, the authors use the correlation and regression analysis to reveal the very existence of cause-and-effect relationships between concepts and to ascertain the direction of these relationships. However, it is well known that high value of a coefficient of correlation between factors as well as reliability of the regression model built on their basis are not sufficient to draw a conclusion of a cause-and-effect relationship between these factors. Moreover, it is impossible to define accurately the direction of this relationship through the specified methods. Second, the authors propose to use regression equation coefficient values as connection weights between concepts. Yet, the weights obtained can acquire values outside the range [-1, 1], which contradicts the formal definition of Sylov's FCMs. Finally, these papers don't examine the

multicollinearity problem, i.e. a high degree of intercorrelation among explanatory variables in regression models. This inevitably leads to abundance of redundant connections in the FCMs obtained.

The approach described in [10] is also based on the multiple regression analysis but is free of the enumerated drawbacks. Nevertheless, there are still a number of current problems connected with the identification of fuzzy cognitive models on the basis of statistical data.

First of all, it should be noted that since the modeled systems are dynamic (i.e. their state changes in the course of time), statistical information about them is likely to be represented in the form of time series in most cases. In this context, the regression analysis is not viable because one of its conditions is representation of data in the form of spatial sampling. Considering this, methods should be developed to identify connection weights between concepts on the basis of time series analysis. This issue was partially addressed in [11], but no approach to identification based directly on time series analysis was proposed in it – the primary focus was on correlation analysis.

Development of FCM structural identification methods based on statistical data is another advanced problem. As has been mentioned above, methods based on spatial sampling analysis do not enable us to establish cause-and-effect relationships between concepts, while time series analysis methods provide us with such an opportunity.

Further, we describe approaches to solving the specified problems.

4. Application of Granger causality test to structural identification

As noted above, the decision of adding a connection between two concepts to a cognitive model is made on the basis of expert notion of the system modeled. Even if there are statistical data about the concepts in the form of spatial sampling, it is impossible to establish either a cause-and-effect relationship between concepts or its direction. In this case, statistical information is used only for the sign and influence intensity identification if there is such influence in the expert's opinion.

If there are statistical data about some concepts *X* and *Y* in the form of time series, then Granger causality test can be used for verification of feasibility and viability of adding a connection between them [12].

The idea of the test is as follows: if *X* influences *Y*, then a change in *X* must precede a change in *Y*, but not vice versa. Moreover, the following two conditions must be met:

- *X* must contribute significantly to the prediction of *Y*;
- *Y* must not contribute significantly to the prediction of *X*.

If every variable contributes significantly to the prediction of the other one, there are two options possible:

- there is a two-way causality between them;
- there is a third variable influencing both.

Two null hypotheses are sequentially checked in Granger test:

- "*X* does not Granger-cause *Y*";
- "*Y* does not Granger-cause *X*".

To test these hypotheses two regressions are built; in each of them, the regressand is one of the variables tested for causality, and the regressors are the lags of both variables:

$$y_t = a_0 + a_1 y_{t-1} + \ldots + a_p y_{t-p} + b_1 x_{t-1} + \ldots + b_p x_{t-p} + \varepsilon_t; \quad (1)$$

$$x_t = c_0 + c_1 x_{t-1} + \ldots + c_p x_{t-p} + d_1 y_{t-1} + \ldots + d_p y_{t-p} + u_t. \quad (2)$$

For each regression (1) and (2) the null hypothesis is that the coefficients of the lagged values of the second variable simultaneously equal zero:

$$H_0^1 : b_1 = \ldots = b_p = 0; \quad (3)$$

$$H_0^2 : d_1 = \ldots = d_p = 0. \quad (4)$$

To test hypotheses (3) and (4), an F-test should be performed. To arrive at the conclusion that *X* influences *Y*, it is necessary that the first hypothesis be rejected and the second one be accepted (generally, at significance value 0.05).

The number of lagged variables included in the regressions influences the result of the test. Therefore the test is recommended to be performed at a variety of *p*.

Granger causality between variables doesn't guarantee a cause-and-effect relationship between them but implies the *possibility* of such relationship. Meanwhile, no Granger causality guarantees absence of such relationship. In other words, Granger causality between time series is an *essential but not sufficient* condition for a cause-and-effect relationship between the corresponding concepts. Thus, the final decision whether to add a connection to the cognitive model remains with the expert.

Suppose a cognitive model includes concepts X_1, K, X_n and there are data about them in the form of time series $x_t^1,\ldots,x_t^n$. Then, at the structural identification stage, it is required to establish between which pairs of concepts connections should be added. For this purpose, the described test should be conducted between the series corresponding to each pair of concepts.

Moreover, it is important to consider the fact that influence between concepts can be realized not only directly but also transitively. Granger causality will be also detected in the latter case but with a longer lag (i.e. at larger *p*) than under the direct influence. Since existence of a relationship between concepts in a cognitive model means that a change of cause concept state leads to a change of effect concept state in one step, the question of adding a connection should be raised only if causality is detected between time series at minimum value of *p*.

5. Parametric identification based on distributed-lag time series model

At the parametric identification stage, it is necessary to determine signs and weights of all connections between concepts added to the model following the results of the structural identification stage.

Choosing a time series model for weighting FCM connections, it is required to correlate it with the impulse process model which is supposed to be used for the dynamic analysis of the map under study. Research on various impulse process models can be found in one of the authors' previous papers [13].

Further description of the proposed approach is given by the example of the most common impulse process model, namely an additive model with absolute changes. Within this model, a change of concept *Y* state in a given step *t* is supposed to be determined (except control and external actions) by absolute changes of influencing concept states in a previous step ($t - 1$). Meanwhile, previous state changes of concept *Y* itself are not taken into account. Considering this, for the simplest case (such as when influence on concept *Y* is realized from one concept *X*) we obtain the expression:

$$\Delta y_t = a_1 \Delta x_{t-1} + \varepsilon_t, \tag{5}$$

where $\Delta y_t = y_t - y_{t-1}$; $\Delta x_{t-1} = x_{t-1} - x_{t-2}$; a_1 is a coefficient, determining the intensity of influence transmission from *X* to *Y*; ε_t is an error.

It is easily seen that model (5) is equivalent to the following:

$$y_t = a_0 + a_1 x_{t-1} + u_t, \tag{6}$$

where a_0 is an absolute term; u_t is an error.

The described model (6) is a special case of a distributed-lag time series model (*DL*), which, in its turn, can be represented as a special case of autoregressive distributed lag model (*ADL*) [12].

This model can be assessed by the least squares method (in fact it is a model of concept *Y* values regression on previous values of the influencing concept *X*), so that we can obtain the desired value of a_1 (regression coefficient).

The described model is naturally generalized in case of several influencing concepts. In this case, we receive a multiple regression model. Application of this model to the parametric identification was viewed in detail by the authors in [10].

The same principle of transition from regression coefficients to connection weights should be applied as in the regression analysis. The principle was also described in detail in [10].

6. Experimental validation of the proposed approaches to structural and parametric identification

Suppose concepts *X* and *Y* are added to a fuzzy cognitive model, and there is statistical information about them in the form of time series. Time series corresponding to the concepts are illustrated by graphs in figure 1.

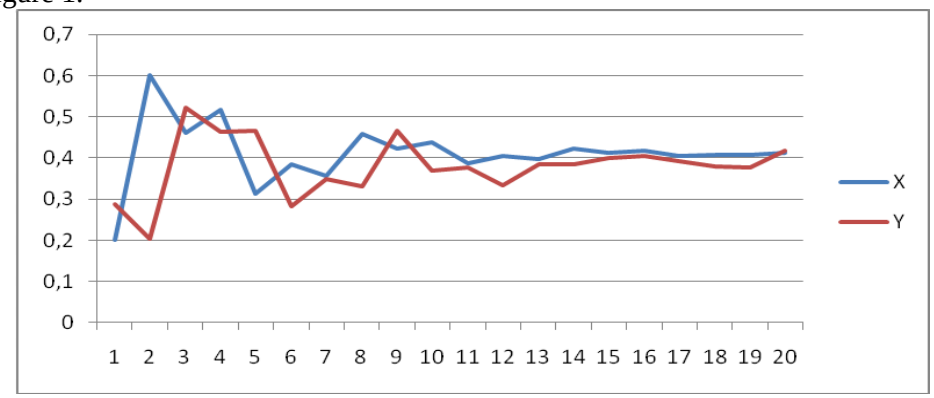


Figure 1. Time series corresponding to concepts *X* and *Y*.

According to the described approach to the structural identification and taking $p = 1$, let us plot the following regressions:

$$y_t = a_0 + a_1 y_{t-1} + b_1 x_{t-1} + \varepsilon_t; \quad (7)$$

$$x_t = c_0 + c_1 x_{t-1} + d_1 y_{t-1} + u_t. \quad (8)$$

By assessing models (7) and (8) by the least squares method, we obtain:

$$y_t = 0.011 + 0.053 y_{t-1} + 0.858 x_{t-1}; \quad (9)$$

$$x_t = 0.61 - 0.31 x_{t-1} - 0.159 y_{t-1}. \quad (10)$$

Next, for each regression (9) and (10), it is necessary to test the hypothesis of the coefficient equal to zero with the second variable lagged, that is $H_0^1 : b_1 = 0$ and $H_0^2 : d_1 = 0$. F-test reveals that the first hypothesis is rejected at significance value 0.05 and the second one is accepted. Thus, *X* Granger causes *Y* if $p = 1$.

Detection of Granger causality between these concepts at minimum value of *p* is a reason to question the expert whether to add a connection directed from concept *X* to concept *Y* to the fuzzy cognitive model.

Suppose the expert decides to add such connection to the model. In this case, the same statistical data about the concepts used at the previous stage can be applied to identify the sign and influence intensity between the concepts. For this, let us develop a model using the existing time series

$$y_t = a_0 + a_1 x_{t-1} + u_t, \quad (11)$$

Having assessed the model by the least squares method, we obtain $a_1 = 0.857$. Determination coefficient R^2 of model (11) equals 0.9, which indicates its acceptable quality. Besides, the obtained value a_1 is significant according to Student's t-test. With the help of transformations described in [10], let us pass from the regression coefficient obtained to the influence intensity of concept *X* on concept *Y*. As a result, we obtain $w_{XY} = 0.88$ (with normalizing function parameter $b = 3$).

7. Conclusion

The paper deals with problems of structural and parametric identification of fuzzy cognitive models and the existing problem-solving techniques: expert and statistical. The viability of new approaches to solving these problems on the basis of time series analysis is substantiated. An approach to solving the problem of structural identification is proposed, based on Granger causality test. Also a possible approach to parametric identification on the basis of distributed-lag time series model is studied. The results of experimental validation of the proposed approaches are presented.

Let us consider possible directions for further research which are of major interest.

First, one of the features of data analysis represented in the form of time series is that, besides the measurements themselves (levels of a series), there is information about real time moments at which these measurements were obtained. Knowing the difference in time between two successive levels of time series (and consequently knowing the time of influence spreading between directly connected concepts), we can approximately correlate model time steps with the real time of the simulated system and thus improve accuracy and concreteness of prediction resulting from the cognitive model dynamic (scenario) analysis.

Second, it is worthwhile developing the existing impulse process models towards taking account of different rates of influences: between different pairs of concepts, influences can spread at a variable speed (at a varying number of simulation steps). Meanwhile, rates of influence spread between pairs of concepts are determined on the basis of time series pairs corresponding to them.

Finally, use of statistical data for fuzzy cognitive model identification in the form of time series provides the model verification with new opportunities. In case of comprehensive statistical data, model identification can be performed using only a part of them; the rest can be used for its verification. Degree of the model adequacy will be determined by the accuracy of recalling in the process of dynamic simulation the data by which it was trained and by the efficiency of predicting data, which were not accounted for while training.

8. References

[1] Avdeeva Z K, Kovriga S V and Makarenko D I 2007 Cognitive modeling approach to control of semi-structured systems (situations) *Managing Large Systems* **16** 26-39

[2] Borisov V V, Kruglov V V and Fedulov A S 2012 *Fuzzy Models and Network*s (Moscow: “Goryachaya Liniya – Telekom” Publisher)

[3] Sylov V B 1995 *Strategic Decision Making in Fuzzy Environment* (Moscow: “INPRO-RES” Publisher)

[4] Roberts F S 1976 *Discrete Mathematical Models with Application to Social, Biological and Environmental Problems* (Englewood Cliffs, N.J.: Prentice-Hall)

[5] Erokhin D V, Lagerev D G, Laricheva E A and Podvesovskii A G 2010 *Strategic Enterprise Innovation Managemnet: Monograph* (Bryansk: Bryansk State Technical University Press)

[6] Denisova A Y and Sergeev V V 2015 Impulse response identification for remote sensing images using gis data *Computer Optics* **39(4)** 557-563 DOI: 10.18287/0134-2452-2015-39-4-557-563

[7] Averchenkov V I, Kozhukhar V M, Podvesovskii A G and Sazonova A S 2010 *Monitoring and Prediction of Regional Demand for Highest Scientific Degree Specialists: Monograph* (Bryansk: Bryansk State Technical University Press)

[8] Makarova E A, Gabdullina E R, Zakieva E Sh and Valiullina K M 2016 Algorithms for intelligent analysis of life quality in the domain of public health on a regional level *Proc. of the 6th Int. Conf. on Information Technologies for Intelligent Decision Making Support* **2** 222-228

[9] Makarova E A, Zakieva E Sh, Gabdullina E R and Makhmutova A E 2016 Knowledge generation algorithms for construction of cognitive model of life quality in the domain of high education on a regional level *Proc. of the 6th Int. Conf. on Information Technologies for Intelligent Decision Making Support* **2** 54-59

[10] Podvesovskii A G and Isaev R A 2016 Application of multiple regression analysis for parametric identification of fuzzy cognitive models *Proc. of the 6th Int. Conf. on Information Technologies for Intelligent Decision Making Support* **2** 28-33

[11] Rogachyov A F and Melikhova E V 2014 Problems of statistical estimation of cognitive map characteristics on the basis of correlation analysis *Proc. of the Int. Conf. “Physico-Mathematical Sciences: Theory and Practice”* 55-62

[12] Magnus Ya R, Katyshev P K and Persetskii A A 2004 *Econometrics: Basic Course* (Moscow: “Delo” Publisher)

[13] Isaev R A and Podvesovskii A G 2017 Generalized model of pulse process for dynamic analysis of Sylov’s fuzzy cognitive maps *CEUR Workshop Proceedings* **1904** 57-63

The application of Kohonen Self-Organizing Maps for the classification of the electronic components and reliability improvement of onboard equipment

R O Mishanov[1]

[1]Samara National Research University, Moskovskoe shosse 34, Samara, Russia, 443086

Abstract. The technique of the electronic products classification into classes of the acceptable and potentially unacceptable instances using Kohonen Self-Organizing Maps (SOM) is given. The methodology was tested on two samples of special-purpose electronic components using application software. The analysis of the SOM is given. The classification accuracy is estimated and a comparison table, that includes the results of cluster analysis, is given. The recommendations for improving the classification quality are developed.

1. Introduction

The development and perform actions for the improving reliability of the radioelectronic equipment, installed on the spacecraft, is one of the most important tasks assigned to scientists and specialists in the space industry. The task of increasing reliability is most successfully solving by forecasting the future state of equipment [1]. The search for the new forecasting methods, based on the use of mathematical models and most applicable to non-renewable equipment, is relevant at this stage in the development of science and technology [2]. Such methods include the individual forecasting (IF). Its main idea is that the value of the informative parameter or the results of monitoring each instance using the forecast model formulates a conclusion about the potential reliability of this instance [1].

In paper [3] forecast models are presented for the CMOS chips and zener diodes (stabilitrons) samples using regression models. Verification of the results was carried out using the method of discriminant functions. The methods allowing to split the initial samples of electronic components into classes of reliable and potentially unreliable elements for verifying the results of the IF are presented in works [4, 5]. These methods are based on the cluster analysis using the k-means clustering and based on the agglomerate methods of hierarchical clustering with subsequent evaluation of the results.

This paper discusses the possibility of using Self-Organizing Maps (SOM) to classify electronic components into classes of reliable and potentially unreliable instances. Furthermore, the technique of using such network to classify samples of electronic components is presented.

2. Kohonen networks

The Kohonen networks are referred to as self-organizing neural networks, which allow to identify groups of input data vectors with similar properties [6,7].

The Kohonen network is a single-layer network constructed from WTA neurons (Winner Takes All). Figure 1 shows the structure of such network.

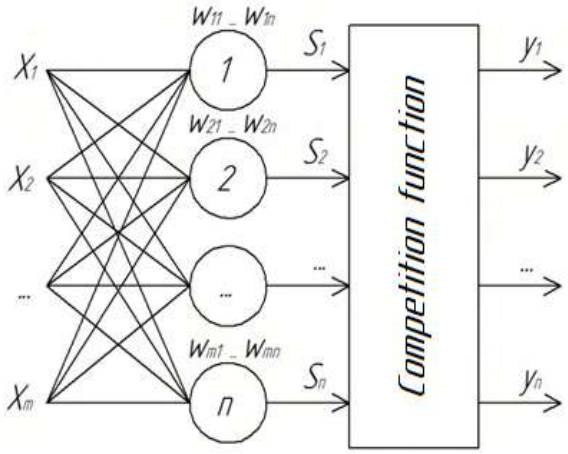


Figure 1. A structure of the Kohonen network.

Each neuron is associated with each component of the input data $x_1, x_2,\ldots,x_m$ (the input data vector), which in this case is a set of informative parameters for each instance of the sample. Each neuron is presented as a linear combiner:

$$S_j = b_j + \sum_{i=1}^{m} w_{ij}\, x_i, \tag{1}$$

whereS_j – is an output result of the combiner; j – is a neuron number; i – is a number of the informative parameter; b_i - is a threshold; w_{ij} – is a weight of the i[th] parameter of the j[th] neuron.

From the output of each neuron the result goes to the competition function, which defines the neuron with the maximum result value at the output and assigns it a value of one. The remaining output signals are assigned a value of zero. For the neuron-winner, the following condition is fulfilled:

$$d(x, w_j) = \min_{1 \le i \le n} d(x, w_i), \tag{2}$$

where$d(x, w)$ – is a distance between the vectors of the input data and the vector of the synaptic weights, j – is a number of the neuron-winner. It should be noted that the distance in the selected metric is an Euclidean distance.

The neuron-winner is a neuron, the vector of synaptic weights w_i of which differs least from the vector of input data. In the case of determining the maximum result at the outputs of several neurons, the result equal to one is assigned to one of them, and the outputs of the remaining neurons are assigned a value of zero. Neuron-winner defines such a group, to which the vector of input data is closest .The scales are adjusted each training cycle:

$$w_j^{(k+1)} = w_j^{(k)} + \eta_j^{(k)} \left[X - w_j^{(k)}\right], \tag{3}$$

where$k$ – is a number of the training cycle; X – is an input vector; $\eta_i^{(k)}$ – is a coefficient of learning rate of the j[th] neuron in the k[th] cycle.

Thus, the neuron, whose synaptic weights vector was closer to the vector of the input data, is corrected and becomes even closer. Also, the weights of the nearest neurons are corrected to the neuron-winner. Thus, the probability of this neuron to become a neuron-winner is increased when a closer vector of data is input.

The training of the Kohonen network consists in the selection of weight values when errors are minimized.

3. Classification technique

The developed classification technique based on Kohonen self-organizing maps is as follows:

- determination of the vector of input data based on the results of the learning experiment;
- definition of network construction parameters;
- analysis of the results with an assessment of the accuracy of the classification.

Two samples of CMOS chips, obtained in [3-5], were chosen as the study samples. The volume of both samples is 50 instances. Time delay on the leading edge of the signal $x_1(t^+_p$, μs) and the critical supply voltage $x_2(V_{cs}$, V) are used as the informative parameters. The leakage current drift y is assumed as the forecasting parameter.

Thus, the vector of input data is a set of informative parameters x_1 and x_2 of each sample.

The analytical platform "Deductor Academic 5.1" was chosen as a software tool for building Kohonen network. The following building parameters need to be defined:

- size of the training and test set;
- a network size (number of neurons);
- a recognition criterion (the amount of recognition error);
- rate and radius of learning;
- number of clusters.

Because the samples size is small, the training set should be as large as possible and sufficient for the correct building of the network. In this case, the test set should be sufficient for testing the network with 100% recognition. Based on the set conditions, the size of the training set was 45 instances (90% of the sample), the test set size – 5 instances (10% of the sample).

The network size is chosen in such a way as to exclude the appearance of "dead" neurons (vector of synaptic weights is significantly removed from the vector of input data). Such neurons turn out to be inactive in network training, because they cannot win competition from the nearest neurons. This situation leads to the fact that the input data is interpreted by a smaller number of neurons than was originally set, which introduces distortions into the final result.

Networks with a dimension of 5×4 neurons (cells) were chosen for testing the presented samples. The type of cells is selected hexagonal: it more accurately displays the Cartesian distance between the cell centres. The remaining parameters of the Kohonen network building are presented in table 1.

Table 1. Kohonen network building parameters.

Network parameter	Criterion
Recognition error value	An example is considered recognized if the recognition error is less than 0,05
Number of the stages	200
Training rate η	In the beginning of the training: 0,3 In the end of the training: 0,005
Training radius	In the beginning of the training: 2 In the end of the training: 0,2
Number of clusters	Fixed number: 2

From the whole set of variants of networks built according to the given parameters, the only option was chosen for each sample, at which the minimum average recognition error for the training set was achieved with 100% recognition of both the training set and the test set.

Because a fixed number of clusters equals two, according to the classification table it is not difficult to determine to which class of products the network relates each instance.

To estimate the accuracy of the classification, the technique given in [4, 8] was applied.

4. Results of building a network for sample №1

Figure 2 shows the maps of the input data x_1, x_2 and the output data y for the network. The maps of such parameters are called component planes. Figure 3 shows the density hit matrix, the matrix of distances between the centres of neurons, and also a map of cell decomposition into two clusters.

Figure 2 shows that the maps of the neuron inputs (maps X1, X2) and the output map have a similar appearance: on the left there are cells that characterize neurons with smaller weights, and on the right there are cells characterizing neurons with large weights. Because the output map of neurons (map Y) is a projection of the maps of neuron inputs, its appearance will be similar.

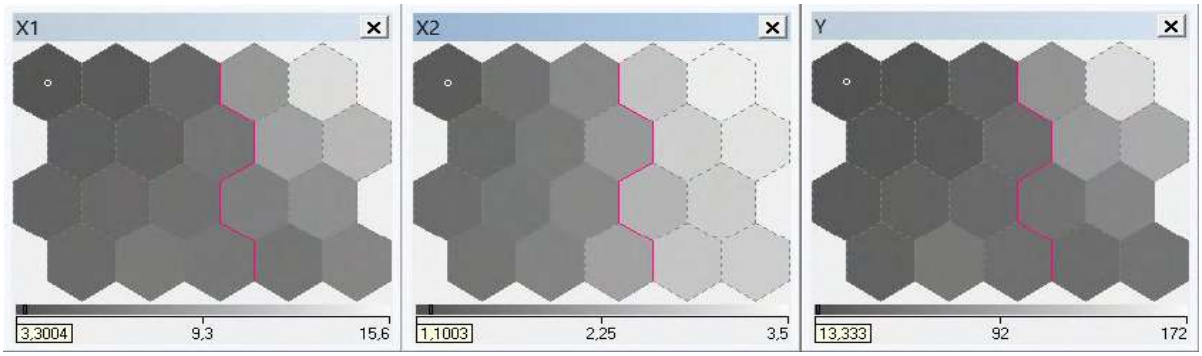


Figure 2. The maps of the input data x_1, x_2, output data y for the sample №1.

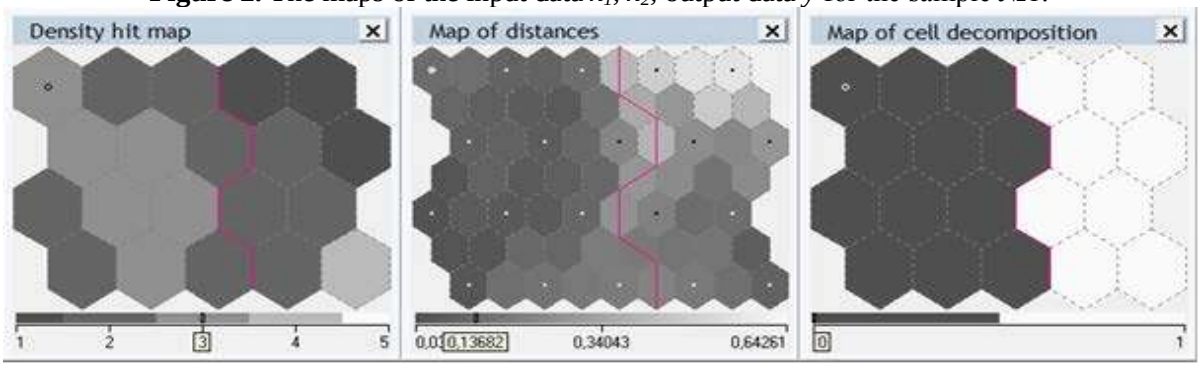


Figure 3. The density hit map, the map of distances, the map of cell decomposition for the sample №1.

Figure 3 shows that the map of cell decomposition is divided into 2 areas that characterize clusters. In this case, cluster 0 is a class of the reliable instances (class 1), cluster 1 is a class of the potentially unreliable instances (class 2). Table 2 provides the information on assigning each instance to a specific class, and K_{fact} means the actual belonging of the instance to a class, $K_{rec.}$ – is an assignment of the instance to a class based on recognition results using Kohonen maps.

Table 2 shows that for some instances, the actual class and the recognition class do not match. Therefore, it is necessary to estimate the accuracy of this classification. For this the technique presented in [4, 8] may be applied. The results of the classification accuracy evaluation are presented in table 3.

Table 2. The ownership of the sample №1 instances to classes.

№ instance	K_{fact}	$K_{rec.}$	№ instance	K_{fact}	$K_{rec.}$	№ instance	K_{fact}	$K_{rec.}$
1	1	1	**18**	1	1	**35**	2	2
2	1	2	**19**	1	1	**36**	1	1
3	1	1	**20**	1	1	**37**	1	1
4	2	1	**21**	1	1	**38**	1	2
5	1	1	**22**	2	2	**39**	2	2
6	1	1	**23**	1	1	**40**	2	2
7	2	1	**24**	1	1	**41**	1	1
8	1	1	**25**	2	2	**42**	1	1
9	1	2	**26**	1	2	**43**	2	2
10	1	1	**27**	1	1	**44**	2	2
11	1	1	**28**	1	1	**45**	1	1
12	1	1	**29**	1	1	**46**	2	2
13	1	1	**30**	1	1	**47**	2	2
14	2	1	**31**	1	1	**48**	1	1
15	2	2	**32**	2	2	**49**	1	1
16	1	2	**33**	1	1	**50**	1	1
17	1	2	**34**	2	2			

Table 3. The results of the classification accuracy evaluation for the sample №1.

Criterion	Value
Producer's risk (α-risk) $P_{prod.}$	0,333
Consumer's risk (β-risk) $P_{cons.}$	0,094
Risk of the incorrect decision $P_{inc.\ d.}$	0,18

5. Results of building a network for sample №2

Figure 4 shows the maps of the input data x_1, x_2 and the output data y for the network. Figure 5 shows the density hit matrix, the matrix of distances between the centres of neurons, and also a map of cell decomposition into two clusters.

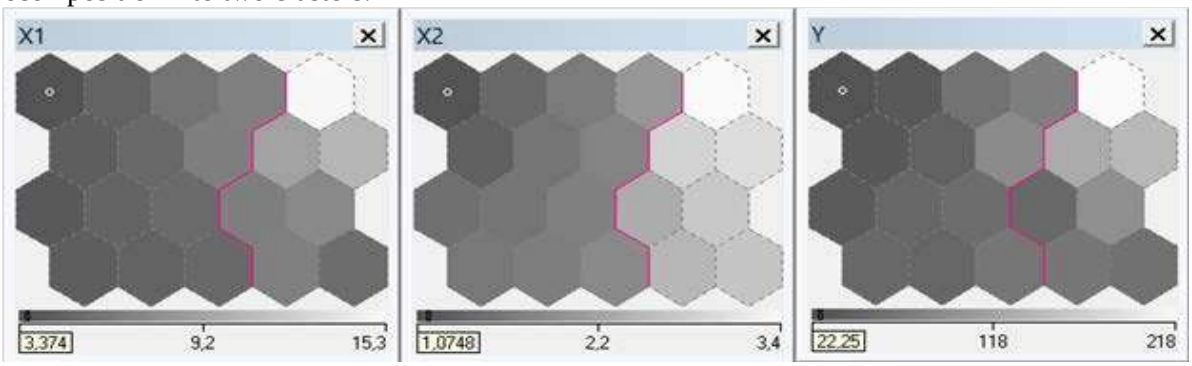


Figure 4. The maps of the input data x_1, x_2, output data y for the sample №2.

Figure 4 shows that the neuron input maps and the output map also have a similar appearance: on the left are cells that characterize neurons with smaller weights, and on the right are cells characterizing neurons with large weights.

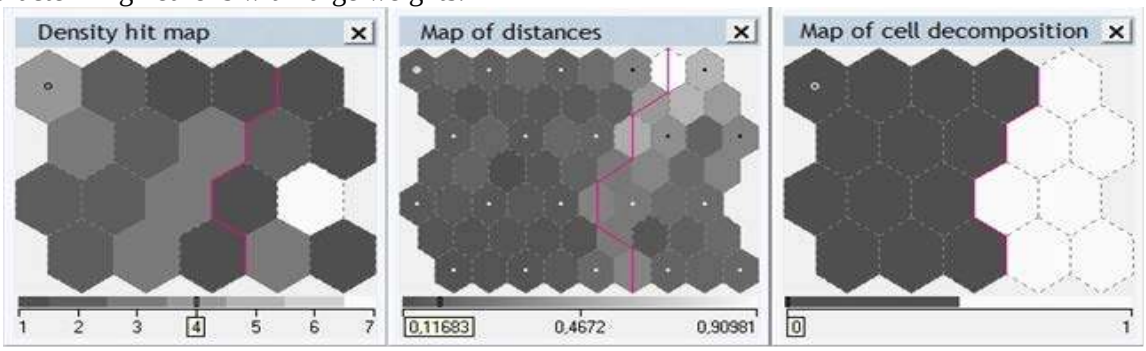


Figure 5. The density hit map, the map of distances, the map of cell decomposition for the sample №2.

Figure 5 shows that the map of cell decomposition is divided into 2 areas. Cluster 0 is a class of reliable instances (class 1), cluster 1 is a class of potentially unreliable instances (class 2). Table 4 provides information on assigning each instance to a specific class. Table 4 shows that for some instances, the actual class and the recognition class do not match. Therefore, it is necessary to estimate the accuracy of this classification.

Table 5 compares theaccuracy estimatesofthe classification with the best results obtained in [4].

6. Conclusion

The results are explained by the fact that small samples were used as input data, and, consequently, small training sets are used. The size of the networks was also limited, because with increasing network size the "dead" neurons appear, which should be avoided. One of the advantages of the developed technique is that the trained network is able to recognize any instance of the dot of products that is not part of the sample, which could not be done by cluster analysis algorithms. Thus, the

developed technique can be recommended both as a method for verifying the results of the IF, and as the technique of the IF.

Table 4. The ownership of the sample №2 instances to classes.

№ instance	K_{fact}	$K_{rec.}$	№ instance	K_{fact}	$K_{rec.}$	№ instance	K_{fact}	$K_{rec.}$
1	2	2	**18**	2	1	**35**	1	1
2	1	1	**19**	1	1	**36**	1	1
3	2	2	**20**	1	1	**37**	1	1
4	1	1	**21**	2	1	**38**	1	1
5	1	1	**22**	1	1	**39**	1	1
6	1	1	**23**	2	2	**40**	1	2
7	1	1	**24**	1	1	**41**	2	1
8	1	2	**25**	1	1	**42**	1	1
9	2	2	**26**	2	2	**43**	1	1
10	2	2	**27**	1	2	**44**	2	2
11	1	1	**28**	1	2	**45**	2	2
12	1	1	**29**	1	1	**46**	2	2
13	1	1	**30**	1	1	**47**	2	2
14	1	1	**31**	2	2	**48**	2	2
15	1	1	**32**	2	2	**49**	1	1
16	1	2	**33**	1	1	**50**	1	1
17	1	1	**34**	2	2			

Table 5. Comparison table of the classification accuracy.

Criterion	SOM	Cluster analysis algorithms		
		Var. 1	Var. 2	Var. 3
Producer's risk (α-risk) $P_{prod.}$	0,263	0,238	0,238	0,263
Consumer's risk (β-risk) $P_{cons.}$	0,097	0,069	0,034	0,129
Risk of the incorrect decision $P_{inc.\ d.}$	0,16	0,14	0,12	0,18

7. References

[1] Piganov M N 2009 Reliability prediction of electronic equipment *Scientific and Technical Bulletins SPbGPU. Series "Informatics. Telecommunications. Management"* **1** 175-182
[2] Piganov M N, Tyulevin S V and Erantseva E S 2015 *The Experience of Designing and Application of CAD Systems in Microelectronics* (Lviv: IEEE) 367-371
[3] Mishanov R O, Tyulevin S V, Piganov M N and Erantseva E S 2017 Forecasting models generation of the electronic means quality *Mathematical Modeling* **1** 124-129
[4] Mishanov R O and Piganov M N 2017 Method of verification of the results of individual forecasting of products of space radio-electronic equipment with the help of algorithms of cluster analysis *Physics of wave processes and radio engineering systems* **20** 55-63
[5] Mishanov R O 2017 Method of applying cluster analysis for the classification of electronic products and improving the reliability of equipment *Izvestiya of the Samara Scientific Center RAS* **19** 414-419
[6] Kohonen T 2008 *Self-organizing maps* (Moscow: BINOM Laboratoriya znanij) p 655
[7] Haykin S 2006 *Neural networks: full course* (Moscow: Vilyams) p 1104
[8] Tyulevin S V 2016 *Individual forecasting of electronic means: schoolbook* (Samara: Samara University) p 65

Development of Traffic Signs Recognition WebService based on Convolutional Neural Networks

K A Pronchuk[1] and P Y Yakimov[1]

[1]Samara National Research University, Moskovskoye Shosse 34A, Samara, Russia, 443086

Abstract. Image classification is one of the most important applications of neural networks. In this paper,we propose the classification algorithm for traffic signs recognition based on convolutional neural networks.The designed CNN is implemented using the TensorFlow framework, and the inference is performed using CUDA. To utilize the Connected Cars concept, we also developeda webservice to remotely process images, obtained by a camera installed into a vehicle. The experimental results show that the proposed algorithm, implemented using the CUBA library for developingclient-server apps, shows high efficiency and is applicable for a connected vehicle.

1. Introduction

The Advanced driver assistance systems (ADAS) are high-tech vehicle systems that are designed to improve road safety, help drivers better understand the road and its potential dangers. The recognition of road signs is one of the important subsystems of ADAS. Implementing a real time traffic sign recognition system is usually divided into two steps: localization of a road sign and its classification. The detection is often performed using simple computational algorithms, such as the color thresholding with shape detection [1], [2], [3]. After that, the classification takes place using more complex, and at the same time more accurate algorithms.

Recently, with the development of high speed networks like 5G, the concept of Connected Cars [4] is becoming more popular and, more important, more possible.The term ‘Connected Car’ refers to applications, services, and technologies that connect a vehicle to its surroundings. A connected vehicle is basically the presence of devices in a vehicle that connect to other devices within the same vehicle and/or devices, networks, applications, and services outside the vehicle. The applications of the concept include everything from traffic safety and efficiency, infotainment, parking assistance, roadside assistance, remote diagnostics, and telematics to autonomous self-driving vehicles and global positioning systems (GPS). Typically, vehicles that include interactive advanced driver-assistance systems and cooperative intelligent transport systems (C-ITS) can be regarded as connected.

Usually, on-board computers of connected vehicles have rather low performance, which is quite enough to implement various network connection functions, including sending the data from its sensors and receiving some data from other vehicles.But using it for such tasks as traffic sign recognition is impossible. However, nowadays it becomes possible to use remote computing power to solve this problem: the service can send the image to a remote server for processing and further obtaining a response about the presence of road signs in it. Thus, in the traffic sign recognition task, the web service only transfersan image from the camera installed into a vehicle to remote server, gets

the response from it and displays it on an interactive map, which then can used by other cars in the whole system.

2. Recognition algorithm

2.1. HAAR cascade classifier for sign detection

The Haar feature-based Cascade Classifier initially proposed by Viola and Jones [5], [6] classifies objects using a series of edges, lines, and center-surround features that scan throughout the image to construct ROI features. The name cascade means that the resultant classifier consists of several simpler classifiers that are applied to the image one at a time in the order of their classification effectiveness.To speed up the detection process, when earlier classifiers fail to produce a match, the remaining classifiers are no longer applied.We implemented Haar cascade detection for all 43 classes of signs. The number of classes was chosen because we used a dataset containing 43 types of signs.

2.2. Convolutional Neural Neworks

Recently, neural networks have been actively used in classification tasks. The neural network is a mathematical model built on the principle of the organization and functioning of biological neural networks. Neurons are organized into layers: input, hidden and output. The input layer does not consist of complete neurons, but rather consists of values that are inputs to the next layer. The next layer is hidden. In one neural network, there may be several hidden layers. The last level is the output layer, where each class corresponds to one node [7].

The key part to understand, which distinguishes CNN from traditional neural networks, is the convolution operation. Having an image at the input, CNN scans it many times to look for certain features. This convolution can be set with 2 main parameters: stride and padding type.

2.3. Proposed implementation

In this paper, we chose a classic LeNet model which contains only two convolution layers. To train the model we use TensorFlow – the deep training library. Training and testing were performed on the data set of The German Traffic Sign Recognition Benchmark [6]. In this version of the system, it recognizes the 43 types of traffic signs.

Table 1 shows the architecture of the developed convolutional neural network. The network is a sequence of interconnected layers, starting with a convolutional layer and ending with a layer of softmax. The parameter of the convolutional layer - step - determines the step of the rolling convolution window. In case the step is greater than 1, the convolutional layer will combine with polling. The softmax layer performs the normalization of the results of the previous layer in such a way that its probabilities will form the probabilities of the object's relation to one of 43 classes.

Table 1. LeNet-5 architecture.

Layer	Description
Input	32x32x1 Greyscale image
Convolution 5x5	1x1 stride, valid padding, outputs 28x28x6 + ReLU
Avg pooling	2x2 stride, outputs 14x14x6
Convolution 5x5	1x1 stride, valid padding, outputs 10x10x16 + ReLU
Avg pooling	2x2 stride, outputs 5x5x16
Fully connected	inputs 400, outputs 120 + ReLU
Fully connected	inputs 120, outputs 84 + ReLU
Fully connected	inputs 84, outputs 43
Softmax	

The model can be divided into 2 blocks:the convolution block and the fully connected block. TensorFlow includes network mapping tools that allow you to visualize the model at different levels of abstraction, down to low-level mathematical operations.

To generate so-called augmented data, we randomly chose images to copy. To provide additional information to the model, we randomly rotated this copy and changed its brightness. To implement all these operations,the OpenCV library was used. We performed these operations until each label had 3200 samples. This increased the training set to 139.148 samples. As an illustration, here's a drawing of a sample traffic sign with generated images (rotated and different brightness).

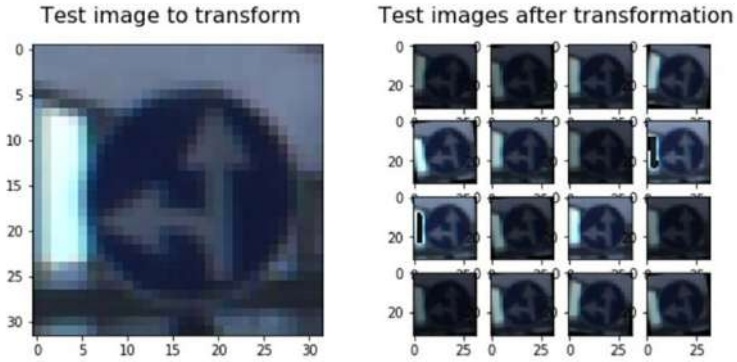


Figure 1. Image transformation examples.

Further, there are the details of intermediate steps and the corresponding validation accuracies after 100 epochs of training.

- Initial LeNet model, choosing input images color representation - 91 %
- Input images normalization - ~91 %
- Training set augmantation - 93 %
- Learn rate optimization, from this stage I tested for 100 epochs - 95 %
- Finding optimum image transformations during training set augmentation - 96 %
- Trying different pool methods, trying dropout, choosing L2 loss, tuning learn rate again - 96.8

Final model results were as follows:

- Training set accuracy of 99.5 %
- Validation set accuracy of 96.8 %
- Test set accuracy of 94.6 %

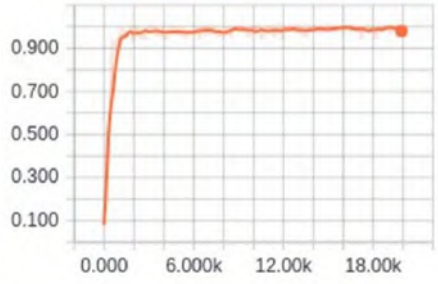


Figure 2. LeNet-5 accuracy.

The convolutional networks also have drawbacks, one of which is the max-pooling layer and the invariance of detection only to the position in the image. The network can falsely recognize the same type of image with a specific rotation, so training is accompanied by an increasing test data and the addition of a max-pooling layer. The layer is used to reduce the dimensions of the outputs of convolutional layers, ignoring small differences in the spatial structure of images, and some information is lost.

In the past decade, the CNNs have really become a revolutionary tool for computer vision developers. However, some new types of neural networks are still introduced. Let's take a closer lookat the architecture of the capsular neural networks described by Hinton in [8]. Figure 3 shows an example architecture of a capsular neural network.

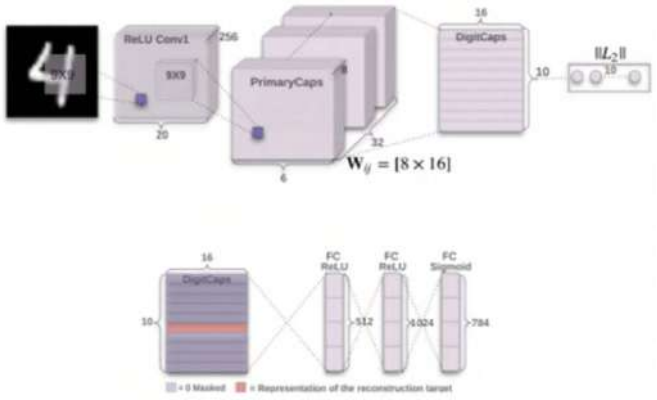


Figure 3. CapsNet architecture.

First, there is a standard separation of features that are invariant to translation in an image, using a convolutional layer. Then another 32 features layers are isolated, each of which is connected to the first convolutional layer. This is necessary for dynamic routing. Next comes the "digitCaps" layer,thelayer that gave the name for the whole network. Each capsule has a strict meaning: the amplitude of the vector in each of them corresponds to the probability of the presence in the image of one of the desired classes. At the bottom of the image, an important part of the architecture is represented: a decoder, on the input of which a layer of DigitCaps is supplied, with all the capsules except the one with the largest amplitude being reset. Decoders are essentially an inverted ANN (artificial neural network), at the input of which a small vector was applied, and at the output - the original image. This decoder provides the second most important part of this architecture: the contents of one capsule should fully describe the subset of a particular class being fed to the input. Otherwise, the decoder can not restore the original image in any way. There is another useful property of the decoder - this is the method of regularization. Close codes in the capsules will be close to Euclid images.

In this paper, we also describe a capsular neural network,trained and then used for signs recognition. The network was trained on a 39,209 images database and contains 43 traffic sign classes. Figure 4 shows an increase in the accuracy of classification over time.

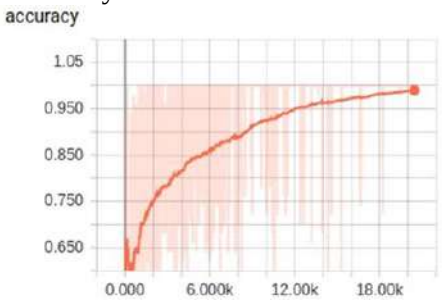


Figure 4. CapsNet accuracy.

3. Web service development

To use the trained neural network, we developeda web service based on the CUBA platform. CUBA is the platform for rapid development business applications in Java, which allows to quickly create a UI and logic for working with data. The Vaadin framework is used in the default web client to display the generic user interface. Vaadin offers a comprehensive set of extensible components and supports SCSS for UI customization. At the same time its server-based implementation model improves security and allows to use Java for both client and server-side code. Spring container provides core infrastructure for the middle tier and the application's client blocks. The framework is also used to establish remote interaction between blocks and for web services implementation.

Shown in figure 5, CUBA Studio is a specialized tool for rapid application development, that streamlines building applications upon CUBA Platform. This is a separate application, which the

developer uses in parallel with the regular Java IDE. The studio provides a graphical interface to the platform mechanisms that allows you to create a data model using the mouse, generate DDL scripts for the database, draw screens in the WYSIWYG editor, and make the middle layer services.

Applications based on CUBA have a standard three-layer architecture. The connecting element of the system is metadata - information about the data model of the application. Relates to metadata, visual components know what data they are working with. In the same way, metadata helps the visual components to work with the database through the ORM, specifying the graphs of objects that need to be downloaded or updated. The same principle applies to the security subsystem, the generation of reports and other parts of the platform.

Applications created with the CUBA platform can be deployed (figure 6) in various configurations, from starting all the components of the application on a single server and up to configurations that provide high fault tolerance, with separated web, middle and database layers. Out of the box supports PostgreSQL, Oracle Database, Microsoft SQL Server, MySQL and HSQL (the latter is usually used for prototyping), between which you can switch as the project grows. It is also important to note that CUBA applications can be deployed on any Java EE Web Profile server, including Jetty, Tomcat, Glassfish, Websphere, etc.

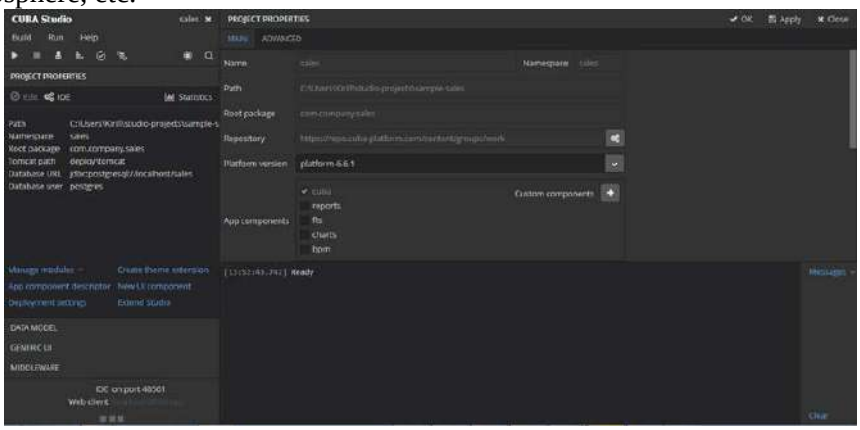


Figure 5. Project main page.

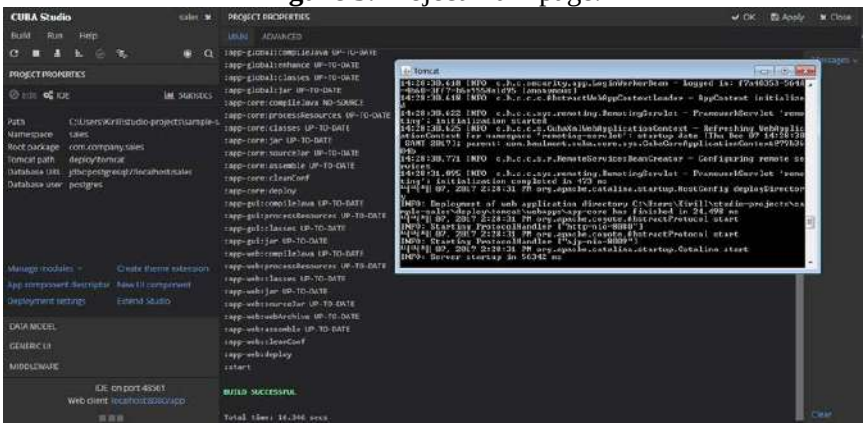


Figure 6. Project deployment.

4. Experiments

The German database of annotated images containing road signs was used [3] to test the accuracy of the detection algorithm. It contains more than 50,000 images with traffic signs registered in various conditions. To assess the quality of detection, the number of images with correctly localized and classified traffic signs was calculated. The experiments showed 96.8% of correctly localized and classified prohibitory and warning road signs.

To test the performance of the developed web service we used images containing a 30 km / h speed limit sign. Various stages of the road sign classification process, such as: loading the trained Tensorflow model, loading a cascade classifier and detecting the presence of a sign on the image,

classifying the sign (if present on the image), the total classification time were evaluated during the experiment.

Table 2. LeNet-5 architecture.

	TF model loading (*ms*)	Sign detection(*ms*)	Sign classification(*ms*)	Total time(*ms*)
1	101	142	12	300
2	113	134	1	315
3	109	157	10	330
4	117	149	12	319
5	100	124	10	423
6	-	-	-	-

Figure 7 shows an example of a successful recognized traffic sign. The information about the sign type can now be transferred to all connected vehicles specifyingits geographical coordinates. Figures 7 and 8 shows the user interface for the developed web service.

Figure 7. Example of recognized image.

Figure 8. Web service UI.

5. Conclusion

In this project, we developed a method for the classification of traffic signs using convolutional neural networks and integrated it into web service based on CUBA Platform.

The experiments have shown high classification efficiency of 96.8%. Moreover, capsular neural networks showed better result compared with convolutional neural networks. This became possible thanks to the more complex CapsNet device, but the performance also decreased. In final version of application, we usedthe LeNet-5.

In addition to solving the problem of classification of traffic signs, this paper also proposes a method for their localization using the cascade classifier based on the Haar features. This method has successfully proven itself to high productivity and accuracy, which was also confirmed in this work.

The developed web service is intended for use in the Connected Car concept. This service makes it possible to create and fill interactive maps, inform other vehicles about road events, and later can be used for unmanned vehicles.

In future work, we plan to refine the existing solution by adding REST API and the ability to classify characters on video recordings.

6. References

[1] Shustanov A and Yakimov P 2017 CNN Design for Real-Time Traffic Sign Recognition *Procedia Engineering* **201** 718-725
[2] Ruta A, Porikli F, Watanabe S and Li Y 2011 In-vehicle camera traffic sign detection and recognition *Mach Vis Appl.* **22** 359-375 DOI: 10.1007/s00138-009-0231-x
[3] Belaroussi R, Foucher P, Tarel J-P, Soheilian B, Charbonnier P and Paparoditis N 2010 Road Sign Detection in Images: A Case Study *Pattern Recognition 20th International Conference ICPR* 484-488
[4] Koesdwiady A, Soua R and Karray F 2016 Improving Traffic Flow Prediction with Weather Information in Connected Cars: A Deep Learning Approach *IEEE Trans Veh Technol*. 9508-9517
[5] Yakimov P Y 2015 Tracking traffic signs in video sequences based on a vehicle velocity *Computer Optics* **39(5)** 795-800
[6] Houben S, Stallkamp J, Salmen J, Schlipsing M and Igel C 2013 Detection of traffic signs in real-world images: The German traffic sign detection benchmark *Proceedings of the International Joint Conference on Neural Networks* 715-722
[7] Sermanet P and LeCun Y 2011 Traffic sign recognition with multi-scale Convolutional Networks *Neural Networks The International Joint Conference (IJCNN)*
[8] Hinton G E, Sabour S, Frosst N 2017 Dynamic Routing Between Capsules *Proceedings of the 30th Conference on Neural Information Processing Systems (NIPS)*

Acknowledgments

This work was supported by the Russian Foundation for Basic Research - Project # 16-37-60106 mol_a_dk.

Information-analytical system of chemical technology processes modeling by the use of parallel calculations

I V Akhmetov[1] and I M Gubaydullin[1,2]

[1]Ufa State Petroleum Technological University, Kosmonavtov St. 1, Ufa, Russia, 450062
[2]Institute of Petrochemistry and Catalysis of the Russian Academy of Sciences, Prospekt Oktyabrya 141, Ufa, Russia, 450075

Abstract. An information-analytical system for modelling of chemical technology processes by the use of parallel computing has been developed. On the basis of this system a kinetic model for the synthesis of benzylidene benzylamine has been built. The optimal conditions to maximum yield of target product were found.

1. Introduction

In order to explore mechanisms for chemical technology processes, the kinetic models should be developed. The development of a kinetic model requires solving the inverse kinetic task. This stage the most laborious and time consuming [1-4]. The use of parallel computing is becoming increasingly instrumental method of mathematical processing of experimental data [5, 6].

Inverse chemical kinetics problems suppose a significant amount of computation. The use of high-performance computing systems has fundamentally changed the ability to analyze complex chemical processes. A detailed analysis of rather complex kinetic models with a large amount of experimental information became available. The time required to build kinetic models has been significantly reduced. The accuracy of solutions has been improved.

Authors of this work has developed and tested the information-analytical system for modeling chemical technology processes. On the basis of the technology of parallel calculations the actual method of search of kinetic parameters is offered.

The use of high-performance systems allows us to reduce the time of study and development of new chemical processes.

2. Information-analytical system for modelling chemical technology processes

The information-analytical system (IAS) is proposed to use for studying chemical technology processes. Its structure is given on the figure 1.

Input information flows include:

- Experimental conditions: reactor type, process flow mode, temperature, pressure, volume of reaction mixture.
- Chemical formula involved in the reaction.
- Concentrations at the initial time.
- Concentrations at some fixed points in time.

- Evaluation of kinetic parameters: ratios between constants, certainty intervals of some kinetic parameters.
- Proposed mechanisms of chemical reactions.

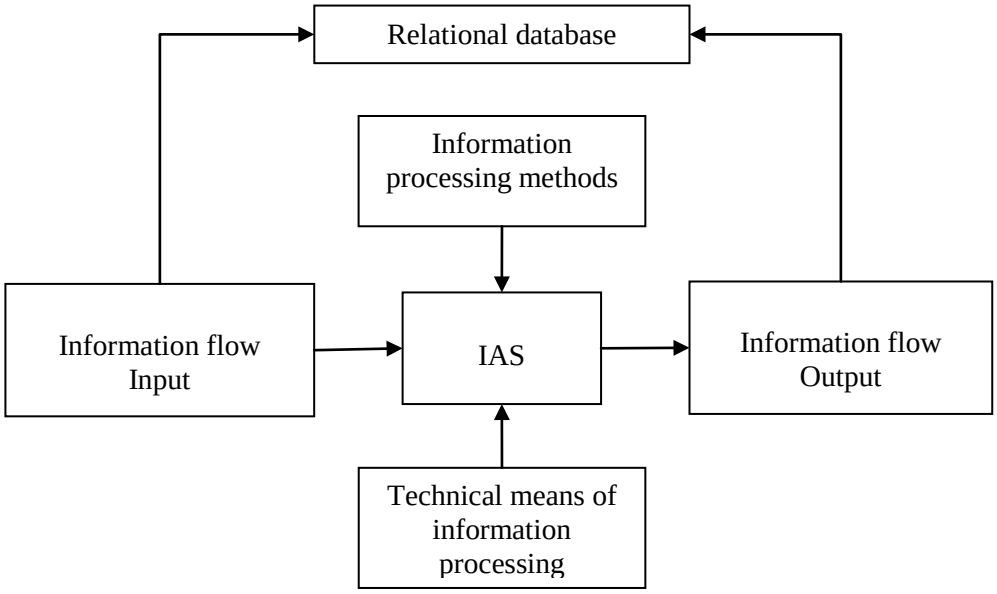


Figure 1. Information-analytical system structure for modeling chemical technology processes.

Information processing methods include:

- The algorithm for solving the direct kinetic problem.
- Parameters of computing control in solving the inverse kinetic problem.
- The algorithm for solving the inverse problem.
- Graph plotting for comparison the calculated and experimental concentrations of the reacting substances.
- The algorithm for determining the activation energies of the reaction stages.

Output information flows include:

- Calculated concentrations of all reacting substances.
- Calculated activation energies of all reaction stages.
- Graphs for comparison the calculated and experimental concentrations of the reacting substances.
- Graphs of change in reaction stage rate.

Technical means of information processing include:

- Experimental unit.
- Industrial reactor.
- Computing system: personal computer, multiprocessor computer system.

3. Application of parallel computing

The genetic algorithm for the parallel solution of the inverse problem of chemical kinetics is the most effective. This algorithm is borrowed from biology. It is based on the idea of selection, i.e. the predominant reproduction of the most adapted individuals [7]. The practical application of the genetic algorithm in all known cases has led to positive results [8, 9]. The genetic algorithm is a universal method for finding the optimum regardless of the complexity of the functions [9]. The basis of the genetic algorithm presents the following sequence of operations.

On the first step the algorithm randomly creates the initial population of N individuals. Then N points are formed in the space of kinetic parameters. Each point has m coordinates, i.e. values of the parameters.

At the stage of mutation, individuals of the population change in accordance with a predetermined operation. The coordinates/parabolic descent from the points of the space were taken as the operation of mutation.

At the stage of selection, a certain fraction of the entire population is selected. The chosen fraction will remain "alive" at this stage of evolution. The probability of survival of the individual depends on the value of the adaptiveness function for this individual. The function of adaptiveness is characterized by the residual function. The fraction of survivors (s) is a parameter of the genetic algorithm. According to the results of the selection from N individuals of the population the total population will comprise s×N individuals. In the considered case s=1/2. The crossing is used to form a new generation. It needs two parents to produce a descendant. To form a new point in the parameter space, one of the "survivors" and one of the "dying" points are selected as parents. The crossing is carried out by choosing m/2 coordinates from the first point and the remaining ones from the second. In this case, the descendant inherits features of both parents. To exclude the degradation of the population, the individuals for reproduction are selected from the entire population. This set of actions is repeated iteratively. By this way the "evolutionary process" is modeled and several life cycles (generations) continue. The algorithm stops when one of the following conditions is met:

- finding a global or suboptimal solution;
- exhaustion of the number of generations for evolution;
- exhaustion of time given for evolution.

Parallelization of the computational process takes place at the stage of initial filling. Pseudo-random points in the parameter space are uniformly distributed over the flows of a multi-core computing system. The mutation is carried out by each flow independently. Data exchange is carried out at the selection stage. The autonomous operation time of flows significantly exceeds the interaction time between flows. Therefore this algorithm is effective.

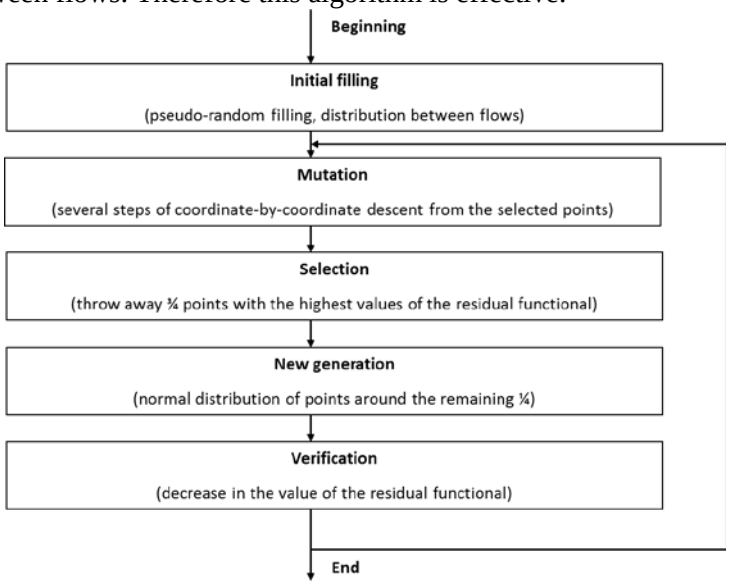
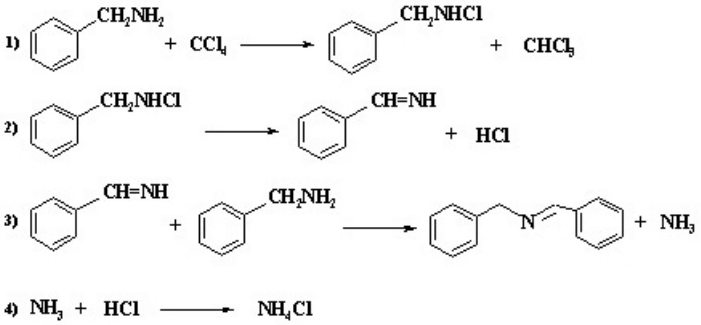


Figure 2. Genetic algorithm.

4. Kinetic model of the reaction for the synthesis of benzylidene benzylamine

The developed information-analytical system was tested when modeling of a kinetic model for the synthesis of benzylidene benzylamine. In the series of Schiff bases, N-benzylidene benzylamine has an important place and has wide application in practice. In particular, N-benzylidene benzylamine is known as an indicator for the quantitative determination of lithium-organic compounds by titrimetric method [10-12]. N-benzylidene benzylamine is used as an initial compound for the synthesis of a number of heterocycles. It has a wide range of physiological activity: analgesic, anticonvulsive, anticonvulsant [13]. Based on the analysis of the experimental data and the results of their

mathematical treatment [14-15] the following sequence of chemical transformations in reactions of benzylideneaniline synthesis is proposed:



The sequence of reaction transformations and their corresponding kinetic equations are represented by equations:

1. $C_1 + C_2 \rightarrow C_3 + C_4$ $W_1=k_1C_1C_2$
2. $C_3 \rightarrow C_5 + C_6$ $W_2=k_2C_3$
3. $C_5 + C_1 \rightarrow C_7 + C_8$ $W_3=k_3C_5C_1$
4. $C_8 + C_6 \rightarrow C_9$ $W_4=k_4C_8C_6$

where C_i – concentration of components, mol/L: $C_1=C_7H_9N$ – benzylamine, $C_2=CCl_4$ – carbon tetrachloride, $C_3=C_7H_8NCl$ – chlorbenzylamine, $C_4=CHCl_3$ – chloroform, $C_5=C_7H_7N$ – 1-phenylmethanimine, $C_6=HCl$ – hydrogen chloride, $C_7=C_{14}H_{13}N$ – benzylidene benzylamine, $C_8=NH_3$ – ammonia, $C_9=NH_4Cl$ – ammonium chloride; W_j – rate of the j-th reaction, mol/(L×h); k_j – kinetic rate constant of the j-th reaction, $l\times mol^{-1}\times h^{-1}$ (j=1, 3, 4), h^{-1}.

The kinetic equations of the transformation sequence are analyzed within the law of acting masses [16]. The correct description of a laboratory reactor with a stirrer is the ideal mixing model:

$$\frac{d\overline{N}}{dt} = F_N,\ F_N = \frac{1}{V_o}\sum_{j=1}^{J}\delta_j\omega_j,\ \delta_j = \sum_{i=1}^{I}\nu_{ij}$$

$$\frac{dX_i}{dt} = \frac{F_i - X_i F_N}{\overline{N}}$$

with the initial conditions: at $t = 0, X_i = X_i^o,\ \overline{N} = 1,$ где $\overline{N}$ = C/C_o – the relative change in the number of moles of the reaction mixture; C and C_o – molar density and its initial value, mol/L; $X_i=C_i/C$ – concentration of components, mole fractions; V_o – volume of reaction space, L; $\omega_j=W_j/C_o$ – adjucted reaction rates, h^{-1}; j – number of stages of chemical transformation; i – number of components.

The right-hand sides of the system of nonlinear differential equations have the following form:

$F_1 = -\omega_1 - \omega_3$; $F_2 = -\omega_1$; $F_3 = \omega_1 - \omega_2$; $F_4 = \omega_1$; $F_5 = \omega_2 - \omega_3$; $F_6 = \omega_2 - \omega_4$; $F_7 = \omega_1$; $F_8 = \omega_3 - \omega_4$; $F_9 = \omega_4$; $F_n = \omega_2 - \omega_4$.

To solve a system of ordinary nonlinear differential equations with initial conditions at t=0 – $X_i = X_i^o,\ \overline{N} = 1$ the modified fifth-order Kutta-Merson method was chosen [17].

Algorithm for solving a system of differential equations $\frac{dx}{dt} = f(x)$ requires a fivefold calculation of the right-hand sides f(x):

$F_1=f(x_k)\cdot h,$
$F_2=f(x_k+F_1/3)\cdot h,$
$F_3=f(x_k+(F_1+F_2)/6)\cdot h,$
$F_4=f(x_k+(F_1+3\cdot F_2)/8)\cdot h$
$x_{k+1}=x_k+(F_1-3\cdot F_2+4\cdot F_4)/2$
$F_5=f(x_{k+1})\cdot h$
$x^*_{k+1}=x_k+(F_1+4F_4+F_5)$

where h – the step of integration over the space or time coordinate, x_k – values of variables at the beginning of the integration interval. At the end of the integration interval, two solutions are obtained: x_{k+1} – by the Runge-Kutta method with accuracy $O(h^4)$ and x^*_{k+1} – by the Kutta-Merson method with accuracy $O(h^5)$.

Based on the developed information-analytical system, the stage rate constants and activation energies are found. The values are given in Table 1.

Table 1. Kinetic parameters for the synthesis of benzylidene benzylamine at temperature of 23° C.

	Kinetic constants, h^{-1}	Activation energy, kcal/mol
k_1	1.5×10^{-2}	10.6
k_2	4.7	7.7
k_3	13.4	1.6
k_4	0.6	0.4

For this reaction, the adjusted and true constants are related by the following equations: $K_i=k_i \cdot C_o$ (i=1, 2, 3, 4), k_i ($L\times mol^{-1}\times h^{-1}$); $k_5=K_5$ (h^{-1}).

Table 1 shows the constants at reference temperature T_{ref}=23°C [73]. Recalculation of constants for any temperature is performed based on the selected reference temperature by expressions:

$$K_i(T) = K_i(T_{ref}) \cdot exp\left[\frac{E_i}{RT_{ref}}\left(1-\frac{T_{ref}}{T}\right)\right]; \quad K_i(T_{ref}) = K_i^o \cdot exp\left(-\frac{E_i}{RT_{ref}}\right)$$

To verify the adequacy of the constructed kinetic model, the calculated and experimental data are compared (Fig. 3, Fig. 4).

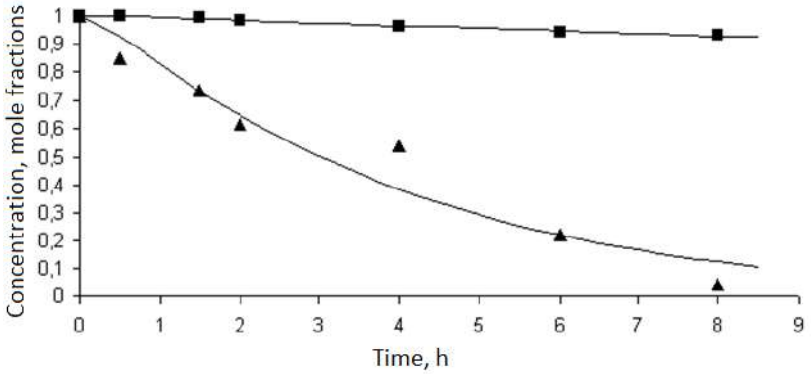


Figure 3. Calculated and experimental data on benzylamine consumption at temperatures: 1 (■) – 23° C; 2 (▲) – 85° C.

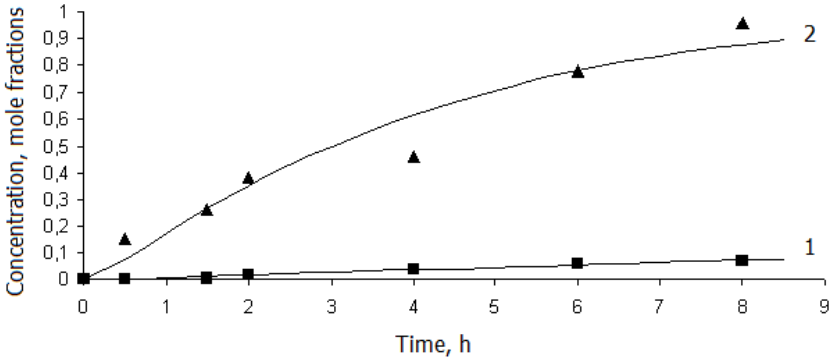


Figure 4. Calculated and experimental data on the formation of benzylidene benzylamine at temperatures: 1 (■) – 23° C; 2 (▲) – 85° C.

The developed kinetic model adequately describes the experimental observations. The divergence between the calculated data and the experimental changes in concentrations over time does not exceed the error in the quantitative analysis.

5. Optimal reaction conditions

Based on the developed kinetic model for the synthesis of benzylidene benzylamine, computational experiments were made. The molar ratio of the initial substances: benzylamine (X_1) and carbon tetrachloride (X_2) was chosen as the optimization criterion. The results of computational experiments on change in yield of the target product depending on the ratio X_1:X_2 at two temperatures are given in Table 2 and Table 3.

Table 2. Dependence of the target product yield on the initial substances molar ratio at 23º C.

Benzylamine (X_1), molar fraction	Carbon tetrachloride (X_2), molar fraction	Benzylidene benzylamine yield (X_7), %mol.
0.45	0.55	6.3
0.40	0.60	6.9
0.37	0.63	7.2
0.35	0.65	7.4
0.30	0.70	8.0
0.25	0.75	8.5
0.10	**0.90**	**9.7**
0.05	0.95	6.3

Table 3. Dependence of the target product yield on the initial substances molar ratio at 85º C.

Benzylamine (X_1), molar fraction	Carbon tetrachloride (X_2), molar fraction	Benzylidene benzylamine yield (X_7), %mol.
0.45	0.55	80.6
0.40	0.60	85.5
0.37	0.63	87.9
0.35	0.65	89.0
0.30	0.70	92.3
0.25	0.75	94.5
0.10	**0.90**	**97.6**
0.05	0.95	97.1

It was found that to reach the maximum product yield, the optimal molar ratio of initial substances is X_1:X_2=1:9.

6. Conclusions

The information-analytical system for modeling chemical technology processes has been developed. The parallel genetic algorithm for solving inverse kinetic problems is implemented. On the basis of information-analytical system the kinetic model for the synthesis of benzylidene benzylamine has been built. The adequacy of the built kinetic model has been established. In accordance with developed model the optimal molar ratio of the initial substances has been found as benzylamine: carbon tetrachloride = 1:9.

7. References

[1] Akhmetov I V and Gubaydullin I M 2016 Analysis of methods for solving inverse problems of chemical kinetics with the use of parallel computing *CEUR Workshop Proceedings* **1576** 402-410

[2] Akhmetov I V and Gubaydullin I M 2017 Parallel calculations in the construction of the kinetic model of benzylidene benzylamine synthesis *CEUR Workshop Proceedings* **1902** 1-4

[3] Akhmetov I V, Bobrenova YU O, Gubaydullin I M and Novichkova A V 2013 Mathematical modeling of complex chemical reactions in the presence of metal complex catalysts on the basis of multi-cycle computing systems *Control Systems and Information Technology* **2.1** 111-115
[4] Akhmetov I V and Gubaydullin I M 2012 Development of kinetic models using parallel computations on multi-core systems *Bulletin of Omsk University* **2** 172-174
[5] Voevodin V V and Voevodin Vl V 2002 *Parallel computing* (SPb: BHV-Petersburg) p 608
[6] Sadovnichy V A, Savin G I and Voevodin Vl V 2009 *Supercomputer technologies in science, education and industry* (Moscow: MSU Publishing House) p 232
[7] Holland J 1975 *Adaptation in natural and artificial systems* (Ann Arbor: University of Michigan Press) p 96
[8] Nikitin A V and Nikitina L I 2002 Evolutionary model for optimizing modular associative memory for data flow machines based on the genetic algorithm *Programming* **6** 31-42
[9] Chernyshev O and Borisov A 2007 Comparative analysis of solving optimization problems by genetic and gradient methods *Transport and Telecommunication* **8** 40-52
[10] Singn G and Zimmer H J 1979 *Org. Chem.* **44** 417
[11] Khusnutdinov R I, Bayguzina A R and Aminov R I Synthesis of N-benzylamine benzilidenbenzilamina under the action of iron catalysts in CCl4 *Russian Journal of Organic Chemistry* **48** 1063-1065
[12] Kauffmann T, Habersaat K and Koppelmann E 1977 *Chem. Ber.* **110** 638
[13] Burdzhiev N T and Stanoeva E R 2006 *Tetrahedron* **62** 8318-8326
[14] Mazitov A A, Osipova A G, Akhmetov I V and Gubaydullin I M 2016 Solution of the inverse problem of chemical kinetics using the example of the benzylidene benzylamine synthesis reaction *Journal of the Middle Volga Mathematical Society* **18** 145-152
[15] Akhmetov I V, Gubaydullin I M, Koledina K F and Safin R R 2015 Mathematical modeling and optimization of reactions of synthesis of aromatic compounds *Electrical and Information Facilities and Systems* **2** 53-58
[16] Khusnutdinov R I, Baiguzina A R, Mukminov R R, Akhmetov I V, Gubaidullin I M, Spivak S I and Dzhemilev U M 2010 New synthesis of pyrrole-2-carboxylic and pyrrole-2,5-dicarboxylic acid esters in the presence of iron-containing catalysts *Russian Journal of Organic Chemistry* **46** 1053-1059
[17] Polak L S, Goldenberg M Ya and Levitskiy A A 1984 *Computational methods in chemical kinetics* (Moscow: Nauka) p 280

Acknowledgments
The reported study was funded by RFBR according to the research projects № 18-07-00341.

Implementation of the finite-difference method for solving Maxwell`s equations in MATLAB language on a GPU

N D Morunov[1]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086

Abstract. The FDTD method of solving the Maxwell equations and methods for specifying a plane wave for it are considered. In addition, the features of implementation of calculations in MATLAB via graphics processing unit were considered. The comparative analysis of the difference solution of the Maxwell equations in the MATLAB language via the GPU for different ways of specifying a plane wave using an inexpensive graphics card is proposed. As a result, it is possible to speedup calculations to 60 times in the MATLAB language via the user video card NVIDIA GTX 960m.

1. Introduction

Numerical methods for solving the Maxwell equations require a large amount of time to solve simple problems of electrodynamics, not to mention complex ones. The most successful reduction in costs over time is parallel programming on GPUs and it is alsocalled general-purpose computing for graphics processing units (GPGPU). The well-known GPGPU programming tools are CUDA, OpenCL and computational shaders. Of these, CUDA is by far the most common and effective in the vast majority of cases [1].

It is necessary to know the C language to use CUDA for the calculations, which is more low-level language than MATLAB language. You have to deal with such difficulties as debugging of undefined behavior and memory leaks for programming mathematical problems in C language. In addition, not all scientists well-know the popular programming language C. The syntax of the MATLAB language is very simple, easy to learn, and is ideal for such scientists. There is a special package for parallel computing in MATLABcalled Parallel Computing Toolbox. It is possible to easily implement calculations on multi-core processors, on clusters and on GPUs via using this package and simple syntax.

The book A. Elsherbeni [2] described in detail the FDTD-method with implementation in the MATLAB language, but the author does not describe the implementation of parallel computations on a GPU in the same language. This was because at the time of writing the book MATLAB could not perform calculations on the graphics processing unit. This gap is filled with this article.

In the current article, we discuss in more detail the methods for specifying a plane wave in MATLAB using an inexpensive custom graphics card NVIDIA GTX 960m. Similar characteristics are GTX series 600-700 series for PC and 800m-900m series for laptops, the price of such a video card is low, about 200$.

2. Organization of parallel computations in MATLAB using a GPU

In additional materials to the book A. Taflove [3] there are scripts written by his assistant S. Hagness in the language MATLAB. There are only three of them: for one-dimensional, two-dimensional and three-dimensional cases of the FDTD method. Calculations in scripts are performed on the central processor. Consider the algorithm in MATLAB language for the three-dimensional case. At each step in time, the calculated values of the electric field components are written into the animation matrix, and then the animation is played using the movie function.

The algorithm of the program in the MATLAB language is presented below:

```
ex=zeros(ie,jb,kb);
…
hz=zeros(ie,je,kb);
for n=1:nmax
ex(1:ie,2:je,2:ke)=ca*ex(1:ie,2:je,2:ke)+...
cb*(hz(1:ie,2:je,2:ke)-hz(1:ie,1:je-1,2:ke)+...
hy(1:ie,2:je,1:ke-1)-hy(1:ie,2:je,2:ke));
…
ez(is,js,1:ke)=ez(is,js,1:ke)+...
srcconst*(n-ndelay)*exp(-((n-ndelay)^2/tau^2));
…
hz(1:ie,1:je,2:ke)=hz(1:ie,1:je,2:ke)+...
db*(ex(1:ie,2:jb,2:ke)-ex(1:ie,1:je,2:ke)+...
ey(1:ie,1:je,2:ke)-ey(2:ib,1:je,2:ke));
end;
```

Listing 1. Algorithm for the 3D case of FDTD method in MATLAB language.

In the MATLAB Parallel Computing Toolbox, there are three ways to implement parallel computations on a GPU: using a new type of gpuArray variable and various functions defined for this type; using arrayfun and bsxfun functions, which allow to define own function for processing GPU arrays; writing its CUDA-kernel and using it in calculations.

First of all, in order to remake the MATLAB code written for the CPU in the code written for the GPU, you need to convert all large arrays to the gpuArray type, you can do this either by explicitly casting the type to gpuArray, or by creating the array directly in the graphics card's memory (Listing 2). To transfer an array from the internal memory of the GPU to the RAM should be use the gather function. However, to display the array as a graph, this operation is not required. When drawing graphs, a set of vertices that define geometric primitives must be in the memory of the video card, so it is not meaningful to make data exchange.

```
ex=zeros(ie,jb,kb,'gpuArray');
…
hz= zeros(ie,je,kb,'gpuArray');
```

Listing 2. Declaring arrays for electric and magnetic fields directly in graphics memory.

Particularly useful functions are arrayfun and bsxfun, which very similar to the CUDA kernels in their work, the difference is the strict correspondence of the thread number to the indexes of the input arrays. The most effective way to implement parallel computations on a GPU in MATLAB is to reduce multiple synchronous operations over arrays to one function, where these operations are performed asynchronously, and using it in arrayfun or bsxfun. In addition, the bsxfun function is ideal for reduce large numbers of operations between small-sized vectors, for example, this technique is used to optimize the calculation of CPML.

```
ex(il:ir-1,jl:jr,kl) = bsxfun (@plus,ex(il:ir-1,jl:jr,kl), cb*hy_1D(il:ir-1));
ex(il:ir-1,jl:jr,kr+1) = bsxfun(@minus,ex(il:ir-1,jl:jr,kr+1), cb*hy_1D(il:ir-1));
```

Listing 3. The update of the electric field at the total field (TF) and the scattered field (SF) boundary using bsxfun.

```
Psi_eyx_xn(1:L,1:je,2:ke) = bsxfun (@times,Psi_eyx_xn(1:L,1:je,2:ke),cpml_b_en(1:L))+...
bsxfun (@times,hz(1:L,1:je,2:ke)-hz(2:L+1,1:je,2:ke),cpml_a_en(1:L));
ey(2:L+1,1:je,2:ke) = ey(2:L+1,1:je,2:ke) + cb_psi(1:L,1:je,2:ke).*Psi_eyx_xn(1:L,1:je,2:ke);
Psi_ezx_xn(1:L,2:je,1:ke) = bsxfun (@times,Psi_ezx_xn(1:L,2:je,1:ke),cpml_b_en(1:L))+...
bsxfun (@times,hy(2:L+1,2:je,1:ke)-hy(1:L,2:je,1:ke),cpml_a_en(1:L));
ez(2:L+1,2:je,1:ke) = ez(2:L+1,2:je,1:ke) + cb_psi(1:L,2:je,1:ke).*Psi_ezx_xn(1:L,2:je,1:ke);
```

Listing 4. Updating Ex CPML components using bsxfun.

MATLAB supports a set of functions for working with CUDA directly: loading the kernel, changing kernel parameters, etc. This method involves writing a kernel in C (the actual procedure that will be performed on the GPU) and connecting it to the MATLAB application. Next, you need to configure the loaded kernel: specify the number of maximum possible parallel threads; specify the number of blocks, and execute the kernel procedure.

The third method though is the best, but the task of the article was to implement the finite-difference solution using MATLAB language. Moreover, it is necessary to write CUDA kernels in the C language before connectthem to MATLAB.

MATLAB is more suitable for vector calculations, so whole code should be vectorized at first. When working with Parallel Computing Toolbox, another one feature was noticed: if a size of the vector is small (about 10 elements), then this vector should not be transferred to the graphics card's memory, since in some cases this will only slow down the work. Moreover, if over a large number of iterations in the algorithm, only one element of the vector is used, and its values doesn`t change during these iterations, it is worthwhile to cut this value from the vector and transfer it to the local work area.

MATLAB uses memory both for storing raw data and result data, and for storing intermediate data used in calculations. Thus, the maximum number of nodes can be defined as:

$$I_{\max} = \frac{V_{\max}}{v(a_{st} + a_{comp})},$$

where V_{max} – available memory capacity; v –amount of data consumed by a single node value; a_{st} – number of arrays using by store data; a_{comp} – number of arrays using by compute data.

Graphical and operative memory are allocated in different ways in the same cases. For example, copies of arrays are created with using RAM, whenever there is a possibility to use the same memory cell two or more times for calculating one expression, i.e. the presence of conflicts when accessing memory. Therefore, the array is duplicated when using indexes and when using non-element operators (*, ^). In the case of graphics memory, arrays are duplicated usually for each individual argument and the return value in the expression. In addition, there is one general rule that can avoid copying an array: if the return value is contained as an argument in the expression, then this argument is not copied.

For example, consider the following expression written in MATLAB language:

```
A = ca.*A + cb.*(B(2:end)-B(1:end-1)).
```

We will assume that the expression is written correctly and the sizes of all arrays used as arguments coincide. If all the arrays are in RAM, then in the process of calculations the program in MATLAB language will copy array B two times. If they are in graphics memory, then they will copy 5 arrays. As a result, the graphics memory will need almost 1.5 times more.

When performing parallel computations on a GPU, MATLAB automatically optimizes the calculations in the cycle, allocating more graphics memory. If the memory enough for arrays copies for all expressions in the loop, then MATLAB additionally parallelizes the calculation of the expressions. In this case, you can set an effective amount of graphics memory and the maximum amount of graphics memory.

In accordance with all the above listed features of the Parallel Computing Toolbox, the algorithm of C. Hagness (Listing 1) was adapted to run on the GPU.

3. Experimental research of FDTD implementation in MATLAB language

An experiment was conducted to investigate the effectiveness of the implementation of the 3D case of the FDTD method.

In the experiment were used an NVIDIA GTX 960m graphics card with GDDR5-2Gb internal memory and a number of CUDA-compatible shader processors - 640 with a base clock speed of 1.1 GHz. Moreover, a 4-core Intel Core i5-6300HQ CPU with a base clock speed of 2.3 GHz and a maximum frequency of 3.2 GHz, with a cache size of 6 MB. Also on the laptop was installed 8GB of RAM type DDR4-2133 andnorth bridge - Intel Skylake-H IMC.

The next step was to develop a other script for profiling the work of the computational process on different processors. Based on the algorithms considered earlier, Functions were written to perform calculations on the CPU and on the GPU. The input of this function is given the size of the grid area and the number of iterations, and the output is the execution time of the calculations.

Consider the profiler code in a simpler form:

```
c_min = 30;c_max = 250;dc = 10;nmax=6;
T = (c_max-c_min)/dc+1;
cells = c_min:dc:c_max;
fori = 1:T
    cpu_time(i) = fdtd3D_cpu(cells(i), nmax);
    gpu_time(i) = fdtd3D_gpu(cells(i),nmax);
end;
speedup = cpu_time./gpu_time;
cells_axis = cells.^3 / 10^6;
cpu_perfomance_axis = cells.^3 ./ cpu_time / 10^6;
gpu_perfomance_axis = cells.^3 ./ gpu_time / 10^6;
figure;
subplot(3,1,1);
plot(cells_axis, cpu_perfomance_axis);
subplot(3,1,2);
plot(cells_axis, gpu_perfomance_axis);
subplot(3,1,3);
plot(cells_axis, speedup);
```

Listing 5. Profiler code in the MATLAB language.

Four combinations of initial and boundary conditions were considered and for each of them two functions fdtd3D_cpu and fdtd3D_gpu were implemented (Listing 5). A source that generated a plane wave in two ways, TFSF and TFRF [4] was given as initial conditions. Perfect electrical conductor (PEC) boundary conditions is defined along the axis of propagation and cyclic boundary conditions are given for other directions for two cases with the TFRF method.Perfect electric conductor (PEC) boundary conditions was defined in the other two cases with the TFSF method. In addition, Convolutional Perfectly Matched Layer (CPML) [2] was defined for each method of specifying a wave.

As a result, of the profiling script execution, plots were obtained. Plots show the dependence between calculation rate and number of cells. Calculation rate is a millions of processed cells per second. Figure 1 shows the characteristics of a computational process that calculates the propagation of a plane wave in a computational domain bounded by a perfect electric conductor boundary without the use of CPML, and in Figure 2 with CPML. As a result, the plot with dependence between speedup and number of cells for each case was calculated and presented in Fig. 3.

There is sharp decrease in the calculation rate on the plots (Figures 1b and 2b) for GPU after some value of the grid size. This is due to the lack of video memory on the graphics card, and the value corresponds to effective memory. On the plots of dependence between CPU calculation rate and the number of nodes (Figures 1a and 2a), a sharp jump is seen at the very beginning. This is because the amount of data necessary for the computing task is placed in the central processing unit’s cache and the access speed to the memory is high. The amount of data required for the computational task

increases and number of cache misses increases too. Moreover, speed of access to memory decreases, as well as the performance of computations.

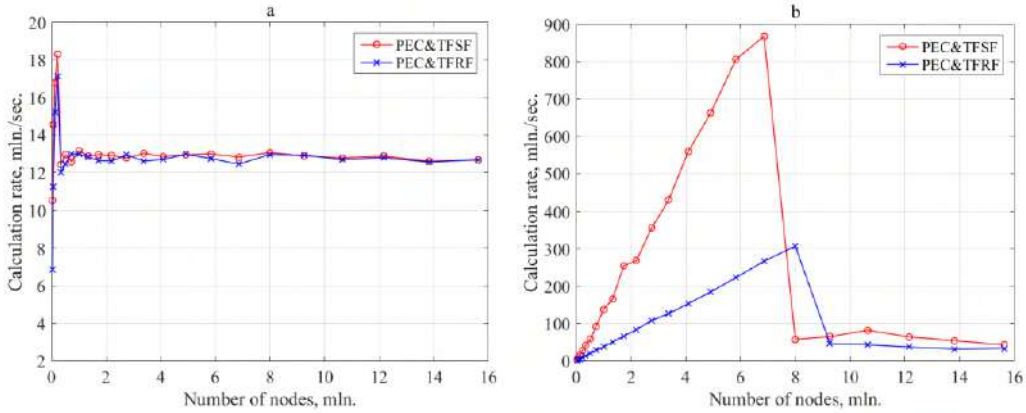


Figure 1. Plots of dependence between calculation rate (million nodes / sec.) and the number of processed nodes (million) withoutCPML: a) on the CPU; b) on the GPU.

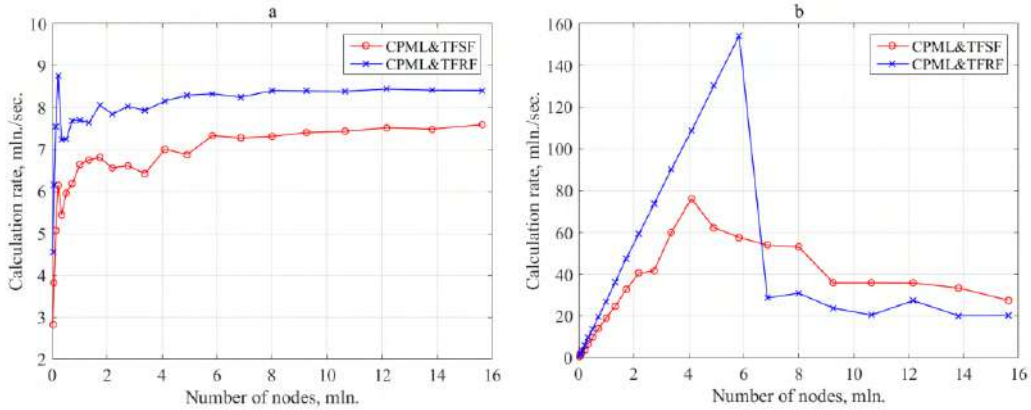


Figure 2. Plots of dependence between calculation rate (million nodes / sec.) and the number of processed nodes (million) with CPML: a) on the CPU; b) on the GPU.

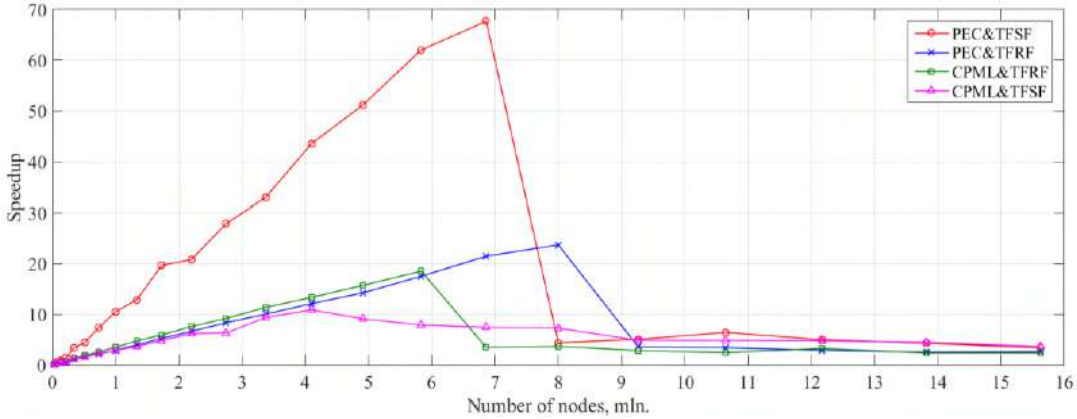


Figure 3. Plots of dependence between speedup and the number of processed nodes (million).

The calculation rate increases well when using the GPU in both cases with TFRF (Figures 1 and 2). However, there is a calculation rate decline when using CPML in cases with TFSF. This behavior can be explained by the following: the boundary conditions in TFSF (an ideal electrical wall) are set easily without adding extra vectors, but CPML is set on each of the 6 faces; TFRF is given periodic boundary conditions, which increases the number of vectors, and CPML is only set on 2 faces. The use of CPML in calculations creates a greater number of small vectors, larger the dimension of the field to

which these perfect matched layer is set. Accordingly, the volume of parallel computations decreases because of the large number of individual vectors.

4. Conclusion

Let's sum up the analysis. It should be noted that if it is required to solve the computational problem of electrodynamics, where a plane wave is used as a source, it is preferable to use the TFRF method. In this paper we obtained the results (Figure 3), based on which it can be stated that it is possible to perform computational tasks of electrodynamics on inexpensive video cards with speedup 20 times using CPML and 60 times without it.However, another problem arises: the limited amount of graphics memory in comparison with the operating memory. Solve this problem is decomposing, splitting the range of values into blocks of such a volume, in which the computational process has the maximum acceleration.

5. References

[1] Golovashkin D L and Kasanskiy N L 2011 Solving diffractive optics problem using graphics processing units *Optical Memory and Neural Networks (Information Optics)* **20(2)** 85-89
[2] Elsherbeni A Z and Demir V 2008 *The finite-difference time-domain method for electromagnetics with MATLAB simulations* (Raleigh: SciTech Publishing) p 426
[3] Taflove A 1995 *Computational Electrodynamics: the finite-difference time-domain method* (London: Artech House) p 599
[4] Soifer V A 2014 *Diffractive nanophotonics* (CRC Press) p 704

Numerical simulation of the ion focusing process in a dust impact time of flight mass spectrometer

I V Piyakov[1], D V Rodin[1], M A Rodina[1] and A M Telegin[1]

[1]Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. The article presents the description of a design of dust-impact time of flight mass-spectrometer. A method for calculating the trajectory of a charged particle on a triangular grid is considered. The implementation of the calculation algorithm for the personal computer, the GPU accelerator and the supercomputer is described. The main differences in implementations, comparison of performance and accuracy of calculation are given. The advantages and disadvantages of each implementation are considered.

1. Introduction

Property analysis of the micrometeoroid and debris particles *in-situ* is impossible without the use of appropriate recording equipment. Thus, for the analysis of the chemical composition of particles, dust-impact time-of-flight mass spectrometers are used. The advantage of this type of analyzers is high speed, the ability to register single events, small size and power consumption. A unique feature of dust-impact time-of-flight mass spectrometers is ionization by a high-velocity collision of a particle-striker and target, which occurs randomly at an arbitrary point of the target. An increase in the probability of ionization in this way is achieved by increasing the active area of the target, which leads to the need for additional spatial focusing of the ions obtained using parabolic reflectors or electrostatic reflectors with a special form of electric field. The plasma produced as a result of a high-velocity collision has a high energy spread, which necessitates the use of electrostatic ion mirrors that provide a temporary focusing of ion packets.

The principle of operation of time-of-flight devices is based on measuring the time of flight of particles with the known energy value from the ion source to the receiver. In an ideal case, the time of flight depends only on the mass of the ion. The development of real devices is based on the determination of design parameters that ensure maximum independence of the time of flight from the place of collision to the ion receiver from the ion initial energy. For this purpose, a one-dimensional field structure [1] is calculated that provides the required temporal focusing and the requirements for design elements are developed by numerical simulation. It allows to form field structure close to the desired in the maximum volume of the device structure. Verification of each of the design options is carried out by simulating the trajectories of model ion packets, with the subsequent calculation of the resolving power. In this paper we consider the software that implements a massively parallel algorithm for calculating particle trajectories on a triangular grid, which provides the possibility of modeling curvilinear structural elements.

2. Formulation of the problem

The structure diagram of the device under consideration is shown in Figure 1. A detailed description of the device is given in [2].

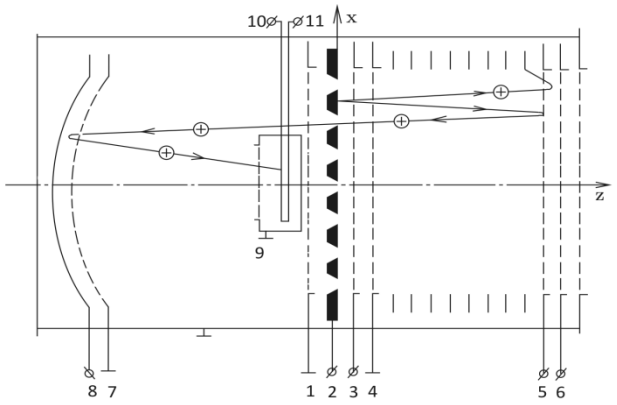


Figure 1. Dust-impact mass spectrometer with a parabolic reflector (2 – target; 3 – accelerating grid; 4 – retarding grid, grounded; 5, 6 – mesh of the ion mirror, 1 – grounded grid; 7, 8 – parabolic reflector, 10, 11 – ion receiver; 9 – receiver casing, grounded).

The device works as follows. When the particle strikes the target, the particle substance is ionized, accelerated by an electric field and enters the ion mirror, bounded by grids 4, 5, 6, where it is reflected and, passing back through the target (by this moment the accelerating potential is turned off), flies through the fieldless space between the meshes 1 and 8, when ion packet is reflected in a parabolic reflector and enters the receiver input 10, an ion pulse is formed at its output 11.

A special feature of the reflector is the shape of a paraboloid with a focus in the ion receiver. This facilitates additional spatial focusing of ion beams. Thus, the loss of ions decreases, leading to increase of the device sensitivity. The energy focusing of ions is performed by the ion mirror.

The main parameters of the time-of-flight mass spectrometer are its resolving power and sensitivity. At the same time, their values depend on many factors:

– design parameters of a time-of-flight mass spectrometer;
– the energy spread of ions formed as a result of high-velocity collision;
– direction of the initial ion velocity vectors;
– coordinate of the particle impact on the target;
– chemical composition of particles

To determine the influence of these factors on the characteristics we are interested in, we need to calculate the trajectories of ions in the device path. For this purpose, for the given mass spectrometer structure, the following parameters are specified: the distances between the grids and the accelerating potential, the target and ring dimensions, the curvature of the parabolic reflector, the diameter of the central target hole, the parameters of the field-setting rings. The grid with the distribution of electrostatic potential inside the device path that is used on the stage of modelling the operation of the device is preliminary calculated.

3. The calculation algorithm

Numerical simulation of the motion of charged particles can be implemented both with the use of a grid of field values obtained by the method of finite differences, and with the use of nodal values obtained by the finite element method. The advantage of the first method is the clarity, simplicity of implementation of the cell search functions and cubic interpolation, ease of mesh joining when calculating in several areas. However, an essential disadvantage of this type of grids is the impossibility of locating nodes on curvilinear surfaces. The finite element method allows you to divide arbitrary regions into separate elements (in certain cases, triangles), which allows to create an exact reconstruction of the shape of the regions of interest. At the same time, the use of finite element grids requires the implementation of additional functions, for example, the search for the triangle to which the point belongs, as well as the more complex than in the finite difference method data structures for storing grid information, which is a graph with a description of nodes and triangles.

Based on the data on the interaction coordinates and velocity, an ion packet with a known value of the energy spread is formed [3]. For each ion in the ion packet with a given time step, the increment of coordinates is calculated. After each iteration, a check is made for impact with the structural element, escape and detection. Times of flights of the detected ions are conserved. To estimate the design characteristics for each mass, it is required to calculate the number of ions, to estimate the collection

coefficient, the mean time of flight, the standard deviation and the resolution of the instrument on two sigma level.

Step by step algorithm of the program on a triangular grid can be described as follows.

1. Download the data about the calculated grid with the field from the file.
2. File parsing - linking nodes, triangles, regions and field values.
3. Creating a Mesh object with a description of the relationships between nodes and triangles.
4. Formation of a model ion packet, starting from the given coordinates and the speed of interaction.
5. Finding the current triangle for each ion.
6. Interpolation of the field at the location of each ion.
7. Calculation of the displacement of each ion in time d*t*.
8. Checking the location of the ion in the current triangle.
9. Recursive search for a new triangle.
10. Check for escape / detection - return to point 6.

In general, the implementation structure is similar to the one described in [4]. So in step 4 the Box-Muller algorithm is similarly used for the formation of a packet with a Maxwellian velocity distribution. The ion displacement can also be calculated by the trapezoidal rule or by the Runge-Kutta method. The difference between this implementation is the use of the bilinear interpolation method for a triangular grid and the need to check whether the coordinates of the particle belong to the current triangle. Both problems are solved using the known approaches with small modifications. To accelerate the interpolation operation, each instance of the triangle class stores the previously calculated interpolation coefficients *s*1, *t*1, *s*2, *t*2, *s*3, *t*3, *A*, coeff and to speed up the search for a new triangle to which the particle coordinates belong. The same instance stores references to neighbouring triangles, which allows to replace the search for all grid triangles by a recursive search from the current triangle.

4. Features of the algorithm implementation for calculating the trajectories for CPU, GPU, HPC

The implementation for the CPU is written in C# and contains four main classes for implementing trajectory calculation. The Point class is used to store graph nodes, contains the node index, coordinate information, field values, and the number of triangles to which this node belongs. The Triangle class contains the index of the triangle, three references to the nodes-vertices of this triangle, information about the domain to which the triangle belongs, three references to adjacent triangles, pre-calculated coefficients for interpolation, and methods for bilinear interpolation of the field and the search for a new triangle when the point doesn't belong to the current triangle. The Mesh class contains arrays of nodes and triangles, as well as the methods for finding the nearest point, triangle, reading and writing files. The Particle class contains information about the mass and charge of a particle, its speed, coordinates, the current triangle, and also a method that implements an iterative calculation of the trajectory on a triangular grid.

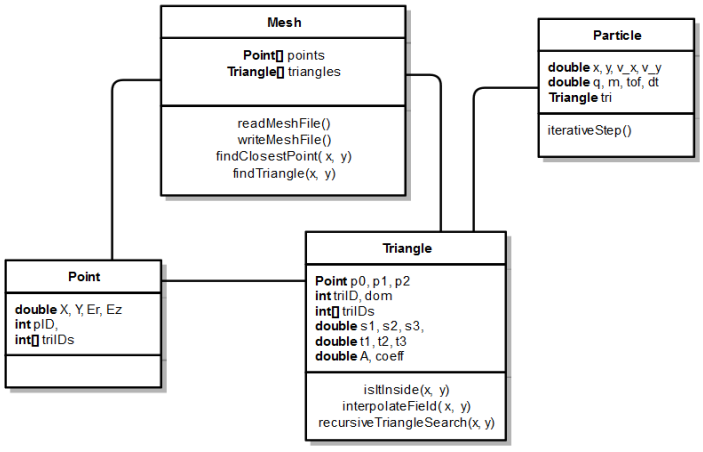


Figure 2. Class structure for storing information about the calculated grid, PC implementation in C#.

The use of a graphic accelerator imposes severe restrictions on the implementation of software, leading to the need, if possible, to represent all information in the form of arrays. Thus, the implementation for the GPU contains the following arrays: triangleNeighbours, trianglePoints - arrays of vector format uint4 for storing information about neighbouring triangles and triangle vertices. Arrays s1_t1, s2_t2, s3_t3, A_coeff of vector format double2 are used for storing information about the coefficients for interpolating the field. Arrays pointsCoordinates and pointsFields of vector format double2 contain information about grid nodes. Arrays particlesCoordinates and particlesSpeeds vector format double2 are used for storing information about the coordinates and velocities of particles. An array of particlesTimes of double1 format is used for storing the current time of the particle. An array of particlesTriangles of the uint1 format contains the index of the current triangle of the particle. The index in the array is used to refer to nodes, triangles, and particles instead of explicit indexes. These arrays are created in the host memory and populated with data from the CPU implementation, after those arrays are created in the GPU memory and data from the host arrays is copied to the GPU. Then, the texture and surface links are bound to the GPU arrays, which allows using the operations of texture samples that implement hardware caching. It should be noted that linking texture references to arrays of vector double types is impossible, so when copying data from the CPU to the GPU, a bit conversion of one double number to a pair of int numbers takes place. After the texture sample, a reverse bit conversion occurs.

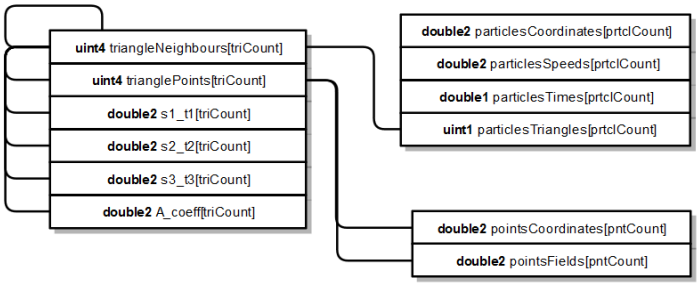


Figure 3. The structure of arrays for storing information about the calculated grid, GPU implementation in the CUDA language.

Implementation for the supercomputer Sergey Korolev differs both from the version for the PC, and from the version for the GPU. Since this implementation is written for the MPICC compiler, and the version of the C language that does not support OOP is used. The data about triangles and their interrelationships and the relations with nodes are stored in triangleNeighbours, trianglePoints of int format, data on coefficients for interpolation in triangleCoeff array, double format , the coordinates of the nodes and field values in them - in the pointsCoordinatesAndFields array of the double format, the particle data in the particles array of the double format. This implementation does not require additional type conversions when accessing memory, all arrays also lack the index of nodes, triangles and particles, instead of them an index in the array is used.

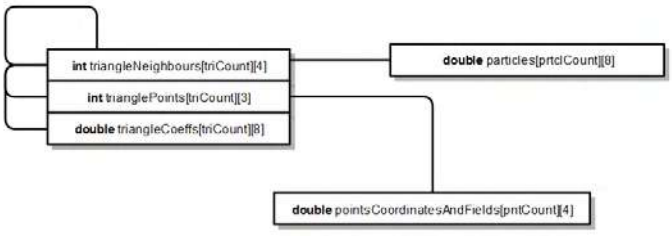


Figure 4. Structure of arrays for storing information about the calculated grid, HPC implementation in C language.

5. Results of calculation and performance comparison

To visualize the modelling process and evaluate the performance of the design, a visualization module was written to depict the movement of ion packets in the volume of the structure. Figure 5 shows several combined frames, the focusing of ion packets over time is clearly visible - two masses are well separated at the receiver plane.

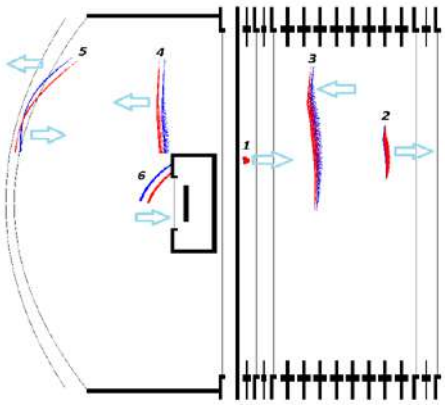


Figure 5. Visualization of the trace of ion packets. A packet of two masses is formed at point 1, accelerates and flies into the ion mirror 2, reflects and flies back to the target 3, flies through the target, part of the ions settles on the receiver shell 4, is reflected in the parabolic reflector 5 and flies into the receiver 6.

The results of modelling the construction of a time-of-flight mass spectrometer with a linear ion mirror and a parabolic reflector are presented in Figures 6 and 7. Calculations were made for 16382 particles without taking into account the transparency coefficients of the grids and the target.

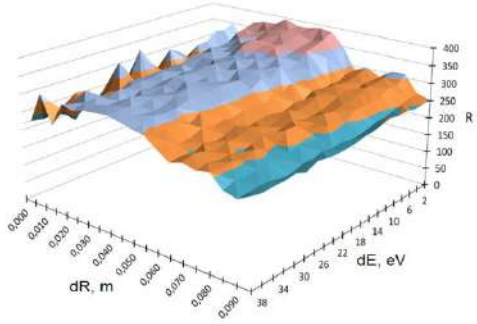


Figure 6. Dependence of the resolving power of the instrument on the collision coordinate and the energy spread.

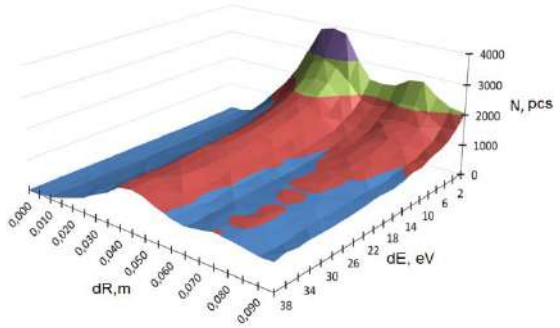


Figure 7. Dependence of the number of received ions on the collision coordinate and energy spread.

The resolving power has a local maximum at zero radial offset. The ion collection coefficient has a maximum of 35 mm and 75 mm radial displacement with minimal energy spread.

Comparison of implementations for a personal computer, accelerator GPU and supercomputer was performed for a different number of particles. The comparison did not reveal any differences in the accuracy of calculations, which is generally explained by the fact that all calculation modules are implemented using variables of the double type. HPC implementation was launched on four processors.

The results of measurements of the execution time of 10,000 iterations are given in Table 1.

6. Conclusion

Using the above mentioned software, the authors simulated various designs of the dust-mass spectrometer. Simulation of a design with a linear mirror and a parabolic reflector showed a relatively low resolving power of about 250-300 units regardless of the energy spread and the collision coordinates. However, for a combination of factors, such as the achieved resolution, ion collection coefficient, ease of adjustment, this design is more preferable than others. Thus, the ion collection coefficient drops sharply in designs with a flat reflector. In constructions with a nonlinear ion mirror, it is not possible to achieve independence of the resolving power from the collision coordinate.

The presented software provides the possibility of modelling structures with axial symmetry, containing curvilinear electrodes and non-uniform electric fields. In addition to modelling point single-shot sources, such as laser or high-speed impact ion sources, it is also possible to simulate the trajectories of charged particles in various instruments forming ion fluxes, such as gas mass spectrometers with continuous ionization or gas-discharge sources of off-electrode plasma [4, 5].

In contrast to using rectangular meshes [6, 7], the implementation of the accelerator for the GPU (2,880 cores per 1020 MHz) did not show a significant increase in performance, which is explained by a higher number of branches, during the execution of the algorithm that implements the calculation on a triangular grid. However, even such a modest increase in the speed of execution can be useful in solving design optimization problems that require multiple tracing of ion packets, with a total number of estimated passes reaching several tens of thousands, and a total execution time of several days or weeks.

The implementation for HPC, proved to be more efficient with the same number of processors, which can be explained by greater memory performance of the binary code produced by the MPICC compiler. Due to the fact that this task is parallel to the initial data, performance can be increased by increasing the number of compute nodes.

Table 1. Results of measuring the execution time of 10,000 trajectory iterations for implementation on PC, GPU, HPC.

	PC parallel. (*4CPU*, 8 threads)	*PC Single thread*	GPU Host calls	GPU Device calls	HPC Sergey Korolev
16384	7с	24с	6с	6с	6с
65536	29с	1м 34с	23с	23с	18с
262144	1м 53с	6м 19с	1м29с	1м29с	1м19с

7. References

[1] Semkin N D, Piyakov I V, Rodin D V and Pomel'nikov R A 2012 Analytical method for computing the electrostatic field distribution in the deflector of the time-of-flight mass spectrometer *Technical Physics. The Russian Journal of Applied Physics* **57(10**) 1400-1405 DOI: 10.1134/S106378421200192

[2] Semkin N D, Piyakov I V, Rodin D V and Pomelnikov R A 2012 An onboard dust-and-dust mass spectrometer for studying the elemental composition of micrometeoroids *Nauchnoe priborostroenie* **22(3**) 13-20

[3] Hornung K, Malama Y G and Kestenboim K S 2000 Impact vaporization and ionization of cosmic dust particles *Astrophysics and Space Science* **274**(1-2) 355-363

[4] Podlipnov V V, Kolpakov V A and Kazanskiy N L 2016 Etching of silicon dioxide in off electrode plasma using a chrome mask *Computer Optics* **40(6)** 830-836 DOI: 10.18287/2412-6179-2016-40-6-830-836

[5] Kazanskiy N L, Kolpakov V A and Podlipnov V V 2014 Gas discharge devices generating the directed fluxes of off-electrode plasma *Vacuum* **101** 291-297

[6] Piyakov A V, Rodin D V, Rodina M A andTelegin A M 2017 Numerical simulation of motion of dust particles in an accelerator path *CEUR Workshop Proceedings* **1902** 55-61

[7] Semkin N D, Rodin D V and Piyakov I V 2013 A method of calculating the potentials of the field-setting elements of a nonlinear ion mirror *Nauchnoe priborostroenie* **23(3**) 69-75

Simulation of the control system of the electrodynamic accelerator of dust particles

A V Piyakov[1], D V Rodin[1], M A Rodina[1], A M Telegin[1] and S N Kondratev[1]

[1]Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. The device and control system of the accelerator of charged micron particles for simulation of micrometeorites and technogenic particles in laboratory conditions are considered. The model of the accelerator control system operation is described, the results of the accelerator control system functioning simulation for various operating modes are given. The results of comparison of mathematical simulation with experimental data are presented.

1. Introduction

Recently, there has been a trend of increasing concentrations of high-speed technogenic dust particles in near-Earth orbits. According to different sources, the concentration of technogenic dust particles already is three times higher than the concentration of natural origin micrometeorites. Considering the increasing demand for the reliability and durability of spacecrafts(SCs), and as well as new materials production, there is a demand for further research on the interaction of high-speed dust particles with materials of spacecrafts elements [1-3].

Currently new materials have been produced and applied in the space industry, which makes it necessary to conduct impact experiments. Conferences are regularly held on the problem of protecting spacecraft construction materials from micrometeorites. The technique of high-speed throwing is of interest not only from the point of view of modeling the factors of the space environment on the materials of spacecraft structure, but also from the point of view of feeding the thermonuclear reactors with fuel. Thus, there is an obvious need is to construct various accelerators that cover the entire range of masses of the required particles.

There are various ways to accelerate dust particles in the laboratory to simulate the interaction of micrometeorites and technogenic particles with the materials of SC construction, and to create and calibrate new micrometeorite sensors. The type and construction of the accelerator is determined by the problem posed and depends on the range of the analyzed masses and the velocities of the accelerated particles.

The main element of the dust particle accelerator is the control system, which forms accelerating voltages on the drift tubes. To solve the problem of the voltages formation accuracy, different measures can be taken: for example, in [4], the authors use the modification of a linear electrostatic accelerator, which exclude all particles with a specific charge different from 30 coulomb per kilogram from the acceleration process. Other research teams [5] use accelerators based on Van de Graaf high-voltage sources, which do not require complex control schemes, but are too complicated.

Thus, the development of high-precision control systems is essential for the construction of compact accelerators for applied research. This article concerns the operation of the electrodynamic

accelerator of dust particles control system in nonstandard operation modes, caused by the dust contamination of the accelerator internal path.

2. The structure of the particle accelerator

We developed and constructed an accelerator [3, 6-8] of high speed particles, which allows to simulate impacts of micrometeorites with a size of 1 ÷ 10 um and speeds of 1-15 km / s with materials of spacecraft structures under laboratory conditions.

Figure 1 shows a block diagram of a linear accelerator for simulation of micrometeorites, the path of which consists of an injector, linear electrostatic accelerator (LESA), linear electrodynamic accelerator (LEDA), three measuring lines and an experiment chamber. Vacuum in the system is provided by two pumping systems consisting of a vacuum diffusion pump AVMD -250, a forvacuum pump NVPR-16 and the necessary valves (KVE-63). The limiting residual pressure in the system is 10^{-5} mm Hg.

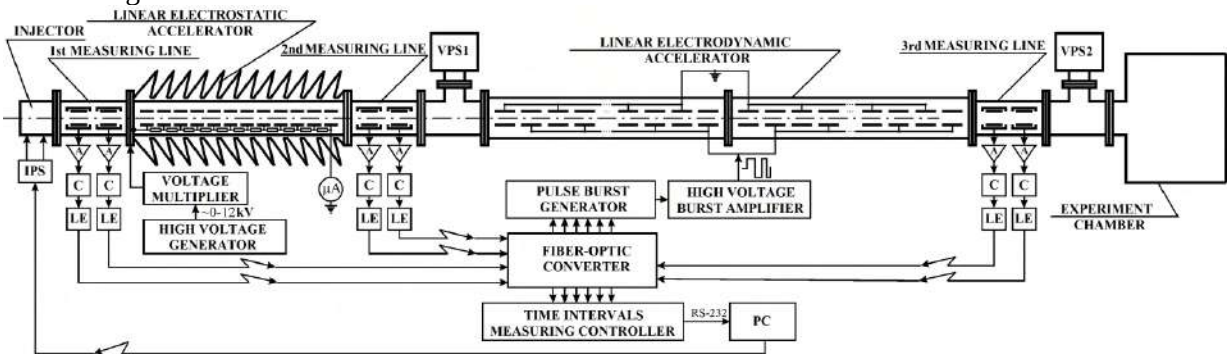


Figure 1. Structural diagram of the linear accelerator of micrometeorites. A – amplifier; C – comparator; LE – laser emitter; VPS1, VPS2 – vacuum pumping system 1 and 2; IPS– power supply of the injector.

The accelerator works as follows. The injector generates charged particles in a given mass range with the frequency of the order of 1 Hz. The charged particle sequentially passes the first measuring line, a linear accelerator, a second measuring line, cylindrical electrodes, the third measuring line and hits the target. The first pair of measuring lines and the linear accelerator are used to determine the particle parameters (specific charge Q / m and initial speed V0). Flying through the measuring line, the particle induces potential to two tubes working as induction sensors.The potentioa sign is opposite to the charge of the particle. Since the tubes are made of metal, their surface is equipotential, which means it does not matter in which part to measure the voltage. According to incoming signals from the measuring line, the speed selector and the selector of specific charges form a digital code of the initial velocity of the particle at their outputs as well as the specific charge code. In the velocity selector, the time intervals of a particle path between two sensor centers for the first and second induction sensors are measured. The measured time intervals are directly proportional to the velocity of the particle. After passing through the linear accelerator, the particle receives a velocity increment. The second measuring line works similarly to the first one. According to the initial velocity and specific charge codes supplied to the variable frequency and duration pulse generator, the burst of voltage impulses is formed on its output which creates accelerating field between each pair of electrodes. This field varies in time according to the position of the particle in the accelerating path. Parameters of the burst are selected from a series of pulse-forming data loaded from a PC into the variable frequency and duration pulse generator. The variable pulse duration amplifier transmits the pulses generated in the variable frequency and duration pulse generator. The amplified pulses are applied to cylindrical electrodes. The third measuring line is connected to the interface unit and serves to obtain the output data. Then the accelerated particle hits the target and the whole process repeats. The computer produces experimental statistics and controls the accelerator dynamically.

3. Description of the electrodynamic accelerator control system

To ensure the functioning of the electrodynamic part of the accelerator of micrometeorites, a control system has been developed, the functional diagram of which is shown in Figure 2.

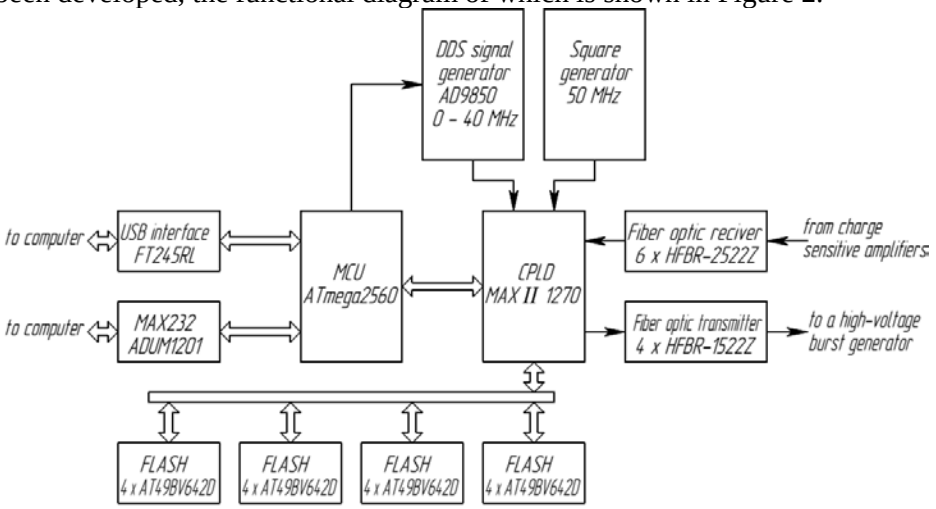


Figure 2. Functional diagram of the electrodynamic accelerator control system.

The control system works as follows. FLASH memory stores the time intervals between voltage changes on the drift tubes of a linear electrodynamic accelerator. The incoming signals from optical receivers to the CPLD MAXII 1270 contain the particle velocity before and after the linear electrostatic accelerator. The speed codes are sent to the address bus of the FLASH memory to select the desired burst of pulses corresponding to the accelerated particle. The counter implemented in CPLD starts counting clock pulses of the DDS generator. When the number of pulses on the counter coincides with the number in the FLASH memory, pulses are formed on the optical transmitters. The microcontroller serves for transferring the data about the system operation to the computer. A high-speed USB interface is used for reprogramming the memory, however, due to interference signals, this interface is not suitable for data transfer during the operation of the accelerator. Therefore, in the control system, a more noise-protected optically isolated RS-232 interface was used.

When the internal path is contaminated with metallic dust particles, the maximum accelerating voltage of the linear electrostatic accelerator is reduced. Thus, the cleaning of the accelerator path is required (Figure 3) [7].

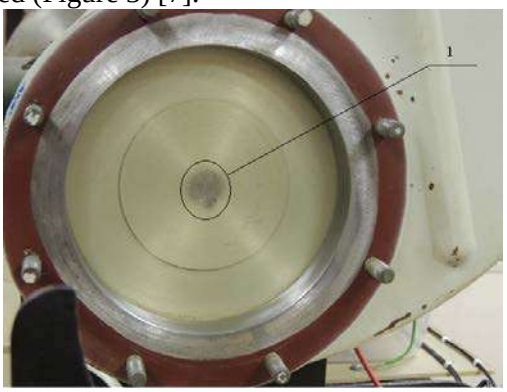


Figure 3. Disassembled accelerator path (valve in a vacuum chamber). 1 – area of microparticle dispersion in the accelerator internal path.

This time-consuming process requires disassembling the vacuum system making it necessary to provide the accelerator operation mode with a reduced accelerating voltage. This operation mode is possible when frequency of the clock generator aimed for generating pulses on the drift tubes of the electrodynamic accelerator is adjustable. For example, when operating at a voltage of 80 kV instead of the calculated 100 kV, frequency tuning from the calculated 10 MHz to 12.403 MHz is required. Let us consider the results of an experiment conducted with the reduced voltage.

Figure 4 shows the time dependence of the detection of particles on the fifth induction sensor on the specific charge of the particles. The dots represent the experimentally measured data, the solid line represents pre-calculated times embedded in the FLASH memory of the generator.

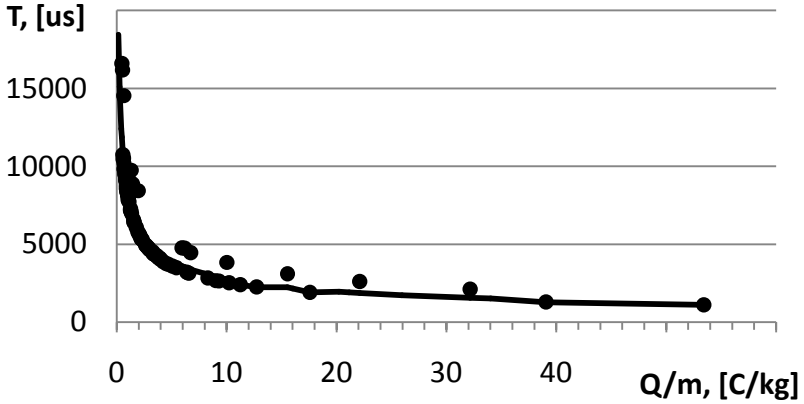


Figure 4. Dependence of the registration time of a particle on the fifth induction sensor on the specific charge.

Figure 5 shows the dependence of the particle velocity on the specific charge. Points stand for experimental data, solid line represents the calculated speed. The points that much higher or lower than the calculated curve represent the particles for which the synchronization of the accelerating pulses is lost.

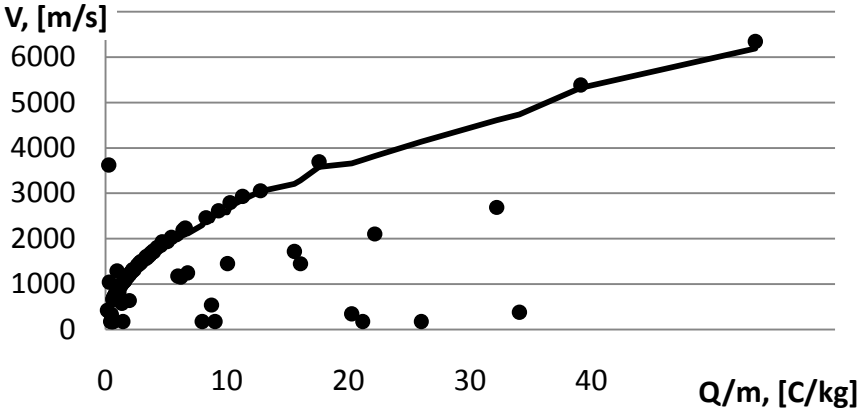


Figure 5. Dependence of the particle registration rate on the third measuring line on the specific charge.

Figures 4 and 5 show that the frequency correction with a reduced voltage of the electrostatic accelerator ensures the normal electrodynamic accelerator operation. Most of the particles pass through the internal path of the dynamic part.

4. Results of simulation

In order to verify the obtained experimental data, the authors modified the software described in [8], which includes:

– a class describing the state of each particle (its velocity components, coordinates, time of flight, mass-to-charge ratio values calculated for 80 kV and 100 kV acceleration modes);

– an array containing a two-dimensional distribution of the field in the drift tubes, calculated on the assumption that the problem is axisymmetric. The grid has a step of 9.775 10-5 m, which corresponds to a division of 10 cm of the path into 1023 intervals or 1024 nodes;

– an iterative method for calculating the trajectory of a particle, based on the Runge-Kutta algorithm. Synchronization check of the particle time of flight with the required switching times on the

drift tubes is added, in case of inconsistency in the calculation of the current step, all field values are inverted, which results in the deceleration of the unsynchronized particles;

– the method of calculating the array with the required switching times of the voltages on the drift tubes, based on the assumption that the moment the particle is in the middle of the drift tube is the moment of commutation, taking into account the difference between the real and the measured ratio of mass to charge;

– a method for generating model packets with the Maxwellian velocity distribution corresponding to the distribution of the flux obtained in the particle injector. This method is implemented using the Box-Muller algorithm with subsequent summation of the velocity vector components and normalization for the most probable energy;

– methods for writing and reading files with particle parameters, trajectory points, and generation of header files for parallel implementation of the algorithm.

The parallel implementation of the algorithm differs singlethreaded one by storing the original matrices with particle parameters and field grid values in the form of external header files, written in plain C. The algorithm implements only the modules required directly for calculating trajectories. Multithreading was provided by connecting the MPI library. The calculation was carried out for mutually independent particles, so each involved node was used to calculate only a certain part of the trajectories.

At the initial stage, we simulated the trajectories of the real particles with the known parameters of input and output velocities and the mass to charge ratio. For each particle, the switching times were calculated, all the particles started from the center of the first drift tube with the given velocities. The particle trajectories were calculated by an iterative method, the field interpolation was carried out on the assumption that the particles have only a positive coordinate along the radial axis. For this reason, the operation of taking the module from the radial coordinate of the particle is added to the interpolator function. Interpolation was carried out for a field section of 1 cm x 10 cm, respectively, the x coordinate within the interpolator function should always lie in the range 0 ÷ 0.1 m. The field interpolation operation for particles having a negative radial coordinate returned an inverted radial field component, flag of negative radial coordinate was used. The synchronous motion of the particle in the path was checked by comparing the current time of flight of the particle with the switching times from the array for the current drift tube number. The inversion of the field components in this case occurred if the time of flight was less than the smaller value or larger than the larger neighboring switching time. The intermediate points of the trajectories were stored during the transition of the centers of the drift tubes. The calculation was terminated either by a surpassing the inner radius of the tube, equal to 1 cm, or by the transit of the center of the 40th tube. The results of the simulation are shown in Figure 6. The experimental results are in good agreement with the calculated data, the error of calculation does not exceed 3.5%.

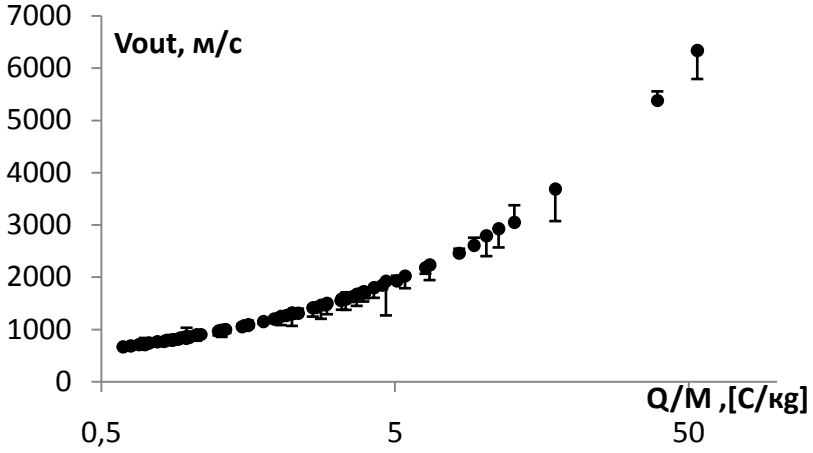


Figure 6. Comparison of experimental and model data on the velocity of the particle at the exit from the dynamic path, the absolute error bars have a scale of 10 to 1.

To test the performance of the acceleration system for other combinations of particle parameters, a set of 16384 particles with different initial characteristics and different specific charges was formed. Simulation of particle motion in the tract of the dynamic part of the accelerator was carried out using the implementation of the software for a personal computer and for a supercomputer Sergey Korolev. The simulation results are shown in Figures 7 and 8.

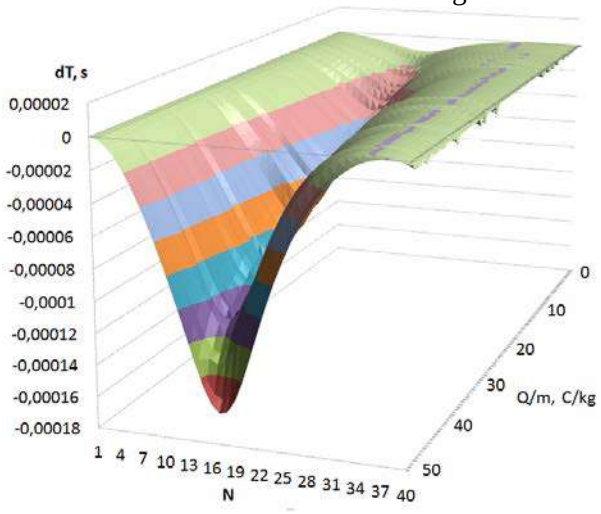


Figure 7. Difference between the calculated voltage switching times on the drift tubes of the electrodynamic accelerator, embedded in the FLASH memory of the control system and the time of particles occurrence in the centers of the drift tubes. Results are obtained by simulating the particle trajectories on the supercomputer Sergey Korolev**.**

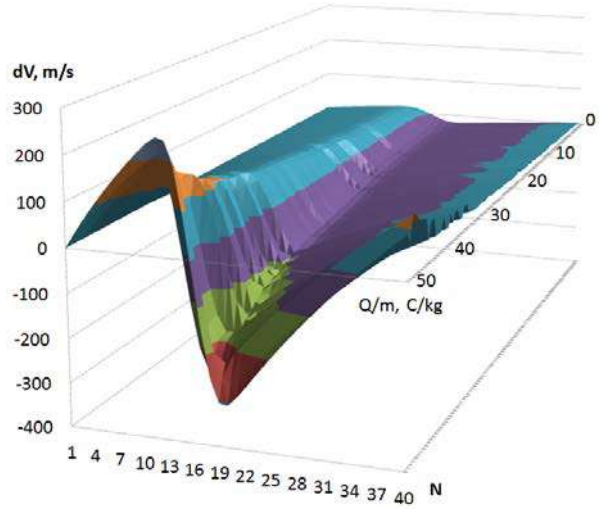


Figure 8. The difference between the calculated speeds of particles during the voltage switching on the drift tubes of the electrodynamic accelerator embedded in the FLASH memory of the control system and the particle velocities during their occurrence in the centers of the drift tubes. Results are obtained by simulating particle trajectories on the supercomputer Sergey Korolev.

Figures 7 and 8 show that the first half of the path particles are slightly ahead of the calculated switching times and, as a consequence, at the initial stage they accelerated more than the calculated data suggests, but in the second half of the tract they are slowed down, and the difference between the calculated and real data is leveled.

5. Conclusion

Based on the results of simulation the of particles trajectories in the path of a linear electrodynamic accelerator, it can be concluded that the control system with adjustable pulse frequency of the master DDS generator from 10MHz to 12.403MHz ensures the operability of the accelerator when the real accelerating voltage of the electrostatic section deviates from the calculated one by 20%. The obtained simulation results for real particles are in good agreement with experimental values, and the simulation of the system operation for a large range of input values will allow to choose the optimal switching frequency without conducting a full-scale experiment.

The implementation of software for a personal computer and supercomputer showed the same accuracy of calculations, which is explained by the choice of the same parameters of the computational grid and the time step. The version for the personal computer showed a longer

execution time of 56 minutes 16 seconds compared to 5 minutes 46 seconds for the supercomputer, but in terms of one processor performance the first one provides a count of 1.2 million iterations per second, and the second one 740 thousand iterations per second. The loss of performance can be explained by the fact that the version for the personal computer retains all results in RAM, the same version for the supercomputer requires transferring the results of calculations between the nodes.

This problem contains a lot of mutually independent processes, so it is easy to implement the algorithm in parallel programming languages. Performance can be improved by increasing the number of computational nodes. Independence of calculating threads provides a proportional increase in performance with an increase in the number of nodes. The application of caching and preliminary processing of results on computing nodes can minimize the amount of data transferred and also increase the speed of execution.

6. References

[1] Semkin N D, Voronov K E and Novikov L S 2005 *Registration of dust and gas particles in laboratory and space conditions* (Samara: Samara State Aerospace University) p 470
[2] Raikunov G G 2014 *Space debris. In 2 books. Book 1. Methods of observation and models of space debris* (Moscow: FIZMATLIT) p 248
[3] Semkin N D, Kalaev M P, Telegin A M, Pijakov A V and Rodin D V 2012 Multilayer film structures under the influence of micrometeoroids and space debris *Applied Physics* **2** 104-115
[4] Slattery J C, Becker D G, Hamermesh B and Roy N L 1973 A linear accelerator for simulated micrometeors *Review of Scientific Instruments* **44** 755-762
[5] Thomas E, Simolka J, DeLuca M, Horányi M, Janches D, Marshall R A, Munsat T, Plane J M C, and Sternovsky Z 2017 *Review of Scientific Instruments* 1-12
[6] Piyakov A V, Rodin D V, Rodina M A and Telegin A M 2017 Numerical simulation of motion of dust particles in an accelerator path *CEUR Workshop Proceedings* **1902** 55-61
[7] Semkin, N D and Piyakov A V 2015 Measurements of particle distributions over the cross section of the accelerator channel for simulating micrometeorites *Instruments and Experimental Techniques* **58(5)** 703-707
[8] Semkin, N D, Voronov K E, Piyakov A V and Piyakov I V 2009 Simulation of micrometeorites using an electrodynamical accelerator *Instruments and Experimental Techniques* **52(4)** 595-601
[9] Semkin N D, Piyakov A V, Voronov K E, Bogoyavlenskii N L and Goryunov D V 2007 A linear accelerator for simulating micrometeorites *Instruments and Experimental Techniques* **50(2)** 275-281
[10] Telegin A M and Piyakov A V 2017 A study of the performance of an induction sensor for an accelerator of charged microparticles *Instruments and Experimental Techniques* **60(6)** 875-879

A research of classification algorithm of spatial information on the basis of methods of persistent homology and random forest

S V Eremeev[1], K V Kuptsov[1] and Yu A Kovalev [1]

[1]Vladimir State University named after Alexander and Nikolay Stoletovs, Gorky street 87, Vladimir, Russia, 600000

Abstract. The classification problem of spatial data is one of the most difficult challenges in the field of the analysis and processing of spatial information. A new approach to a solution of the classification problem of spatial data is presented in article. The offered classification technology of objects will be based on algebraic topology, namely on methods of persistent homology. A barcode is a qualifier of a spatial object. It is determined by computation of topological features of a classified object. The distinctive feature of the offered algorithm is its invariance to affine and topological transformations. The research on results of classification algorithm operation on a set of spatial objects of different classes is carried out.

1. Introduction

Automatic digitization of maps is one of global problems in geographic information systems [1, 2, 3]. Questions of identification [4] and classification of cartographical information appear within this problem. The problem of classification of spatial data on object classes is one of the most difficult in the field of the analysis and processing of spatial information. Russian and world researchers try to solve this problem and propose a set of application-oriented solutions. Having studied scientific works on this subject it is possible to tell that they solve a problem of object classification with various degree of efficiency. There are different methods of classification of spatial objects.

The method intended for work with topographic maps of average scale is presented in [5].The main application is a classification of the area of objects under construction. The method is based on geometrical structures of data and spatial analytical methods. Advantage is improvement of quality of automation of cards with areas of objects under construction.

The problem of classification of spatial data is also relevant for control of information on exhaustion of reservoirs or, on the contrary, – about their degradation [6]. The technology is applied to spatial objects which have similar spectral features, but various form. The algorithm is realized for classification of reservoirs on Alaska and also is used in Bolivia for classification of pastures.

The analysis of the image is applied in [7] together with network methods of extraction of information within a problem of creation of digital tourist maps. The algorithm classifies spatial objects according to the developed rules of simplification and generalization of maps to emphasize reference points and to reduce a role of less significant objects. The technology is applied to creation of tourist maps of San Francisco.

Processing of satellite images or images with high resolution is made in [8] for classification of the objects which are contained in them. A classification is made for the main classes of objects which are presented on topographic maps of large scale.

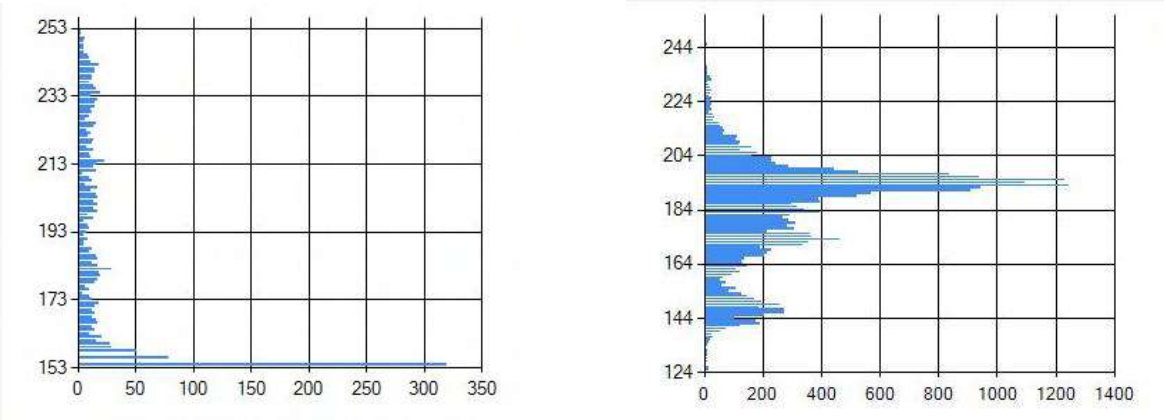
An approach using the example of digitizing distribution maps taken from plant-taxonomic atlases is described in [9]. In result, plant distributions over Europe and Asia have been digitized. The algorithm is a tool to capture data from maps based on obscure projections.

The purpose of work is creation of an algorithm for classification of cartographical information which will make high-quality object classification of various spatial classes and also is invariant [10] to affine transformations and changes of scale.

2. A classification algorithm of spatial data on the basis of methods of a persistent homology and random forest

The offered algorithm of classification of objects is based on algebraic topology, namely by methods of persistent homology. Application of topological characteristics and their analysis is new area of theoretical researches for tasks of the analysis and processing of spatial information. Information from aircraft is processed and analyzed. The allocated objects are distributed on spatial classes in accordance with the classification of spatial information.

A barcode is taken as the qualifier of a class of a spatial object. It is formed by calculation of topological features of the classified object. Set of values of color intensity of all object points is created. Sorting of this set of values according to increase is made. Search of vertices of some intensity is run step by step. It is noted in the list of vertices when finding such point. If this vertice appeared in the neighborhood of Moore of already noted point, then they are connected a line. The triangle is formed at emergence of three such vertices. The number of components (vertices, lines and triangles) at such approach can change on each step of an algorithm. Emergence of vertice adds a component. Emergence of the line connecting different components leads to disappearance of component (two components unite in one). Pass in reversed sequence (on decrease) is the following stage of an algorithm. At the same time the number of holes and their existence time is counted. The hole is formed at emergence of a triangle. The filtration list for holes turns out depending on emergence of new components, their association and other operations. Search of the maximum number of holes and lines is made by the following step. Barcode of the image of an object is calculated on the basis of these numbers (fig. 1).

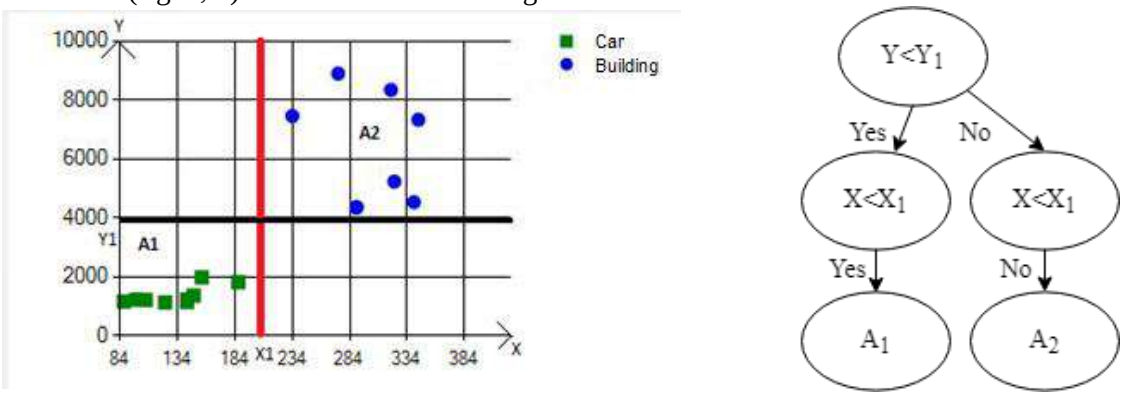


(a) (b)

Figure 1. Barcodes: (a) – the car and (b) – the P-shaped building. The quantity of holes and intensity of points on the color model RGB are displayed on axis X and on axis Y.

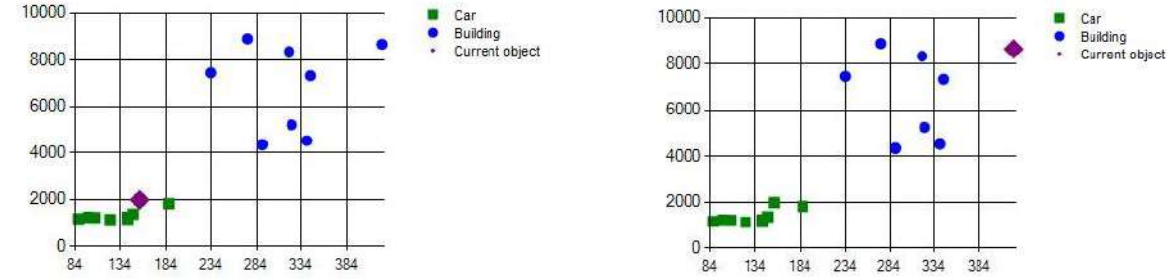
A belonging of object to a spatial class is defined by comparing of barcodes of two objects. Previously training on images of objects of different classes is made (fig. 2, a). Comparing represents check of inclusion of Bettie numbers (the maximum numbers of holes and lines of the image of an object) in the range which characterizes objects of a spatial class (fig.3). The algorithm is complemented with the random forest method [11-13] for optimization of work of an algorithm on

time. It allows to improve speed of an algorithm. The decision tree is formed on the basis of this distribution (fig. 2, b).It is result of work of algorithm.



(a) (b)

Figure 2. (a) Distribution of objects on classes (an algorithm training): A_1 –cars (if $Y < Y_1$and $X<X_1$); A_2 –buildings (if $Y \geq Y_1$and $X \geq X_1$); X – maximum quantity of components; Y – maximum quantity of holes. (b) Decision tree on the basis of this distribution.



(a) (b)

Figure 3. Distribution of objects on classes: (a) – current object is a car and (b) – current object is a building. The maximum quantity of edges and the maximum quantity of holes are displayed on axis X and on axis Y.

The random forest model is applied on the basis of values of Bettie numbers which are taken as features. Random forest is the algorithm consisting of a set of decisive trees. It has been offered by Leo Breiman and Adele Cutler [14]. The algorithm is implemented according to the following scheme.

1. Receiving subselection of the training model. Taking of selection of the training row happens at this stage. The tree is building on its basis.

Basic algorithms have to be unique. Each tree is formed on the training selection for this purpose. There is an element of randomness at the choice of splittings.

The more trees, the quality is better. But time of control and work of Random forest increase in proportion.

2. Viewing of set of random features. It is made for realization of splitting of each branch of a tree.

If the number of features for splitting increases, then time of creation of the forest increases, and trees become "more monotonous".

3. Choice of the best feature and of branch of a tree. Creation of trees is made until disappearance of elements from selection.

Creation of Random forest in direct ratio to selection depth. The selection is deeper, the algorithm is longer executed. The quality on training sharply increases at increase in selection depth. But it usually increases on control selection.

It is recommended to use the maximum depth of trees at realization of this algorithm. Change of the parameters connected with restriction of number of objects doesn't result in significant effect when using superficial trees. Each feature has different degree of importance therefore a part of features can not bear advantages.

3. A research of work of classification algorithm on real spatial objects

The research of the offered algorithm is carried out. Testing of spatial information on the basis of images from aircraft from some height is executed. Classes of spatial objects such as vehicles and buildings are considered. The detailing is made on each class. Vehicles such ascars (are presented on fig.4,a-f), light commercial vehicles and minibuses(fig.4,g-j), buses (fig.4,k-l) are considered. Rectangular, G-shaped, P-shaped, private types of buildings are presented (fig.5).The quantity of the training and test selections in the sum is equal 100 images of objects.

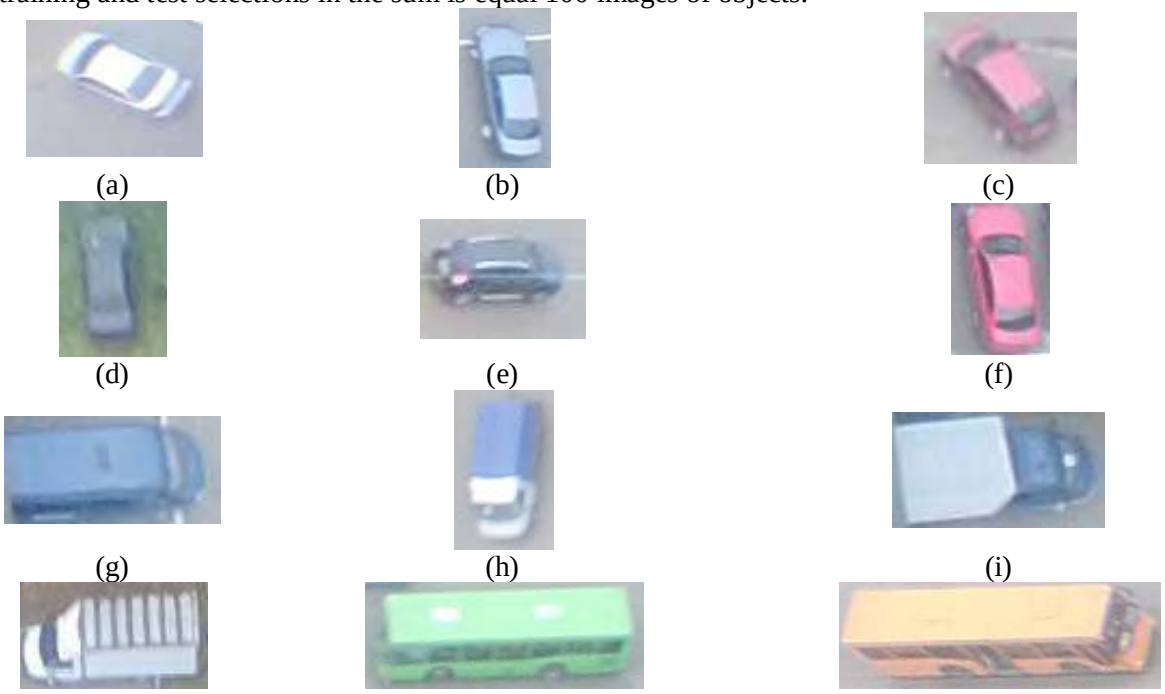


Figure 4. Spatial class of objects "Vehicles". (a-f) – cars; (g-j) – light commercial vehicles and minibuses; (k-l) – buses.

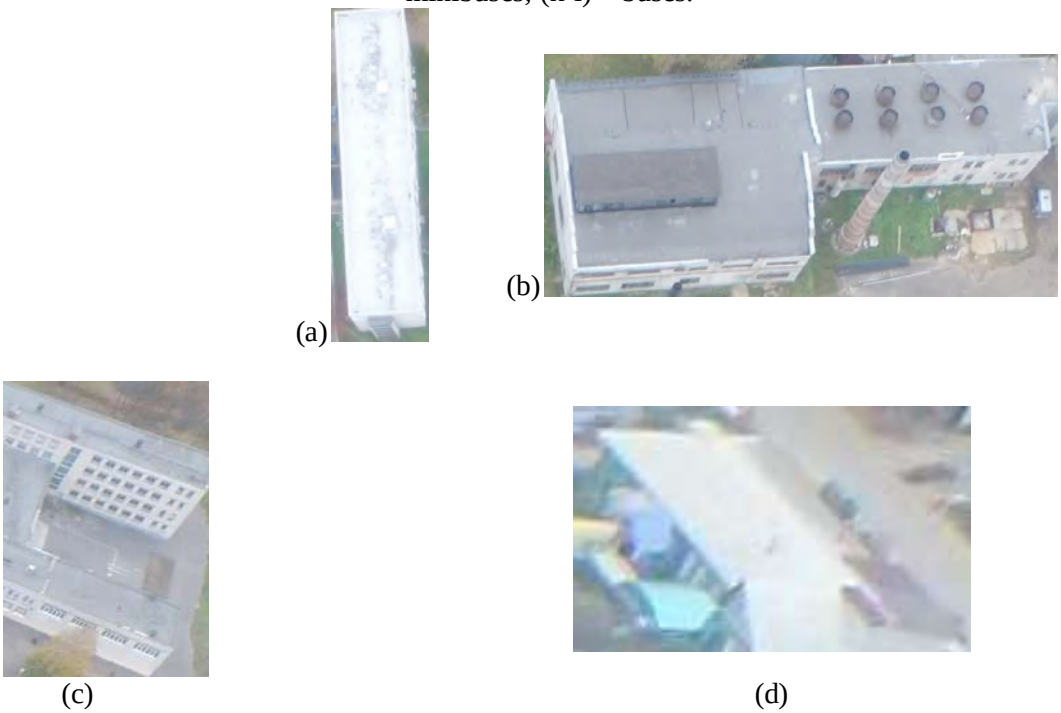


Figure 5. Spatial class of objects of "Building". (a) – rectangular; (b) – G-shaped; (c) – P-shaped; (d) – private houses.

Research results are presented in table 1. The algorithm showed the largest accuracy when determining light commercial vehicles and buses in category of vehicles and for rectangular, G-shaped and P-shaped buildings (100%).

Table 1. Classification of objects by spatial classes.

Object class	Classification accuracy, (%)
Vehicles	98.33
Cars	95.00
Light commercial vehicles and minibuses	100.00
Buses	100.00
Building	97.50
Rectangular;	100.00
G-shaped	100.00
P-shaped	100.00
Private	90.00

Algorithm is invariant to affine transformations. Therefore tests with the different angles of shooting are carried out for the offered types of an object (90˚,180˚,270˚, 360˚– fig.6). Research results of invariancy of an algorithm to affine transformations are presented in table 2.

(a) (b)

(c) (d)

Figure 6. A spatial object of the class "Vehicle" in turn on (a) – 90˚; (b) – 180˚; (c) –270˚; (d) – 360˚.

Table 2. Classification of objects by spatial classes with turn of the image.

Object class	Classification accuracy for turn on 90˚, (%)	Classification accuracy for turn on 180˚, (%)	Classification accuracy for turn on 270˚, (%)	Classification accuracy for turn on 360˚, (%)
Vehicles	100.00	100.00	100.00	100.00
Cars	100.00	100.00	100.00	100.00
Light commercial vehicles and minibuses	100.00	100.00	100.00	100.00
Buses	100.00	100.00	100.00	100.00
Building	100.00	100.00	100.00	100.00
Rectangular;	100.00	100.00	100.00	100.00
G-shaped	100.00	100.00	100.00	100.00
P-shaped	100.00	100.00	100.00	100.00
Private	100.00	100.00	100.00	100.00

Also an algorithm is invariant to scale. Therefore tests in images of the same area of different scales are carried out (table 3). The algorithm hasn't made mistake for two subclasses of vehicles (from three) and for two subclasses of buildings (from four).

Table 3. Classification of objects by spatial classes with images of an object at different scales.

Object class	Classification accuracy, (%)
Vehicles	96.67
Cars	90.00
Light commercial vehicles and minibuses	100.00
Buses	100.00
Building	96.25
Rectangular;	100.00
G-shaped	95.00
P-shaped	100.00
Private	90.00

4. Conclusion

Existing approaches of classification of spatial data are considered in article. New approach of realization of an classification algorithm on the basis of topological characteristics of the image is offered. At the heart of an algorithm methods of a persistent homology and the Random Forest method are put. Results of researches have shown expediency of application of the developed classification algorithm of spatial information as it invariant to affine transformations and changes of scale.

A deformation (distortion) of source data is one of problems of map object classification. A stretching of images of spatial objects is one of types of deformation.The solution of this problem is important within the solution of a problem of classification of spatial information and automatic digitization of maps.Algorithm modernization in the sphere of processing of spatial data with deformations of various types is the following step in improvement of work of the offered algorithm.

5. References

[1] Andrianov D, Eremeev S and Kuptsov K 2017 Identification of spatial objects on digital maps *CEUR Workshop Proceedings* **1940** 1-7
[2] Shekhar S and Xiong H 2008 *Digitization of Maps Encyclopedia of GIS* (Springer, Boston, MA)
[3] Fursov V A, Goshin Ye V and Kotov A P 2016 The hybrid CPU/GPU implementation of the computational procedure for digital terrain models generation from satellite images *Computer Optics* **40(5)** 721-728 DOI: 10.18287/2412-6179-2016-40-5-721-728
[4] Vizilter Y V, Gorbatsevich V S, Vorotnikov A V and Kostromov N A 2017 Real-time face identification VIA CNN and boosted hashing forest *Computer Optics* **41(2)** 254-265
[5] Basaraner M and Selcuk M 2008 A structure recognition technique in contextual generalisation of buildings and built-up areas *Cartographic Journal* **45(4)** 274-285
[6] Frohn R C 2006 The use of landscape pattern metrics in remote sensing image classification *International Journal of Remote Sensing* **27(10)** 2025-2032
[7] Grabler F, Agrawala M, Sumner R W and Pauly M 2008 Automatic generation of tourist maps *ACM Transactions on Graphics* **27(3)** 100
[8] Guienko G and Doytsher Y 2003 Geographic information system data for supporting feature extraction from high-resolution aerial and satellite images *Journal of Surveying Engineering* **129(4)** 158-164
[9] Scholzel C A, Hense A, Hubl P, Kuhl N and Litt T 2002 Digitization and geo-referencing of botanical distribution maps *Journal of Biogeography* **29(7)** 851-856

[10] Fedotov N G, Syemov A A and Moiseev A V 2016 Analysis of conditions that influence the properties of the consructed 3D-image features *Computer Optics* **40(6)** 887-894 DOI: 10.18287/2412-6179-2016-40-6-887-894
[11] Abdulsalam H, Skillicorn D B and Martin P 2011 Classification Using Streaming Random Forests *IEEE Transactions on Knowledge and Data Engineering* **23** 22-36
[12] Biau G, Devroye L and Lugosi G 2008 Consistency of random forests and other averaging classifiers *Journal of Machine Learning Research* **9** 2015-2033
[13] Hastie T, Tibshirani R and Friedman J 2009 *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer-Verlag) p 746
[14] Breiman L 2001 Random Forests *Machine Learning* **45(1)** 5-32

Acknowledgment
The reported study was funded by RFBR and Vladimir region according to the research project № 17-47-330387. The reported study was funded by Vladimir region according to the research project № 326 of 29.09.2017.

Analysis of the personal information from social networks to solve the problems of criminology

E A Gambarova[1], V A Bakaev[1], N V Olinder[1], A V Blagov[1] and M E Naumov[1]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086

Abstract. The article discusses the need to use social networks in the cognitive activities of participants in the criminal process, and suggests that it is possible to use information obtained from social networks in the investigation of crimes. Two approaches are compared: expert and automated. The authors offer tools for data collection and analyzing personal data from social networks.

1. Introduction

An important condition for improving the effectiveness of combating modern crime is the continuous improvement of theoretical and practical knowledge of the investigator, investigator and operative worker on the use of modern technologies in the investigation of crimes, as well as the search for new ways to collect information.

For example, recently much attention has been paid to technical systems for face recognition [1,2,3].

The information is central to the cognitive activity of the investigator, so the search for ways to obtain the information more quickly and fully is an important area in criminology. Great opportunities for working with information are provided by the Internet, in particular, social media (social networks.

At present, it is necessary to note the change in the approach of citizens to the methods and forms of communication, information circulation, etc., partly due to the development of the global Internet, the development of virtual relations. Communication via social networks and messengers is gaining more and more momentum [4]. The growing popularity of social media (social networking, instant messengers, etc.) [5], the spread of "virtual" databases, online banking, cloud storage and other tools used for more comfortable and fast communication and receipt (supply) of services leads not only to the need for normative regulation of these relations, but also determines the creation of new approaches to virtual space in criminology.

The need to use information from social networks in the investigation says the current investigator on particularly important cases of the Republic of Belarus, K. Yu.n. Yu. F. Kamenetsky [6]. In his dissertation "the methodology of the initial phase of the investigation of embezzlement by abuse of official authority in the public sector", he writes that "Anticipation of action by the embezzler and his corrupt ties directly linked to the analytical work to establish social and family ties to the looting, the adoption of measures of preventive character. One of the effective ways of investigating the investigator's activities, along with the classical measures, is the monitoring of social networks available on the global computer network.

This problem is typical not only for forensic and procedural science in the Russian Federation, but also for other countries, as "virtualization" is an integral part of the globalization process taking place

around the world. According to Professor Volchetskaya T. S., the problems associated with virtual space are particularly promising for scientific development. So some steps to forensic and procedural knowledge of the process of virtualization for several years, Russian scientists Efimov V. Yu, Vekhov V. B., Volchetskaya.S., Ishin A. M., Meshcheryakov V. A., Olander N. In. Smushkin A. B. etc. So, Meshcheryakov V. A. one of the first in his work "bases of a technique of investigation of crimes in computer information sphere" referred to "virtual tracks». Subsequently, the section "forensic study of traces" was supplemented with scientific developments related to the study of nature, nature, species, processes of formation, identification and consolidation of virtual traces in the conduct of individual investigations.

When planning certain investigative actions, the investigator is tasked with choosing the most effective ways to achieve this goal. As a rule, one of the tasks of planning investigative actions is to collect information that will help the investigator to choose the tactics of this or that investigative action in the future. Due to the fact that the investigator does not always have a large amount of time to search for information about the event of interest or personality, it is necessary to choose ways that will help to reduce the time of obtaining information. It seems promising to use information from the Internet, including the monitoring of social networks, when planning investigative actions, in particular, interrogation.

2. Methods of social network data collection and processing

In criminology, the organization of the investigation is necessary to quickly obtain reliable information about a person or group of people. Efficiency, reliability and timeliness are key factors. Therefore, it seems reasonable to develop the most effective technology of obtaining and processing information, thus, present the necessary methodological and methodical issues related to the application of the tools of the Internet (social networking, etc.), elaboration of algorithm of search and verification of information. The process of constructing an algorithm for solving a professional problem, the result of which is the allocation of stages of the data processing process, the formal definition of their content and the order of their execution, the development of a template of actions and/or mental operations, allows to optimize the search-cognitive, organizational and technological component of the active interaction of the person conducting the investigation with the objects of the world associated with a criminal event [7].

The task of collecting the necessary information from social networks can be divided into data collection, filtering, processing and subsequent analysis.

The authors of the study set the task of collecting data of social networks expertly, as well as using the developed software. In the first case, a group of experts was determined who systematically searched for the necessary personal information without the use of additional automated services. In the second case, the following approach was used.

Based on the task, the developed software package implements the following functionality:

- analysis of all profiles of target social networks (Vkontakte, Twitter, Instagram, LinkedIn) in order to save open information in the database;
- matching the profiles belonging to one person in the group;
- making assumptions about the user's income level.

For the implementation of a software product, we used the following technology stack: Scala, Python, PostgreSQL, ApacheStorm, CatBoost. This choice is due to the requirement for horizontal scaling of the system.

CatBoost is used to build a mathematical model that determines a person's income level by such parameters as: gender, age, education, field of activity, position, city, family status.

To search for a person's accounts in other social networks, we use a two-layer perseptron, comparing profiles by name, nickname, email, etc.

Primary data are collected as follows [8].

At the first stage the system analyzes all users of social networks VKontakte, Twitter and Instagram and groups them in the following rules:

- in each group there are no more than one profile from each social network;

- all profiles in one group belong to one person.

This problem is solved by means of a program framework of Apache Spark (in particular, superstructures of Spark Streaming intended for stream data processing see fig. 1) and the broker of messages RabbitMQ realizing delivery of basic data in Spark Streaming.

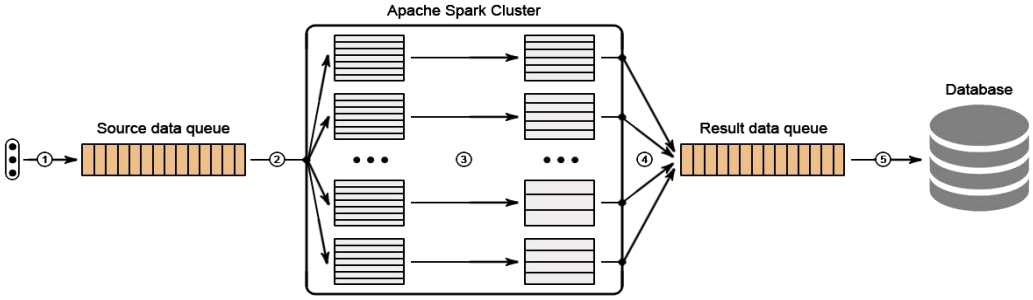


Figure 1. Architecture of the aggregator of users of social networks (pipeline).

Description of steps:

1. Adding of data from different sources in queue for later processing. Data represent a set of couples (network_id, user_id) containing information on profiles which are required to be analyzed.
2. RDD (Resilient Distributed Dataset) formation by a packing of the basic data which are in queue for increase in productivity.
3. RDD (mapping) conversion. For each couple (network_id, user_id) the algorithm finds and groups profiles on other social networks, and also additional information on the person to whom belongs the initial account. The algorithm is restarted for each found profile until all available information on the user is found. As sources can be: the public information specified on the page of the user (the status, contact information, entries in the film, etc.).
4. Export of data retrieveds from RDD in queue for the subsequent saving.
5. Saving results in NoSQL to the MongoDB database in the form of documents with structure, the reflected in table 1.

The speed of data processing makes about 120-130 profiles a second. For work the Microsoft Azure A2 v2 virtual computer was used (2 kernels, 4 GB of RAM, 20 GB of SSD). Casual users of social network VKontakte (1.000.000 profiles) were analyzed.

Thus, if to assume that speeds of processing of profiles of VKontakte, Instagram and Twitter are equal, we will receive an approximate assessment of time which will be required for the analysis of all users of target social networks:

$$T(n) = \frac{4*10^8+6*10^8+13*10^8}{120n} \approx 5324hours \approx 221day, \quad (1)$$

where n - the number of servers in a cluster with a similar configuration.

In case of horizontal scaling of a cluster the linear dependence between the number of servers and processing rate of profiles is watched.

Example of the reference to the table: results of an experiment are reflected in table 1.

Table 1. Data storage structure of profiles.

Name of the field	Type	Description
_id	ObjectId	the document identifier in a collection
vk_id	Int32	the profile identifier in VKontake
facebook_id	Int64	the profile identifier in Facebook network
instagram_id	Int64	the profile identifier in Instagram network
twitter_id	Int64	the profile identifier in Twitter network
other	Object	The additional information (phone number, e-mail address, skype, etc.)

The further task comes down to expansion of the received base by association of profiles by the rules described earlier on which pages bek-links aren't specified other social networks.

Data grouping is based on the analysis of common features [8]. The following procedure applies:

Construction of a full multi-column graph, which stores information about the profiles of social networks and the potential that characterizes the probability of their belonging to one person.

At the vertices of a graph contains information about the profile that is used when the comparison is made.

To compare two profiles, a multilayer neural network is used. The input network layer is fed with a vector of dimension 12, which contains the following data:

Name ↔ Name'
max(Name → Username', Name' → Username)
max(Name → E-mail', Name' → E-mail)
max(Name → Skype', Name' → Skype)
Username ↔ Username'
max(Username → E-mail', Username' → E-mail)
Username ↔ Skype'
max(Skype → Username', Skype' → Username)
max(Skype → E-mail', Skype' → E-mail)
E-mail ↔ E-mail'
Phone ↔ Phone'
Website ↔ Website'

Next, it defines the fullness of occurrences of a in b:

$$a \to b = 1 - \frac{d+r+s}{len(a)} \in [0,1], \quad (2)$$

where d – is the number of delete operations to convert a to b; r – number of replacement operations to convert a to b; s - number of transposition operations to convert a to b; len(x) – function to calculate the length of the argument.

Comparison a and b:

$$\forall i \in [1, len(a)], j \in [1, len(b)] d[i,j] = 1 - \frac{dist(a[i],b[j])}{len(b[j])} \in [0,1], \quad (3)$$

$$a \leftrightarrow b = \frac{\sum_1^{len(a)} d[i, fit(i)]}{min(len(a),len(b))} \in [0,1],$$

where dist(a, b) - is a function that calculates the Damerau-Levenshtein distance [9] for lines a and b; fit(i) - a function that returns the index of the word of the string b, put in accordance with the word a[i].

The comparison operation does not consider the word order. All the words in the source strings are compared in pairs, and then, using the algorithm of Kuhn-Munkres [10], each word of a string a is defined in accordance with the word line b so that the sum of the similarity for all pairs of words was the maximum. Also, punctuation marks and other symbols (except letters and numbers) are not taken into account.

Training and control samples are collected on the basis of primary data. The size of the training sample ~106 pairs.

Next in the generated graph for each pair of shares the following sequence of actions is performed:

- edges are sorted in descending order of weights;

-edges whose weight is less than the threshold value are removed, or one of the incident vertices is already connected to some vertex of the opposite fraction.

The result of these transformations is a graph in which each component of connectivity is a group of accounts from different social networks that belong to one person.

Due to the fact that a person can belong to several communities at the same time, and if the same group was formed in several communities, it can be assumed that the accounts of this group really belong to one user.

Regular expressions and Scala's built-in mechanism of working with CS-grammars are used to parse contact information.

3. Results and discussions

Using the method of expert data collection and processing of social networks, the following results were obtained. The following experiment was conducted. The expert group (115 people), offered to look for information about certain people (3 people) on the given parameters: the place of residence of the person, the place of study of the person, joining the founders, the presence of property, the presence of debts and fines, participation in trials, travel and business trips, leisure, family, close friends.

An important condition was that the participants of the experiment did not use special technical means and looked for information only in open sources.

As a result of the experiment, it was found that it is easy to find information about the city of residence (67% of participants found), although the specific address was found only 5.7% of participants. 79% of the participants were able to find the place of work and study. Joining the founders, shareholders, the status of an individual entrepreneur, etc.found 22% of the participants of the experiment, the availability of information about the property could find only 10% of the participants. Less than 2% of the participants found debts, fines and loans, less than 1% found participation in trials (as a party), more than half of the participants of the experiment could find travel and business trips, 50.6% could find information about parents, 29.5% of the participants of the experiment , about the spouses of more than half of the participants, 52.5%; about brothers and sisters, 19%, about friends, on average 20.9 % (see table 2).

Table 2. TheComparison of two approaches (expert and software).

	The search option	The percentage of participants in the experiment who discovered the information
1	The city of residence	67
2	The residential Address	5,7
3	The place of work and study	79
4	The joining the founders, shareholders, having the status of an individual entrepreneur	22
5	debts, fines, loans	2
6	The fact of participation in court proceedings (as a party)	1
7	Travelling and business trip	50,6
8	The information about parents	29,5
9	The information about spouses	52,5
10	The information about brothers and sisters	19
11	The information about friends	20, 9

The search and processing of information (including the determination of its reliability also took an average of two and a half hours).

Thus, during the experiment, certain parameters ("beacons") were identified, which are freely available in social networks and can be accessed by any user: address, place of residence, place of study, places of rest, business trips, etc.

On the one hand, the experiment showed how much information about the person is stored in social networks, which is a negative factor, as the level of personal data protection is reduced (although these data are placed by the subjects themselves – freely, at will). On the other hand, such "openness" of information can help in the work of investigative bodies in the investigation of crimes. For example, when collecting information about possible participants in organized crime groups, in preparation for individual investigations (interrogation, confrontation) or in General, when planning the investigation of certain types of crimes.

In Russia, the use of social networks in the investigation and prevention of crimes is not widespread. In order to find out the reason of such unpopularity of the use of social networks in the investigation of crimes in the framework of the study, a survey of investigators of the Samara region was conducted. The reasons for this are: the complexity of the search for information (76%), the duration of time (34%), the difficulties of procedural registration of search and use of such information (91 %).

After analyzing the answers of investigators, we concluded that the reduction of time to search for information, as well as the possibility of using some algorithms or software, could create conditions for a wider use of social networks in the investigation of crimes. In this connection, it is advisable to consider the following method of collecting information in the framework of the study.

The second method using the developed software was obtained as follows. The created system of creating portraits of users of social networks is able to collect the following data:

- user profile identifiers in other social networks (including if they are not explicitly specified on their page);
- other contact information (phone numbers, email addresses, Skype logins);
- User's name and nicknames;
- the date of birth;
- city of residence;
- relatives (parents, children, brothers, sisters);
- education (University, school);
- place of work and position;
- уровень дохода (using HeadHunter and Yandex.Work statistic services).

As an experiment, the participants of the community "Big village" were analyzed (https://vk.com/bigvill). For 197 seconds processed data 48.525 profiles Vkontakte (Instagram – 8734, Twitter – 4367, LinkedIn – 1455).

As a result, you can see that with the help of the software you can collect and process much more information more quickly, while, of course, a more detailed analysis can be carried out expertly. The developed software product can be used in criminology for operational preliminary analysis of personal data, including for checking their reliability (for various parameters, for example, according to the specified dates).

Table 3 presents a comparative analysis of the two approaches: with the help of experts and with the help of software.

Table 3. The comparison of two approaches (expert and software).

	Expert	Using the software
The labor expenditures	115experts, 2 hours	The software for 1 PC, 197 seconds
The number of processed data	the social profiles. networks, websites and other open sources	The profiles:VK – 48 525, Instagram – 8 734, Twitter – 4 367, LinkedIn – 1 455.
The completeness of the information	The high, including information about fines and loans	The average: - personal data, - contacts, - connections, - place of work and approximate income level

In General, we can say that social networks can be considered to solve the problems of criminology and serve as the object of research. On the one hand, openness and, as a consequence, availability of data is a negative factor, as the level of personal data protection is reduced (although these data are placed by the subjects themselves). On the other hand, such "openness" can help in the work of investigative bodies in the investigation of crimes. For example, when collecting information about possible participants of criminal groups, when preparing for certain investigative actions (for example, interrogation, confrontation) or in General, when planning the investigation of certain types of crimes.

4. Conclusions

The result of the work is the study and comparison of methods of collection and processing of personal data of users of social networks to solve the problems of criminology. It was found that using social networks you can find a lot of information that users about themselves, their relatives, their work and studies leave on their own, some information is left about users by other users, for example by posting joint photos, videos. These circumstances together make it possible to make a fairly detailed dossier on active users of social networks, which may be important for the investigator, and the information obtained can provide significant assistance in the investigation of crimes.

It is obvious that the comparison of performance and quality of information received by man and machine, gives a predictable result. However, the use of such software by law enforcement agencies can significantly reduce the time to search for basic information and discard from the sample of people who do not meet the specified criteria. And then additional information can be collected by expert.

5. References

[1] Nemirovskiy V B, Stoyanov A K and Goremykina D S 2016 Face recognition based on the proximity measure clustering *Computer Optics* **40(5)** 740-746 DOI: 10.18287/2412-6179-2016-40-5-740-745

[2] Rybintsev A V, Konushin V S and Konushin A S 2015 Consecutive gender and age classification from facial images based on ranked local binary patterns *Computer Optics* **39(5)** 762-770 DOI: 10.18287/0134-2452-2015-39-5-762-769

[3] Nikitin M Yu, Konushin V S and Konushin A S 2017 Neural network model for video-based face recognition with frames quality assessment *Computer Optics* **41(5)** 732-743 DOI: 10.18287/2412-6179-2017-41-5-732-742

[4] Dupuis M, Samreen K and Joyce H 2017 “I Got the Job!”: An exploratory study examining the psychological factors related to status updates on facebook» *Computers in Human Behavior* **73** 132-140

[5] Olinder N V and Gambarova E A 2017 About the results of the experiment "the search and perception of information about a person on the Internet" and its use in the investigation of crimes *Expert-kriminalist* **4** 29-31

[6] Kamenetskiy Yu F 2016 *Methodology of the initial stage of investigation of theft through abuse of official authority in the budgetary sphere: dis. cand. jurid. science* (Minsk) p 196

[7] Stepanenko D A 2016 Algorithmization of the search and cognitive activity of the person conducting the investigation: the grounds, possibilities, problems *Russian Investigator* **10** 3-7

[8] Bakaev V A and Blagov A V 2017 The analysis of profiles on social networks *CEUR Workshop Proceedings* **1903** 88-91

[9] Smetanin N 2011 *Fuzzy search in the text and the dictionary* (Access mode: https://habrahabr.ru/post/114997) (in Russian)

[10] *Hungarian algorithm for solving the assignment problem* (Access mode: http://e-maxx.ru/algo/assignment_hungary) (in Russian)

Acknowledgments

The work has been performed with partial financial support from the Ministry of Education and Sciences of the Russian Federation within the framework of implementation of the Program for Improving the Samara University Competitiveness among the World's Leading Research and Educational Centers for the Period of 2013-2020s.

The approach to the construction of question-answer systems based on the syntagmatic analysis of the text

A A Zarubin[1], A R Koval[1], A A Filippov[2] and V S Moshkin[2]

[1]The Bonch-Bruevich Saint-Petersburg State University of Telecommunication, Moika street 61, Saint-Petersburg, Russia, 191186
[2]Ulyanovsk State Technical University, Severny Venetz street 32, Ulyanovsk, Russia, 432027

Abstract. This article describes the application of the mechanisms of syntagmatic patterns in the construction of various types of question-answer (QA) systems and expert systems. QA systems are systems that can take questions and respond to them in a natural language. In most cases, the principles of building question-answer systems are used in the development of decision support systems. The mechanism of syntagmatic patterns is used when processing open-ended questions and when extracting answers to it from semi-structured resources.

1. Introduction

There is currently a problem of rapid access to information. This problem occurs, for example, when:

- interaction with customers;
- making managerial decisions;
- technical support services, etc

Question-answer systems can be used to solve the problem of operational access to information [1, 2, 3, 4, 5, 6]. Question-answer systems generate a response based on an analysis of the user question. The UIMA architecture is currently used to build question-answer systems [4, 5, 6, 7]. Modern question-answer systems show significant results [4, 5, 6], but they require the presence of highly skilled specialists for tuning and training [7, 8, 9].

This article describes an attempt to develop a question-answer system using syntagmatic patterns [10]. Syntagmatic pattern is a template for detecting certain syntagmatic units in the text. Syntagmatic unit is a collection of several words united on the principle of semantic-grammatical-phonetic compatibility. Relations between syntagmatic units are taken into account when using syntagmatic patterns. For example, the syntagmatic pattern "building * the knowledge base" will allow you to find sentences containing syntagmatic units:

- building a knowledge base;
- building a corporate knowledge base;
- building a fuzzy knowledge base, etc.

Our approach is based on the following ideas:

1. The knowledge base of the question-answer system can be generated automatically based on the analysis of unstructured text resources.
2. The use of syntagmatic patterns to organize the structure of the knowledge base of the question-answer system allows one to effectively search for answers to questions.

2. A structure of the KB of the QA system

The knowledge base (KB) of our question-answer system has a tree-like structure. Semantic networks are currently actively used in the construction of a KB [8]. Formally, the structure of the KB looks like this:

$$KB = \langle SP, TD, R \rangle,$$

where $SP = \{SP_1, SP_2, \ldots, SP_n\}$ is a set of syntagmatic patterns; $TD = \{TD_1, TD_2, \ldots, TD_n\}$ – is a set of text data (KB content); $R = \{R^{SP}, R^{TD}\}$ – is a set of relations of KB:

$R^{SP} = \{R_1^{SP}, R_2^{SP}, \ldots, R_n^{SP}\}$ – is a set of relations between the internal nodes of the KB tree;

$R^{TD} = \{R_1^{TD}, R_2^{TD}, \ldots, R_n^{TD}\}$ – is a set of relations between the internal and terminal nodes of the KB tree.

The internal nodes of the KB tree contain a syntagmatic pattern as a label. Terminal nodes contain text information. The answer to the question will be extracted from this textual information. An example of the KB structure of our question-answer system is shown in Figure 1.

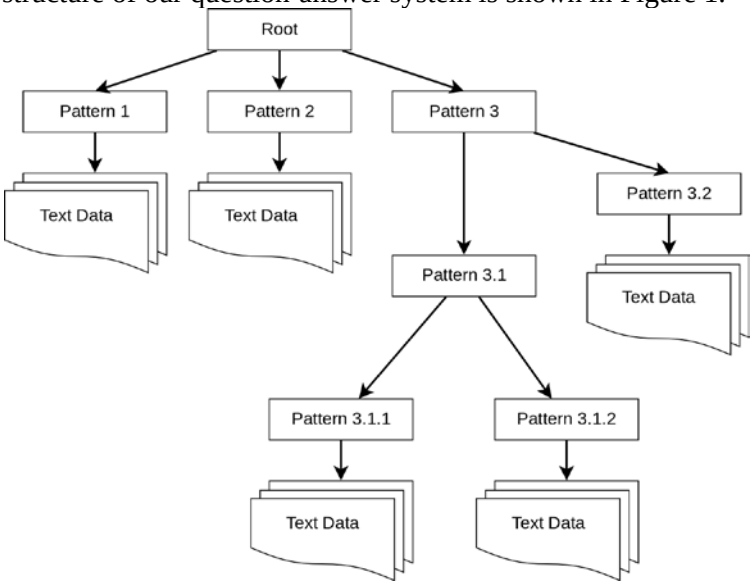


Figure 1. The structure of the knowledge base of our question-answer system.

More general syntagmatic patterns are located closer to the root element of the tree. More precise syntagmatic patterns are located closer to the terminal nodes of the tree. Thus, this knowledge base structure of our question-answer system allows us to find the necessary terminal nodes on the user's question (if the answer to this question is in the knowledge base).

3. Learning of the KB of the QA system

The modified fuzzy C-Means (FCM) fuzzy clustering algorithm is used for learning (building a tree) the KB of our QA system. It is necessary to present each document as an index for the FCM algorithm. Indexing documents consists of the following steps:

1. Download the document.
2. Removing stop words (words that do not have semantic value: prepositions, particles, etc.).
3. Stemming using the Porter algorithm (highlighting the basis of the word).
4. Calculation the frequency of occurrence of words in the document.

The index of the document can be represented as an expression:

$$I_d = \{(w_1^d, \mathrm{f}_\mathrm{i}^\mathrm{d}), (w_2^d, \mathrm{f}_2^\mathrm{d}), \ldots, (w_n^d, \mathrm{f}_\mathrm{n}^\mathrm{d})\},$$

where w_i^d – i-th word of the document d; f_i^d – is the frequency of occurrence of i-th word in the document d; n – is the number of words in the document d.

The modified FCM clustering algorithm is based on minimizing the function:

$$F^{FCM} = \sum_{i=1}^{D}\sum_{j=1}^{C} u_{ij}^{m}\left\|I_i - I_j^c\right\|^2, 1 \le m \le \infty,$$

where D – is the number of document indexes for clustering; C – is a number of clusters; m – is any real number greater than 1; u_{ij} – is the degree to which the document index belongs I_i to the cluster j; I_i – i-th document index; I_j^c – is a center of j-th cluster; $\left\|I_i - I_j^c\right\|$ – the normalized distance between the index of the document and the center of the cluster.

The FCM algorithm consists of the following steps:

1. Initialization of the matrix of indexes belonging to documents to clusters:
$$U = \lfloor u_{ij} \rfloor.$$
2. Calculation of cluster centers:
$$I_j^c = \frac{\sum_{i=1}^{D} u_{ij}^m \cdot I_i}{\sum_{i=1}^{D} u_{ij}^m}.$$
3. Formation of a new membership matrix:
$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\left\|I_i - I_j^c\right\|}{\left\|I_i - I_k^c\right\|}\right)^{\frac{2}{m-1}}}.$$
4. The value of the objective function is calculated. The obtained value is compared with the value at the previous iteration. Clustering is complete if the difference does not exceed the threshold value. Otherwise, go to the second step of the algorithm.

The knowledge base is learned in the process of hierarchical clustering. First, the entire set of document indexes I^0 is clustered. Clusters are formed after the algorithm is executed. Each cluster obtained contains a subset of the documents of the original set: $I^1 \cup I^2 \cup \ldots \cup I^n = I^0$. A new partition is performed for each cluster received. The split continues as long as the value $\sqrt{D/2}$ is greater than or equal to 2.

Thus, a tree is constructed whose internal nodes contain indexes of cluster centers, and terminal nodes contain text data. It is necessary to form internal node labels in the form of syntagmatic patterns based on the contents of internal nodes.

4. An algorithm for constructing syntagmatic patterns

There are currently many approaches to the analysis of texts in natural language [1, 11, 12, 13, 14, 15, 16, 17]. Statistical and/or linguistic methods for the analysis of texts in natural language underlie such approaches. Methods for the analysis of texts on natural language are also used in the development of QA systems [1, 2, 3, 5, 6, 7].

The analysis of texts on natural language consists of the following steps:

1. Grafematic analysis is the selection of structural elements of the text (sentences, names, dates, etc.).
2. Morphological analysis is the definition of the morphological features of the words of the sentence (part of speech, gender, etc.).
3. Parsing is the selection of the syntactic units of the sentence (subject, predicate, etc.).
4. Semantic analysis is the definition of the meaning of the sentence.

It is sufficient to use the first two steps of the process of text analysis in natural language to form syntagmatic patterns: graphematic and morphological analysis.

The algorithm for syntagmatic patterns is represented in Figure 2.

Thus, internal nodes of the knowledge base tree are marked with syntagmatic patterns as a result of this algorithm.

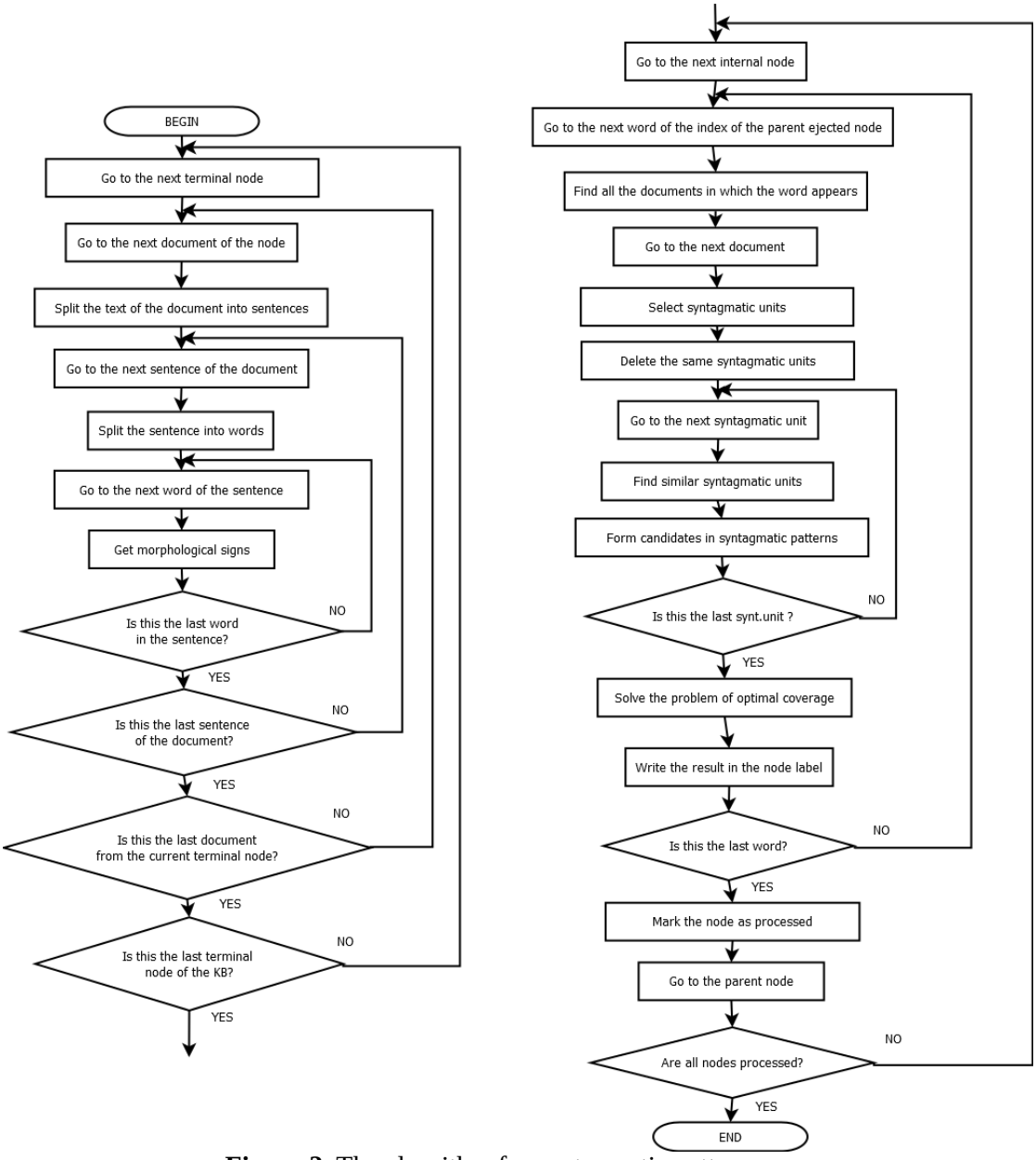


Figure 2. The algorithm for syntagmatic patterns.

5. The search for the answer to the question in the knowledge base

The learned KB of our QA system allows us to find answers to the user's requests. First you need to find the required terminal node of the knowledge base. The internal node labels are used to find the most relevant terminal node. Each internal node of the knowledge base is marked with a syntagmatic pattern.

The search algorithm for the most relevant terminal node of the knowledge base tree is represented in Figure 3.

It is necessary to find in the text documents the most relevant sentence after finding the terminal node. The answer to the question is the most relevant sentence.

The search algorithm for the most relevant sentence from the text documents of the terminal node found is represented in Figure 4.

Thus the two algorithms presented above make it possible to organize the search for the most relevant answer to an incoming question.

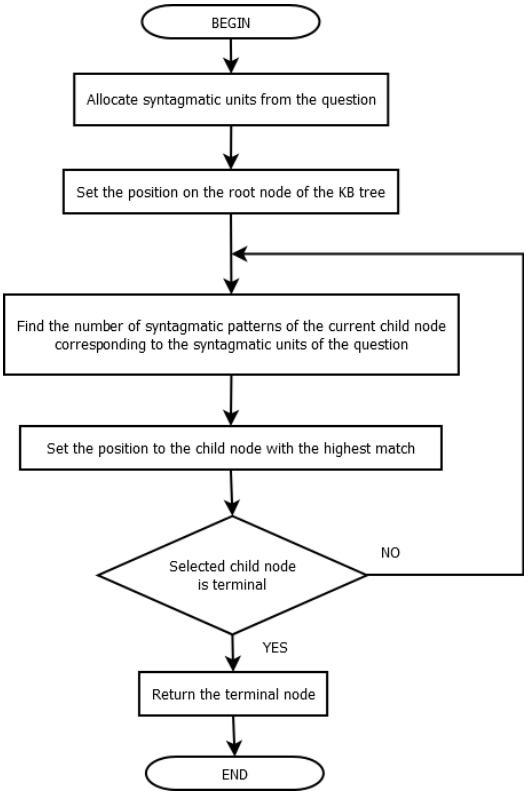


Figure 3. The search algorithm for the most relevant terminal node of the KB tree.

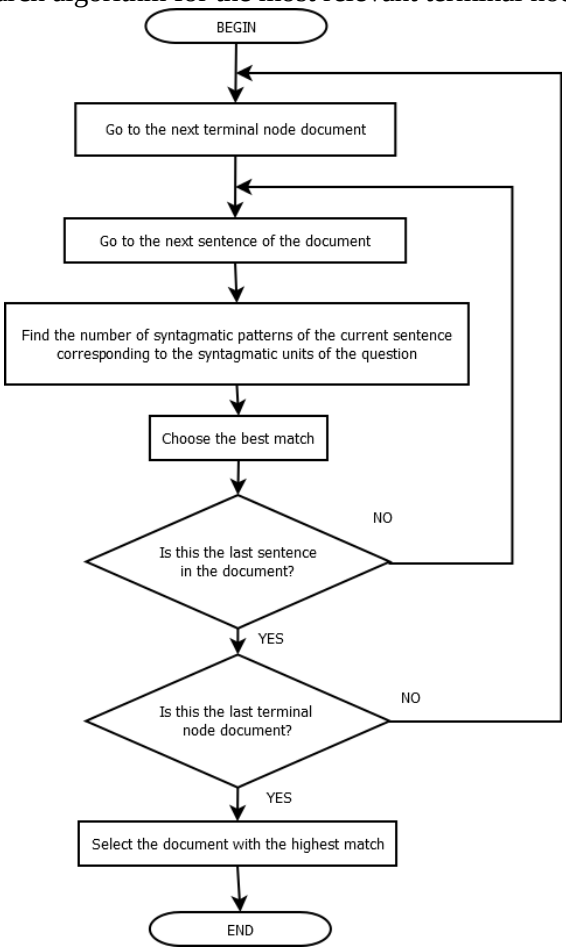


Figure 4. The search algorithm for the most relevant sentence from the text documents.

6. Experiments

The materials of the Sberbank Data Science Contest [18] were used as data for experiments. These materials contain 50,365 entries of the form "paragraph, question, answer." The answer to the question is always the exact text substring of the paragraph, with precision to punctuation and the text register. Each paragraph contains several sentences.

Two question-answer systems were used to conduct the experiment. Their knowledge base was learned using a modified FCM algorithm. The first knowledge base contained many pairs of "term-frequency" as labels of internal nodes. The second knowledge base contained syntagmatic patterns as labels of internal nodes.

A proximity measure was used to find the most relevant document and / or sentence in the first knowledge base. The proximity measure is obtained using the square of the Euclidean distance:

$$Dist(I_i^c, I^q) = \sum_{w=1}^{W} \left(f_w^c - f_w^q \right)^2 ,$$

where I_i^c – is the index of the i-th terminal node document c; I^q – is the index of the received question; W – is the number of words in the index I_i^c; f_w^c, f_w^q – is the frequency of occurrence of the word w in the indexes I_i^c and I^q.

The most relevant document and / or sentence is a document and / or sentence with a minimum proximity measure.

During the experiment 50,365 questions were submitted to both question-answer systems. The sentence from the paragraph was given as an answer to the question. The result was considered successful if the reference answer was a substring of the found sentence.

The results of the experiments are presented in Table 1.

Table 1. The comparison of statistical and syntagmatic approaches to the implementation of the question-answer system.

Type of KB	Number of errors	Percent of errors
Vector	23178	46,0
Syntagmatic	9723	19,3

As can be seen from the results of the experiments, the syntagmatic approach to the implementation of the question-answer system made it possible to reduce the number of errors from 46% to 19.3%.

7. Conclusion

Thus, the developed syntagmatic approach to the development of question-answer systems is effective. This approach can be used to develop the following types of software systems:

- the system for automating the process of interaction with customers based on the analysis of the knowledge base and corporate correspondence;
- decision support system based on the analysis of the knowledge base and use cases;
- the system of verification of information flows of the enterprise to ensure information security;
- the system for automating the work of the technical support service based on the analysis of the knowledge base and use cases.

In the future, we plan to modify the developed approach by finding answers to questions in an implicit form.

8. References

[1] Jurafsky D and Martin J H 2018 *Speech and Language Processing* (Access mode: https://web.stanford.edu/~jurafsky/slp3/28.pdf) (03.05.2018)

[2] Berant J, Chou A, Frostig R and Liang P 2013 Semantic parsing on freebase from question-answer pairs *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)* 1533-1544

[3] Bordes A, Chopra S and Weston J 2018 *Question answering with subgraph embeddings* (Access mode: https://arxiv.org/pdf/1406.3676.pdf) (03.05.2018)
[4] Epstein E A, Schor M I, Iyer B, Lally A, Brown E W and Cwiklik J 2012 Making watson fast *IBM Journal of Research and Development* **56(3.4)** 15-19
[5] Ferrucci D, Brown E, Chu-Carroll J, Fan J, Gondek D, Kalyanpur A A, Lally A, Murdock J W, Nyberg E and Prager J 2010 Building watson: An overview of the deepqa project *AI magazine* **31(3)** 59-79
[6] Gallagher S, Zadrozhny W, Shalaby W and Avadhani A 2018 *Watsonsim: Overview of a question answering engine* (Access mode: https://arxiv.org/pdf/1412.0879.pdf) (03.05.2018)
[7] Ferrucci D and Lally A 2004 UIMA: An architectural approach to unstructured information processing in the corporate research environment *Nat. Lang. Eng* **10(3-4)** 327-348
[8] Bollacker K, Evans C, Paritosh P, Sturge T and Taylor J 2008 Freebase: a collaboratively created graph database for structuring human knowledge *Proceedings of the ACM SIGMOD International Conference on Management of Data* 1247-1250
[9] Chen D and Manning C D 2014 A fast and accurate dependency parser using neural networks *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)* 740-750
[10] Zarubin A, Koval A, Filippov A and Moshkin V 2017 Application of syntagmatic patterns to evaluate answers to open-ended questions *Proceedings of the Communications in Computer and Information Science (CITDS)* 150-162
[11] Zarubin A A, Koval A R, Moshkin V S and Filippov A A 2017 Construction of the problem area ontology based on the syntagmatic analysis of external wiki-resources *CEUR Workshop Proceeding* **1903** 128-134
[12] Boyarskiy K K, Kanevskiy Ye A 2015 Semantic and syntactic parser SemSin *Scientific and Technical Herald of Information Technologies, Mechanics and Optics* **5** 869-876
[13] Artemov M A, Vladimirov A N and Seleznev K E 2013 *Review of Russian NLP systems* (Access mode: http://www.vestnik.vsu.ru/pdf/analiz/2013/02/2013-02-31.pdf) (03.05.2018)
[14] *Automatic text processing* (Access mode: http://aot.ru) (03.05.2018)
[15] Lally A, Prager J M, McCord M C, Boguraev B, Patwardhan S, Fan J, Fodor P and Chu-Carroll J 2012 Question analysis: How watson reads a clue *IBM Journal of Research and Development* **56(3.4)** 2-14
[16] Mikhaylov D V, Kozlov A P and Emelyanov G M 2017 An approach based on analysis of n-grams on links of words to extract the knowledge and relevant linguistic means on subject-oriented text sets *Computer Optics* **41(3)** 461-471 DOI: 10.18287/2412-6179-2017-41-3-461-471
[17] Mikhaylov D V, Kozlov A P and Emelyanov G M 2016 Extraction the knowledge and relevant linguistic means with efficiency estimation for formation of subject-oriented text sets *Computer Optics* **40(4)** 572-582 DOI: 10.18287/2412-6179-2016-40-4-572-582
[18] *Sberbank Data Science Contest* (Access mode: https://contest.sdsj.ru/) (03.05.2018)

Acknowledgments
This paper has been approved within the framework of the federal target project “R&D for Priority Areas of the Russian Science-and-Technology Complex Development for 2014-2020”, government contract No 14.607.21.0164 on the subject “The development of architecture, methods and models to build software and hardware complex semantic analysis of semi-structured information resources on the Russian element base” (Application Code 2016-14-579-0009-0687).

Machine learning in the problem of recognition of pitting corrosion on aluminum surfaces

M Enikeev[1], L Enikeeva[1], M Maleeva[2] and I Gubaydullin[1,3]

[1]Ufa State Petroleum Technological University, Kosmonavtov St. 1, Ufa, Russia, 450062
[2]A.N. Frumkin Institute of Physical Chemistry and Electrochemistry Russian Academy of Sciences, Leninsky prospect 31, Moscow, Russia, 199071
[3]Institute of Petrochemistry and Catalysis, Russian Academy of Sciences, Prospect Oktyabrya 141, Ufa, Russia, 450075

Abstract. The work is devoted to the problem of identification and quantitative estimation of pitting corrosion. The purpose of the work is to design an algorithm and create a set of programs for evaluating and predicting corrosion processes development on the aluminum surface. Object of a research is process of corrosion of aluminum with hydrogen depolarization. Results of work can be used for a research of corrosion processes and their mechanisms on the basis of the visual analysis.

1. Introduction

One of the reasons of decrease in durability and loss of metal designs operability is corrosion destruction of metal. Most of all iron, its alloys and also aluminum suffer from corrosion. Metal designs are subject to generally electrochemical corrosion which arises at contact of metals with electrolytes. Diagnostics and forecasting of corrosion allow estimating service conditions of devices, machines and the equipment. Also they allow defining efficiency of methods and means of protection, to make the decision on their improvement.

Among the non-destructive testing methods for corrosion monitoring [1-4] are scanning reflectometry, confocal microscopy, optical microscopy, ultrasonic inspection, laser scanning microscopy, acoustic emission, vibration analysis, etc. Due to the ease of monitoring, visual inspection is still employed Due to the ease of monitoring, visual inspection is still employed with satisfactory results because it informs about the type and extent of the corrosion. Due to the ease of monitoring, visual inspection is still employed because it informs about the type and level of the corrosion. However, visual inspection has subjective criteria.

The most part of the works connected with detection of corrosion effects by methods of computer vision analyzes cracks and pittings. In most of works, the problem of automation detection and defects measurement was solve successfully. Studying of pitting corrosion of stainless steel in FeCl3 solution became one of the initial experience with computer vision for the metals corrosion research [5]. The ratio of the damaged area to the total image area was determined by photos of corrosion process. Corrosion studying as three-dimensional object, that is determination of depth and a damaged form by computer vision methods merits attention. [6]. In [7] other method of metal corrosion identification based on a minimum distance between the recognition objects is presented. Paper [8] considers fractal geometry methods for the analysis of corrosion process. Researchers don't take into account that the reliable application of image processing needs noise reduction. Abdel-Qader [9] has compared various

edge detection algorithms and has revealed that the Haar's method is the best practice for crack detection. However on noisy data this method is likely to be inefficient as well as the methods based on morphological operations [10]. Prasanna [11] applied histogram analysis and support vector machine (SVM) in classification for the crack detection. The detection rate of the classifier has been found to be approximately 76%. In paper [12] applied the principal component analysis (PCA) for search cracks on the bridge surface. In this case the accuracy of the analysis depends on camera position. Lattanzi and Miller [13] developed the clustering method based on k-means clustering algorithm and the Canny edge detector. In [14] developed algorithm that produces automatically and accurately the quantitative and objective information on the corrosion state of the electroless nickel immersion gold surface using factor analysis (metal turn over and pH of the plating bath). Belim and Kutlunin [15] developed the algorithm of boundary extraction based on image clustering. The mathematical model will allow to describe quantitatively processes on interphase boundaries in a wide number of test environments and metals. The model is based on image processing of corroding surface. The computer vision system can serve as the system realizing such model. In [16] developed design of the device for measurement of the depth of the surface defect of the controlled surface, for example. Vasin [17] describe a machine vision system for the inspection of the railway track condition.

The study of corrosion processes using computer vision methods involves four interconnected processes that create the prerequisites for the creation of data-processing system:

1. Accumulation of experimental data (corroded surface images), which require ordering, structuring and classification for correct use when solving the problem of classifier training.
2. Development of information processing methods depending on the type of problem being solved. In this case, it is the use of image processing methods, mathematical statistics and fractal analysis to evaluate and predict corrosion processes for various types of metals (steel, iron and aluminum).
3. Updating the technical means of processing information, increasing the calculations speed (using graphics processors).
4. Implementation of automatic systems of scientific research in the form of software packages and training of the end user (researcher).

In this paper, we study the results obtained with the help of the created information-computational analytical system (ICAS) for estimating and predicting corrosion (“CorOcenka”). The main results based on the use of ICAS are listed.

2. Pitting corrosion identification

The image of an aluminum surface under corrosion potential in a 0.1 M NaCl solution (pH 11) is shown in figure 1. Process is followed by intensive release of hydrogen. Bubbles of the evolving hydrogen are presented on the image. The task of searching objects in the image was reduced to detection of hydrogen bubbles being evolved. The intensity of this evolution had to be determined from a series of pictures. On the volume of the evolved hydrogen the rate of cathodic reaction is estimated. This rate is equal to the rate of anode reaction and dissolution or corrosion rate.

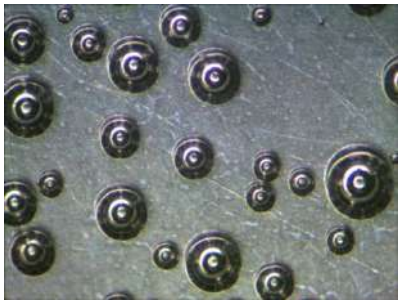


Figure 1. The snapshot of the aluminum surface at a corrosion potential in 0.1 M NaCl (pH 11).

In the study of the corrosion reactions such methods of information processing as image processing, fractal analysis, contour analysis and template matching are used.

Mathematical processing of experimental data consists in work with the image f which is matrix f(M,N):

$$f(x,y)=\begin{bmatrix} f(0,0) & f(0,1) & \text{K} & f(0,N-1) \\ f(1,0) & f(1,1) & \text{K} & f(1,N-1) \\ \text{M} & \text{M} & & \text{M} \\ f(M-1,0) & f(M-1,1) & \text{K} & f(M-1,N-1) \end{bmatrix}$$

This matrix is mathematical model of the image. Each element of the matrix is image pixel.

The following steps can be distinguished in the recognition of surface defects:

1. Image pre-processing (anti-aliasing, noise filtering, increasing contrast);
2. Actual processing (image binarization and isolation of object contours);
3. Image post-processing (filtration of contours on perimeter and area, on form coefficient and fractality, etc).

These operations are specific to each problem of computer visioin. Development and use of the models suitable for the effective solution of a object detection and recognition problems, substantially remains on the verge of science and art. The solution of this problem requires a special know-how or, in other words, knowledge of subject domain reflecting long-term experience of a research on the solution of specific tasks. In more detail this task is considered in work [18].

For search of a bubble on the image we will use the Histogram of Oriented Gradients method (HOG, [19]) The method consists in assessment of normalized gradients histograms of the image brightness. The image is divided into small rectangular regions, in each of which brightness gradient is calculated. For all received gradients in the area the histogram is calculated. The histogram is the descriptor of the image region. For increase in accuracy local histograms use normalization by contrast. In order to take into account changes in illumination and contrast, the strength of the gradient must be locally normalized, which requires grouping the cells into larger, spatially related blocks. Because of its locality, the HOG method is stable to rotations.

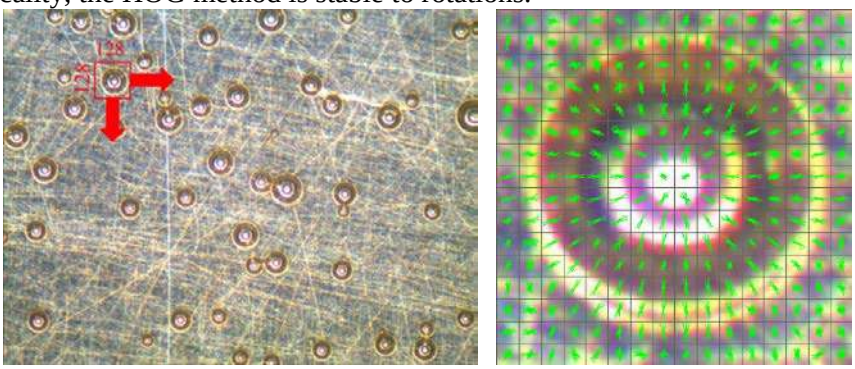


Figure 2. Visualization of the "sliding window" on the example of the image from the data set of the aluminum corrosion with hydrogen depolarization and the construction of oriented gradients histograms for a bubble on the aluminum surface.

We will search for a "bubble" in the image using a "sliding window" with a size of 128 × 128 pixels. We divide it into blocks of 8 × 8. Then one window with a size of 128 × 128 pixels will contain 256 blocks. In each block, we calculate the histogram of oriented gradients with 8 cells. Thus, we get 8 × 256 = 2048 characteristics. In Fig. 2 the visualization of this method is presented, the bubble image has a size of 128 × 128 pixels and is divided into 256 blocks. For illustrative purposes in this case a scaled model is presented. Such a number of characteristics cannot be taken into account without the use of machine learning. An effective method for solving this problem is the support vector machine [20].

Machine learning is the extraction of knowledge from data. This is a scientific field combining the methods of statistics, artificial intelligence and computer science and also known as forecasting analytics or statistical training. The solution of the problem of binary classification using the support

vector method consists in finding a linear function that correctly divides the data set into two classes. Consider the classification problem, where the number of classes is two.

The problem can be formulated as a search for a function y(x) that takes values greater than zero for vectors of the same class (for bubble images) and less than zero for vectors of another class (background). As the initial data for solving the problem, that is, searching for the classifying function *y(x)*, a training set of space vectors is given for which they are known to belong to one of the classes. The collection of classifying functions can be described in terms of the function y(x). The hyperplane is defined by the vector a and the value b, that is, $y(x) = ax + b$. As a result of the SVM-model construction, a function that takes values less than zero for vectors of one class and greater than zero for vectors of another class is found. For each new object, a negative or positive value determines whether the object belongs to one of the classes.

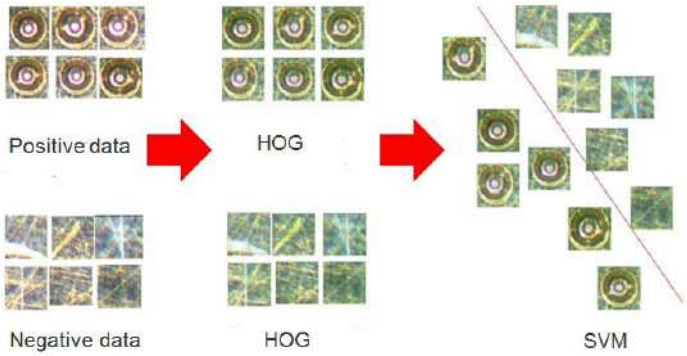


Figure 3. Training of SVM linear classifier.

The sequence of actions for training of the linear SVM qualifier consists of the following steps (Fig. 3):

- We will analyze the image by means of a window of 128*128 pixels in size
- We will break it into blocks of 8*8 in size. Then one window of 128*128 pixels in size will contain 256 blocks.
- We will count the histogram of oriented gradients with 8 cells in each block.
- We receive 8*256 = 2048 properties ıın total.

After the objects in the image are found, in this case a bubble, we calculate its area. Since the contour of the bubble is a circle, then, using the method of least squares for the circle, we can calculate the necessary values for the radius and center of the bubble.

3. Experiment Analysis

The software package that implements ICAS “CorOcenka” was written in C ++ language in the Qt SDK environment using the OpenCV library, and SQLite was used as a database.

To solve the problem of corrosion analysis on the surface of aluminum, a sample was created, consisting of one hundred positive and negative examples.

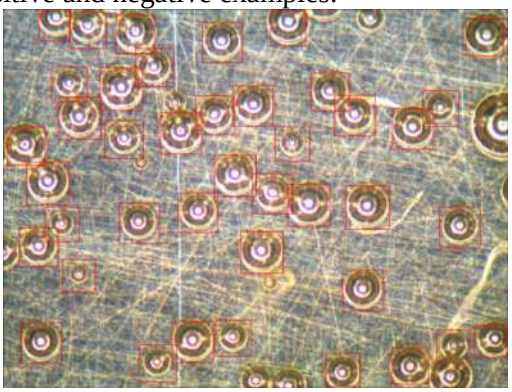


Figure 4. The result of the linear classifier work on the data of the initial experiment.

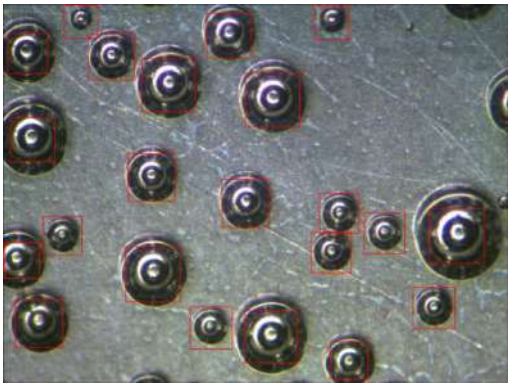


Figure 5. The result of the linear classifier work on the data of an alternative experiment.

On the sample, the SVM linear classifier was trained. Fig. 4 demonstrates that the classifier unsatisfactorily processes bubbles that partially fall into the frame or have small dimensions. This effect can be neglected to obtain an overall picture of the change in the area of the observed bubbles.However, to obtain an overall picture of the change in the area of the observed bubbles, this effect can be neglected.

To verify the performance of the classifier, its work was tested on a set of data from another experiment occurring under different conditions. The result of processing the experimental data is shown in Fig. 5. Despite the fact that the classifier was trained on the data of the first experiment, it showed good performance on an alternative data set.

The use of the parallel implementation of the SVM training had an excellent effect on the calculation time. In general, for the task of identifying bubbles on the image, the acceleration was more than 15 times.

The change in the total area of the bubbles as a function of time was obtained (Fig. 6).

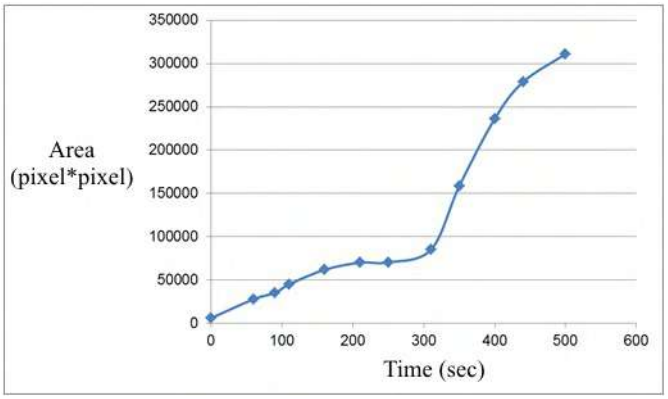


Figure 6. Change in bubble size over time.

Also the possibility of bubbles search by means of contour analysis is realized, but this operation requires the user to introduce the parameters of binarization, noise reduction, morphological analysis and others (Fig. 7). The implementation of this approach is considered in the paper [15].

4. Conclusion

The paper considers the main aspects of image processing and analysis in the studying the corrosion damage mechanism:

- The computer vision and image processing in the problem of detecting bubbles on the aluminum surface are developed;
- The general regularity of corrosion data processing on a aluminum surface is defined;
- The described algorithms are used for detection and recognition of corrosion damage for an aluminum sample by observing the intensity of hydrogen bubbles formation;

• The trained classifier was successfully applied to study several experiments of corrosion on the aluminum surface.

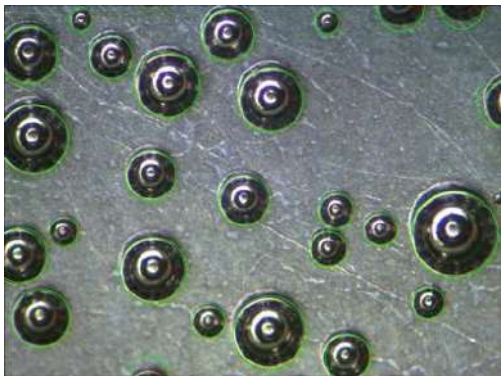


Figure 7. The work of the bubble detector on the aluminum surface.

5. References

[1] Weisner M, Ihlemann J, Muller H, Lankenau E and Huttmann G 2010 Optical Coherence Tomography for Process Control of Laser Micromaching *Rev. Sci. Instrum.* **81** 033705
[2] Kotenev V 2001 Optical scanning of morphology of the protective layers and forecasting of corrosion resistance thermooxidized steel *Protection of Metals* **37(6)** 565-577 (in Russian)
[3] Weber M 2011 Confocal microscopy of the micro and nanostructured surfaces of materials *Nanotech Industry* **28(4)** 42-47 (in Russian)
[4] Kamsu-Foguem B 2012 Knowledge-based support in Non-Destructive Testing for health monitoring of aircraft structures *Advanced Engineering Informatics* **26** 859-869
[5] Itzhar D, Dinstein I and Zilberberg T 1981 Pitting corrosion evaluation by computer image processing *Corrosion Science* **21(1)** 17-22
[6] Quinn M, Bailey M, Ikeda B and Shoesmith D 1993 Image-analysis techniques for investigating localized corrosion processes *Atomic Energy of Canada Limited* p 52
[7] Yan Y, Gao J, Liu Y and Lei S 2002 Recognition and classification of metal fracture surface models based on wavelet transform *Acta Metallurgica Sinica* **38(3)** 309-314
[8] Rudakova O 2012 Fraktal'nyy podkhod k analizu osobennostey ustalostnogo razrusheniya svarnykh shvov *Bulletin of the Perm National Research Polytechnic University. Mechanical Engineering, Materials Science* **14** 102-107 (in Russian)
[9] Abdel-Qader I, Abudayyeh O and Kelly M 2003 Analysis of edge-detection techniques for crack identification in bridges *Journal of Computing in Civil Engineering* **17(4)** 255-263
[10] Yamaguchi T, Nakamura S, Saegusa R and Hashimoto S 2008 Image-Based Crack Detection for Real Concrete Surfaces *IEEJ Trans Elec Electron Eng* **3** 128-135
[11] Prasanna P, Dana K, Gucunski N and Basily B 2012 Computer-vision based crack detection and analysis *Proceedings of SPIE – The International Society for Optical Engineering* **8345** 115
[12] Abdel-Qader I, Pashaie-Rad S, Abudayyeh O and Yehia S 2006 PCA-Based algorithm for unsupervised bridge crack detection *Advances in Engineering Software* **37(12)** 771- 778
[13] Lattanzi D and Miller G 2014 Robust automated concrete damage detection algorithms for field applications *Journal of Computing in Civil Engineering* **28(2)** 253-262
[14] Kantola K and Tenno R 2009 Machine vision in detection of corrosion products on SO2 exposed ENIG surface and an in situ analysis of the corrosion factors *Journal of Materials Processing Technology* **209(5)** 2707-2714
[15] Belim S and Kutlunin P 2015 Boundary extraction in images using clustering algorithm *Computer Optics* **39(1)** 119-124 DOI: 10.18287/0134-2452-2015-39-1-119-124
[16] Abulhanov S, Popov S, Ivliev N and Podlipnov V 2017 Device for Control of Apertures Surface of Pipes of Oil Assortment *Procedia Engineering* **176** 645-652
[17] Vasin N and Diyazitdinov R 2016 A machine vision system for inspection of railway track *Computer Optics* **40(3)** 410-415 DOI: 10.18287/2412-6179-2016-40-3-410-415
[18] Enikeev M, Gubaydullin I and Maleeva M 2017 Analysis of corrosion process development on

metals by means of computer vision *Engineering Journal* **21(4)** 183-192
[19] Dalal N and Triggs B 2005 Histograms of Oriented Gradients for Human Detection *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* **1** 886-893
[20] Vapnik V 1999 An Overview of Statistical Learning Theory *IEEE Transactions on Neural Networks* **10(5)** 988-999

Development and implementation of social network data collection services to improve the human environment

I A Rytsarev[1], A V Blagov[1] and M I Khotilin[1]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086

Abstract. The article discusses the need to develop and implement services for the collection of social networking data for the detection of environmental problems. The authors offer methods and tools for collecting and analyzing data on identified problems from social networks.

1. Introduction

There are many ways of monitoring a person's environment, such as photography from satellites or unmanned aerial vehicles, equipped with cameras, transportation, or installation of stationary cameras, etc. Different methods of approach can be used to detect the necessary objects [1-2]. Often, this problem is solved by the forces of certain municipal services. As a rule, this approach is associated with high labor costs. Such information services as thematic Internet sites, for example, are becoming popular: https://rosyama.ru/, http://moyasamara.com/ etc. It is worth noting that to apply through these sites, you must at least remember the name of the resource, its address. Moreover, creating an application, the author must fill out information about yourself. Among some people, these restrictions can create some difficulties. At the same time, social networks are becoming increasingly popular [3]. At present, social networks are at the peak of popularity: Now millions of people are using Facebook and Twitter. The direction of Big Data, which is related to social media is one of the most perspective areas, and it is developing dynamically. Over the past decade, social networks began to play a huge role: Being the subject of the socialization of people on one hand, and being the most powerful and accessible political, ideological and economical instrument on the other hand. Due to large volumes and continuous regeneration, the research of the data from social networks can be produced by means of methods and instruments of Big Data [3-4]. The term «big data», in information technologies, means a series of approaches, instruments and methods of processing structured and unstructured high volume datasets.

In terms of marketing, social networks have become the most attractive medium for different programs' realization. They are the second place of the quickest means. In Twitter, the social network, the specialists of marketing can communicate with their audience without Service of Public Relations. Due to this, there could be a communication with a specific person, in contrast to depersonalized companies.

Using their branded terms and hash-tags, marketing specialists can learn the customers' opinion about their products, brands and companies. The global attention, which Twitter uses, shows high capabilities of social net-work technologies for public discussion and forming the perception of brands. Besides, the information collected from social networks is important in terms of questions pertaining to national security.

The direction of Big Data related to social media, is one of the most perspective and dynamically developing.

Through gathering and structuring text data from social networks, the attitude of users to any selected issue can be analyzed. Additionally, through analysis, the distribution of data by countries and cities can be received. It helps to estimate the popularity of the selected theme in specific geographical locations.

Many people have smartphones and other personal mobile gadgets, equipped with the ability to access social networks. Due to this, each user generates a large amount of data that may be of interest for different areas and areas of activity. A lot of research works are devoted to the collection, processing and analysis of social network data [4-6]. It is worth noting that for the user of social networks themselves networks can be the most convenient resource for posting information, including on various environmental problems: landfills, accidents, pits, fires, etc.

This article describes the implemented service for the operational collection of information about the problems of the environment in one of the most popular social networks Twitter. The service allows you to collect the necessary information on-line for any necessary geolocation.

2. Data collection algorithm, their processing and classification

There are many social networks (YouTube, Twitter, Facebook, VK, Instagram, etc.). With their help, users exchange various information (text information, video images, etc.). The social network Twitter was chosen for the research. This was done for the following reasons:

- users of this social network mainly share text information which is easier to process;
- the network provides open access to its data (there is no restriction on access to server data streams);
- twitter is the second most popular social network (after Facebook, which does not provide open access to its data) among users worldwide;
- twitteris not a substantive network and most broadly reflects public opinion on issues of interest to us.

The task of collecting the necessary information from social networks on environmental problems can be divided into data collection, filtering, processing and, if necessary, classification. The data generated on-line are of the greatest value for obtaining information on various problems of the habitat. In this case, the filtering can go on a certain, interesting geolocation, and on the subject of the published content.

The following information is generally required when collecting Twitter data:

- idofthemessage;
- messagetext;
- listofmessagehashtags;
- geolocation;
- date and time of the message.

The data collection algorithm developed in the framework of the study, in addition to the indicated filtering, uses an additional element of user feedback through an automatic request for clarification of information about the published message. Figure 1 shows the scheme of the algorithm for collecting the necessary data from the social network.

The social network data generated in on-line mode is collected using the F1 filter for the specified geolocation. Next, using the content filter F2, configured by key words or topics, the necessary data on habitat problems are collected. The data obtained by double filtering can be schematically divided into two parts: the data about the author-A and the message-M. In this case, according to the developed algorithm, the author of the message automatically receives a clarifying request Q, to provide additional information, such as the exact address or coordinates of the problem described in the message M, or its nature. The author a, receiving this request, leaves message N, containing the necessary information. We can say that according to the developed algorithm gathers the following data set: $\sum_{i=1}^{k}(M_{gi} + N_{gi})$, for all k users A_i, the message authors M_{gi}, generated by geolocation g, and send the additional information N_{gi}.

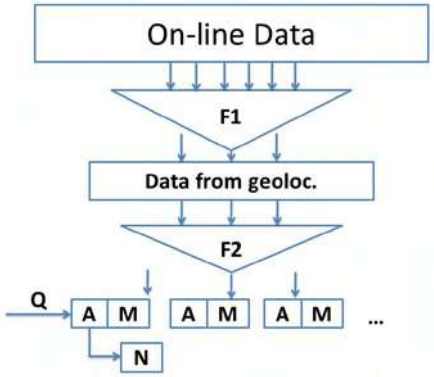


Figure 1. The general scheme of the algorithm of collecting social network data.

After collecting the required data is often required to classify them by reference to certain issues: landfills, pits, accidents, etc. To solve this problem, one can use the key word "hashtag." this message [7], or other approaches, for example N-grams [8]. For a more detailed analysis, the algorithm of collective decision-making developed in the framework of this study can be used.

The method of collective decision - making is based on the uniqueness of each word. It is based on the dictionary, which stores the words and their categories to which they belong.
The dictionary is formed with the participation of an expert who teaches the system, and has the ability to make changes to the dictionary to clarify it.

Any collected post each word from the text message $T = t_1 t_2 t_3 ... t_n$ makes to a dictionary $U = u_1 u_2 u_3 ... u_m$ to obtain a list of categories:

$$\delta(u_i) = \{\underbrace{k_j ... k_h}_{z}\},$$

where $\delta(u_i)$ - is the function of extracting categories of the word u_i from the dictionary; $\underbrace{\{k_j ... k_h\}}_{z}$ - the vector of z elements k (element list categories $K = k_1 k_2 k_3 ... k_l$), to which the word u_i belongs; $1 \le j, h, z \le l$.

If $\gamma(t_i) = u_i$ ($\gamma(t_i)$ - the function of the stemming text, then $\delta(t_i) = \delta(u_j) = \overline{k}$. If there is no word in the dictionary, it is given to the expert for classification with subsequent addition to the dictionary.

A distinctive feature of this algorithm of collective decision-making is the presence of a dynamic threshold value: when determining the threshold value, the number and weight of the categories of words mentioned in the text are taken into account.

As a result of text processing we have a list of categories (which include words) on the basis of which the algorithm calculates the threshold value of the "weight" of the subject and classifies the source text.

3. Results and discussions

The use of sections to divide the text of the paper is optional and left as a decision for the author. Where the author wishes to divide the paper into sections the formatting shown in table 2 should be used.

Data collection on Twitter can be carried out through Apache Ambari and Flume software products, this method is described in more detail in [9]. However, to collect data using a number of filters, it is often more convenient to develop your software product using standard libraries (twitter4j, tweepy, etc.) [10].

Within the framework of this study, a software product was developed in the Python programming language, containing an authorization module, a data collection module and a filtering module. This software allows you to collect data on geolocation, keywords, user, and cache all media files of the

user. To avoid interruptions in the operation of the software product associated with exceeding the limits set by the social network Twitter, a lot of authorization keys are sewn into the software product. The software works in real-time monitoring mode, and can also make requests for information lying on servers.

Launches of the developed software product were made on the basis of Launches of the software complex were made on the hardware and software complex of data processing of the very large volume of the laboratory for data processing of the very large volume of the Samara University. Hard ware and soft ware complex consists of:

- the software and hardware complex of storage and analytical analysis of IBM Puredata for Analytics (Netezza) structured data with a volume of disk space of at least 96Terabyte (taking into account 4-times compression and full data replication);
- Hadoop-clusterраспределённогохраненияианалитическойобработкинеструктурированныхданных (IBM x3630 M4 management server (two Intel Xeon Processor E5-2450v2; 96 GB of memory; 2 600GB disks) and four IBM x3630 M4 data processing servers (two Intel Xeon Processor E5 2450v2 processors; 96 GB of RAM; 8Terrabyte of disk memory)).

Figure 2. The receive notification and clarifying the question on Twitter (example 1).

The hardware and software complex has access to the world wide web at a speed of 10 GB per second, which allows you to work with a large flow of messages.

For each user of the social network Twitter implemented interface for interaction is a bot account @contriol63. The geolocation filter is set to the sixty-third region of the Russian Federation, which is the Samara region.

A user who is going to write in his message about the problem of the environment, can add to your message or the name of the bot account, or its address, or a set of keywords presented in the table №1.

Table 1. Keywords necessary for use in the message.

Problem	Keywords
the unauthorized dump	#dump63, dump63, #trash63, trash63, #svalka63, svalka63, #musor63, musor63,
the fire	#fire63, fire63, #pozhar63, pozhar63,
pitsonroads	#pit63, pit63, #yama63, yama63
breakdown	#breakdown63, breakdown63, #avariya63, avariya63
danger	#danger63, danger63, #opasnost63, opasnost63,
disturb	#disturb63, disturb63, #narushenie63, #narushenie63,
no designation name address of the bot	andcontrol63, @control63, контроль63, @контроль63

The list in table № 1 may be supplemented.

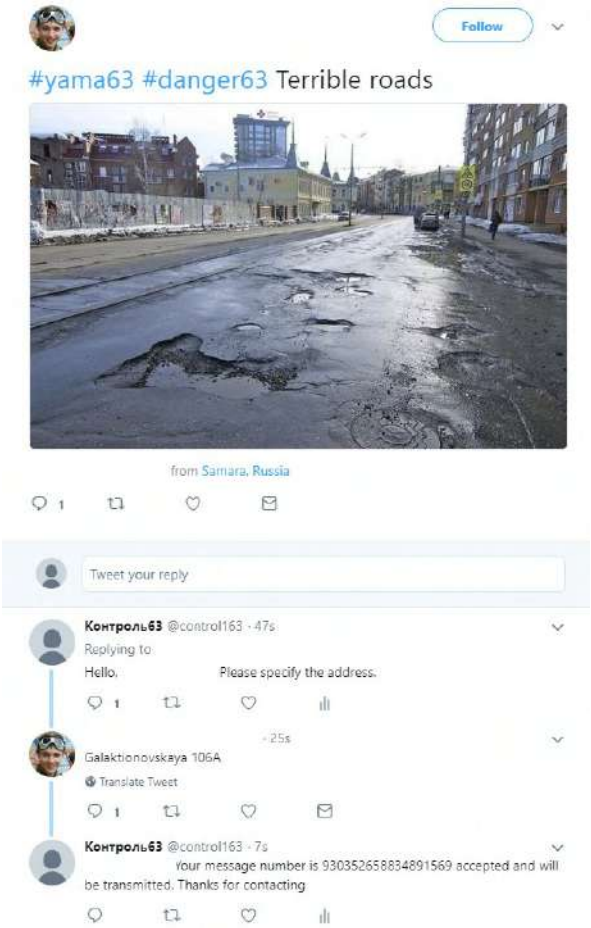


Figure 3. The receive notification and clarifying the question on Twitter (example 2).

A written message containing any of the keywords in table 1, with geolocation by 63 region, is collected and classified by type. The author of the message is automatically notified of the collection and asked a clarifying question (figure 2, figure 3).

The collected pool of reports on various problems of the human environment is valuable information and can serve for the rapid targeted elimination of accidents and violations. At the same time, the service is very convenient for all users who simply continue to use their social networks, putting there information about the problems that have arisen. It is also worth noting that the developed service can be of practical value only if it is supported by structures and services dealing with the problems of the human environment. The main motivation for the users who post information on this topic, can only be a rapid response and subsequent problem solving. At the same time, in case of successful testing and implementation within the Samara region, the project can be scaled to other regions and even countries.

4. Conclusion

The result of the work is the development and implementation of the service to collect the necessary information about accidents and violations in the social network Twitter for a certain geolocation. Additionally, a classification tool for collected messages is also implemented.

In the future, the developed service can be: first, scaled to other social networks, for example, Vkontakte, and synchronized between them, and secondly, it can be supplemented with an image processor (photos attached to the message) and video.

5. References

[1] Epifancev B N, Pyatkov A A and Kopeykin S A 2016 Multi-sensor systems for monitoring access to restricted areas: capabilities of the intrusion detection video analytical channel *Computer Optics* **40(1)** 121-129 DOI: 10.18287/2412-6179-2016-40-1-121-129

[2] Vizilter Y V, Gorbatsevich V S, Vishnyakov B V and Sidyakin S V 2017 Object detection in images using morphlet descriptions *Computer Optics* **41(3)** 406-411 DOI: 10.18287/2412-6179-2017-41-3-406-411

[3] Tan W, Blake M, Saleh I and Dustdar S 2013 Social-network-sourced big data analytics *IEEE Internet Computing* **5** 62-69

[4] Semertzidis K, Pitoura E and Tsaparas P 2013 How people describe themselves on Twitter *Proceedings of the ACM SIGMOD Workshop on Databases and Social Networks* 25-30

[5] Xu X 2007 Scan: a structural clustering algorithm for networks *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining* 824-833

[6] Blagov A, Rytcarev I, Strelkov K and Khotilin M 2015 Big Data Instruments for Social Media Analysis *Proceedings of the 5th International Workshop on Computer Science and Engineering* 179-184

[7] Krokos E , Samet H and Sankaranarayanan J 2014 A look into twitter hashtag discovery and generation *Proceedings of the 7th ACM SIGSPATIAL International Workshop on Location-Based Social Networks* 49-56

[8] Mikhaylov D V, Kozlov A P and Emelyanov G M 2017 An approach based on analysis of n-grams on links of words to extract the knowledge and relevant linguistic means on subject-oriented text sets *Computer Optics* **41(3)** 461-471 DOI: 10.18287/2412-6179-2017-41-3-461-471

[9] Rytsarev I A and Blagov A V 2015 Construction of activity models of users of social networks *Proceedings of the Information Technology and Nanotechnology* (Samara: Samara National Research University) 216-220

[10] Rytsarev I A and Blagov A V 2017 Development and research of algorithms for clustering data of super-large volume *CEUR Workshop Proceedings* **1903** 80-83

Acknowledgments

The work has been performed with partial financial support from the Ministry of Education and Sciences of the Russian Federation within the framework of implementation of the Program for Improving the Samara University Competitiveness among the World's Leading Research and Educational Centers for the Period of 2013-2020s.

Reconstruction of the phonetic composition the recognized word using lexical ontology

V S Moshkin[1], A I Armer[1] and N A Krasheninnikova[2]

[1]Ulyanovsk State Technical University, Severny Venets street 32, Ulyanovsk, Russia, 432027
[2]Ulyanovsk State University, Lev Tolstoy street 42, Ulyanovsk, Russia, 432017

Abstract. It is possible to improve the quality of speech recognition in noisy environments adapting the reconstruction algorithm for the recognized word to the certain peculiarities of usage and application. The article describes an approach to reconstruct the phonetic composition of a recognized word using lexical ontology. The lexical ontology contains lexical links among the words of the subject domain and their phonetic composition in terms of the SAMPA+ for the Russian language.

1. Introduction

Continuous speech recognition is a complex iterative process, based on the sequential division of the total acoustic signal into words, and then the words into phonemes. At the same time, many factors (e.g., noise) significantly complicate the recognition and reduce its accuracy. Thus, after preliminary speech signal processing many algorithms result in a matrix consisting of a phoneme set with their corresponding detection probability in the recognized word.

However, the task of recognizing and converting Russian continuous speech into a text is an urgent problem. It should be solved in different subject domains: medicine, litigation, automated detection of extremist materials, etc.

At the same time, a special set of the most frequently used terms corresponds to each subject domain. The a priori sampling of the basic terms in the analyzed subject domain helps to increase the probability accuracy while determining the patterns of certain phonetic combinations in the process of speech recognition.

One of the ways to adapt speech recognition algorithms to the corresponding subject domains is to apply lexical ontologies of subject domains.

2. The algorithm for the phoneme-by-phoneme recognition of the Russian words

For the initial extraction of speech units from the recognized speech signal, the algorithm of the phoneme-by-phoneme recognition of the Russian words is used. The initially detected speech units will be used to form word groups taking into account both subject domains and the analysis results of a certain lexical ontology.

For the phonemic recognition of the Russian words from an unlimited dictionary, we use the following algorithm: a speech segment is divided into constituents [1, 2]. For this purpose, it is preliminary transformed into a two-dimensional autocorrelation portrait. Then, in the sliding window, which size is equal to the corresponding portrait of a model speech unit, the distance is calculated. During the calculation, the distance between the windows is optimized using the discrete dynamic

programming. For each speech unit, a distance array along the portrait of the analyzed speech segment is determined. The distances corresponding to the same fragments of the analyzed speech segment portrait are compared with each other. As a result, speech unit portraits, which have the smallest distances, form the desired boundaries in such a way that the starting and final readouts of each speech signal are known. In such a method, the average error in determining the beginning of a speech unit is 3579 samples, in the interval [0; 10860] samples, the average error in determining the end of a speech unit is 3724 samples, in the interval [0; 12481] samples; the sampling rate is 44100 S/s.

Then, the speech signal of each speech unit is successively converted into an autocorrelation portrait in the following way. Let $s(i)$ be the *i-th* readout of a digital speech signal; then $s(i+k)$ is a readout spaced k readouts apart $s(i)$. Dependency factor of these readouts is expressed by a sample correlation coefficient:

$$R_s(k) = R[s(i), s(i+k)] = \frac{\mathrm{cov}[s(i), s(i+k)]}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} s^2(i) - m_{s(i)}^2}\sqrt{\frac{1}{N}\sum_{i=1}^{N} s^2(i+k) - m_{s(i+k)}^2}},$$

$$\mathrm{cov}[s(i), s(i+k)] = \frac{1}{N}\sum_{i=1}^{N} s(i)s(i+k) - \left[\frac{1}{N}\sum_{i=1}^{N} s(i)\right]\frac{1}{N}\sum_{i=1}^{N} s(i+k), \quad (1)$$

where N is a number of readouts in the interval, in which the dependency is sought; $\mathrm{cov}[s(i), s(i+k)]$ is the sample covariance $s(i)$ and $s(i+k)$ when $i = 1..N$; $m_{s(i)}$ is a sample mean $s(i)$ when $i = 1..N$; $m_{s(i+k)}$ is a sample mean $s(i+k)$ when $i = 1..N$. Function determined by the sample correlation coefficient using (1) is an autocorrelation function of a signal. While its calculation we perform the transformation of a speech signal $s(i) i = 1..M$ (M is the number of readouts in a speech signal) into a two-dimensional image. For this purpose, $s(i)$ is divided into intervals including $N < M$ readouts. Then, using equation (1) we generate image lines:

$$\begin{matrix} s\left(i_m^j\right), s\left(i_m^j + k\right) & \begin{matrix} k=1..N \\ j=1,N,2N,\ldots \end{matrix} \\ X(j,k) = R \end{matrix}. \quad (2)$$

The two-dimensional image *X (j, k)* obtained from (2), where *i* is the line number, and *k* is the column number, is the autocorrelation portrait (ACP) of a speech signal *s (i)* dimensioned $N \times \frac{M}{N}$.

Model speech units with the same parameter value *N* are transformed into ACPs. These speech units are arranged from the examples of SAMPA + phonetic alphabet. Thus, it is possible to determine to what extent the ACP speech unit corresponds to the model ACP. As a result, the speech unit is considered to correspond to the model with the utmost similarity. The similarity of the ACP is determined by calculating the Euclidean distance between the ACP lines. The position of each line is determined in dynamics [3, 4, 5]. Nevertheless, if in ACPs the number of lines exceeds a certain threshold value they are considered different.

3. A model of lexical ontology

Ontology is a system consisting of many concepts, their definitions and axioms, which are necessary to limit the interpretation and use the concepts [6].

OWL (Ontology Web Language) is the Semantic Web language designed to describe classes and their interrelations. At the heart of the language lies the representation of reality in the "object-property" data model. OWL is a reformulation of the descriptive logic using XML syntax.

Subject domain ontology is a collection of RDF-triples: subject-predicate-object. In this research OWL-ontology was used to solve the problem under consideration [7, 8].

A special type of ontology is lexical (or linguistic). Its distinctive property is to use (lexicalized) concepts (words) together with their linguistic properties in one resource. The main source of concepts in such ontologies are the values of linguistic units. They are also distinguished by a set of relationships, which usually characterize linguistic elements: such as synonymy, hyponymy, meronymy, etc. [9, 10, 11].

To reconstruct the phonetic composition of the recognized word, the elements of phonetic alphabets, establishing letter-sound correspondence, were included in the lexical ontology structure. The most widely-used phonetic alphabets are the International Phonetic Alphabet (IPA) and X-SAMPA (as well as SAMPA + modification, including transcriptions of the Russian language) [12].

The formal model of a certain subject domain lexical ontology may look as follows:

$$O = \left\langle A, C^{A_i}, P^{IPA}, P^{SAMPA+_i}, R^{A_i} \right\rangle, \mathrm{i} = \overline{1, m},$$

where m is the number of subject domains covered by ontology; $A = \{A_1, A_2, K, A_t\}$ is the number of subject domains covered by ontology; $C^{A_i} = \left\{C_1^{A_i}, C_2^{A_i}, K, C_n^{A_i}\right\}$ is a set of terms within the *i-th* subject domain; P^{IPA}= is a set of phonemes peculiar to the Russian language according to IPA; $P^{SAMPA+} =$ is set of phonemes peculiar to the Russian language according to SAMPA+. This alphabet consists of 89 phonemes.

R^{A_i} is a set of ontology ratios within *i-th* subject domain:

$$R^{A_i} = \left\{R_C^{A_i}, R_{C-P}^{A_i}, R_P^{A_i}, R_{P_1P_2}^{A_i}\right\},$$

where $R_C^{A_i}$ is a set of links, which form the hierarchy of ontology terms within *i-th* subject domain; $R_p^{A_i}$ is a set of object properties and data type properties, which determine the relationship between the elements of P^{IPA}, P^{SAMPA+} sets, and also the corresponding properties of these objects. $R_p^{A_i} = \{hasIPA, hasSampaPlus, hasExample...\}$.

$R_{C-P}^{A_i}$ is a set of relations determining the links among the set objects C^{A_i} and P^{IPA}, P^{SAMPA+} within *i-th* subject domain. The given property determines whether the phoneme belongs to the phonetic representation of the corresponding term.

$R_{P_1P_2}^{A_i}$ is a set of links determining the probability that the phoneme P_2 follows the phoneme P_1 within the terms of the *i-th* subject domain. This value depends both on the frequency of the term in the texts of the *i-th* subject domain and on the phoneme sequence in the phonetic representation of a certain term:

$$R_{P_1P_2}^{A_i} = \mu_C^{A_i} \times \rho_{j,k}, \quad (3)$$

where $\mu_C^{A_i}$ is the grade of membership of the term *c* to the subject domain A_i; $\rho_{j,k}$ is the probability that the phoneme k follows the phoneme j within the terminological frames of the subject domain A_i.

The values of these relations are derived from statistical analysis of large text corpus on a given subject domain and from phonetic analysis of each marked term. The task of terminology extraction is solved with the help of semantic algorithms, i.e. thesaurus-based and nested link algorithms [13].

4. Thesaurus-based algorithm

The thesaurus-based algorithm for terminology extraction from a set of words, which belong to a certain text, using the OWL-ontology, calculates the degree of semantic proximity of input word-groups to the terms of the subject domain. This algorithm chooses from the set of incoming words/ word groups only those terms and expressions, which belong to the given subject domain.

The degree of semantic proximity of the input word-group to the subject domain k_{ont} can belong to the interval from 0 to 1: the closer is the obtained value to 1, the greater is the possibility that this word/ word group is a term.

The thesaurus-based algorithm suggests a direct search for input word lemmas and their combinations among the terms defined in a certain ontology. For this purpose, for every type of ontology, it is necessary to define the property “containsLemma”, which has a line value obtained by lemmatizing (reducing to the initial form) the object name with the help of Mystem (Yandex product). The lemmatization is carried out according to certain morphological peculiarities of the term.

The thesaurus-based algorithm consists of the following stages:

1. evaluation of the degree of proximity of the input word / word group to each ontology object;
2. search for the OWL-ontology core object, which is most closely associated with the input word / word group.

The scheme of the algorithm is shown in Figure 1.

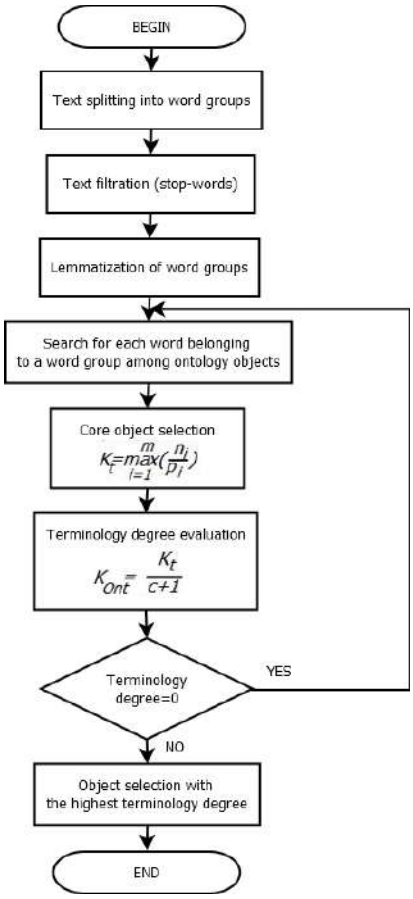


Figure 1. Schematic block diagram of the thesaurus algorithm for extracting terminology.

The reference core object of the expandable ontology, used in the further analysis, has a degree of proximity to the input word / word group, which is calculated by the following formula:

$$k_t = \max_{i=1}^{m}(\frac{n_i}{p_i}),$$

where m is the number of all objects of the OWL-ontology core; p_i is the number of words in the lemma of the reference OWL-ontology core object; n_i is the number of words from the input word group lemma, which are found in the lemma of the OWL ontology core object.

The general scheme to evaluate the degree of proximity of the word groups to the terms of the subject domain is shown in Figure 2.

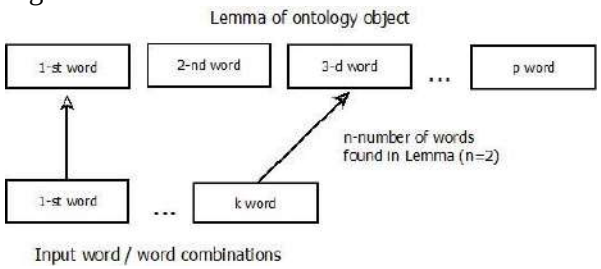


Figure 2. Thesaurus-based criterion. Search for reference object.

In this case, the word order in the group of words in the reference object should not change.

If several different ontology objects have the same value of the coefficient k_t,, then the object that corresponds to the maximum n_i. will be considered as a reference one. If there are several maximum objects, then all of them will be considered reference ones and the analysis according to ontological criterion will be carried out for each of them.

The ontology structure of the analyzed subject domain assumes that each of its objects has a Datatype Property “a Term”, which is of a logical type. This property is auxiliary and is determined by the expert who distinguishes to what extent this ontology object is peculiar for a certain subject domain.

Thus, the terminology degree of the input word group is calculated in the following way:

$$k_{Ont} = \frac{k_t}{c+1},$$

where k_t is a value obtained during the first stage of the algorithm execution; c is the number of ratios between the reference object and the nearest ontology object that has a Data Property “a Term”=true (if this property is true for the reference object, then c=0). The search scheme is shown in Figure 3.

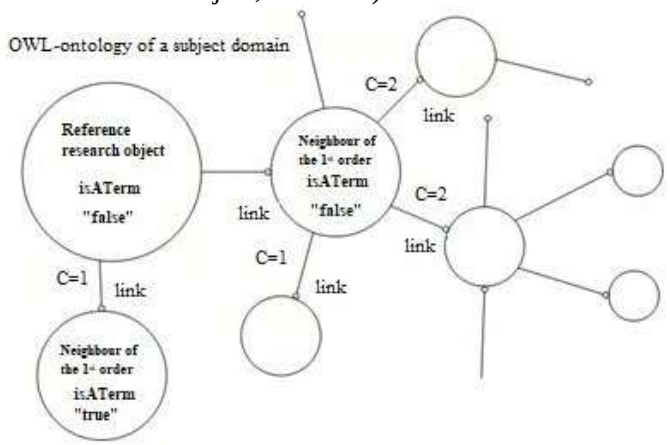


Figure 3. Terminology degree detection of input word groups.

If for the reference object “a Term”=false, and in addition the object has no links with other ontology objects, or all related objects are also distinguished as “false”, then it is necessary to find another reference object for the given word / word group and the evaluate it. In a similar situation with other reference objects, or in the case of their absence, the input word / word group is said to be “not a term” (k_{Ont}=0).

Thus, the process of the terminology degree evaluation of the input word group is a movement along the graph, at the nodes of which there are objects of the corresponding ontology classes.

5. Nested link algorithm

In addition to the terminology degree evaluation of a word / word group, the developed metrics makes it possible to extract terms from the text by comparing them with the existing objects and combinations of lemmas of the corresponding objects with the help of R_{add} ratios, which expand the set of objects of the described subject domain by combining lemmas of related objects. For example: the properties “has a Link” and “is a Part”.

Thus, while comparing the input word groups and the objects of the subject domain, which are related to each other by unidirectional relations R_{add}, the word group will be said to belong to a certain subject domain, if its lemma completely corresponds to the set of lemmas of the corresponding ontology objects.

The nested link method makes it possible to extract terms from the text by comparing them with the existing objects and combinations of lemmas of the corresponding objects using the links defined in the ontology.

The scheme of the algorithm is shown in Fig. 4. The peculiarity of this method is the necessity to represent the ontology objects mainly in the form of single words with the maximum number of links among objects. The determining factors for this method are R_{add}, links, and it is possible to form word combinations in a natural way using these links.

$$t_1 + R_1 + t_2 + R_2 + \ldots + t_i + R_j + \ldots + t_m + R_n,$$

where $R_i \in R_{add}$, $t_j \in T$,T are the terms of the application area, which the ontology describes.

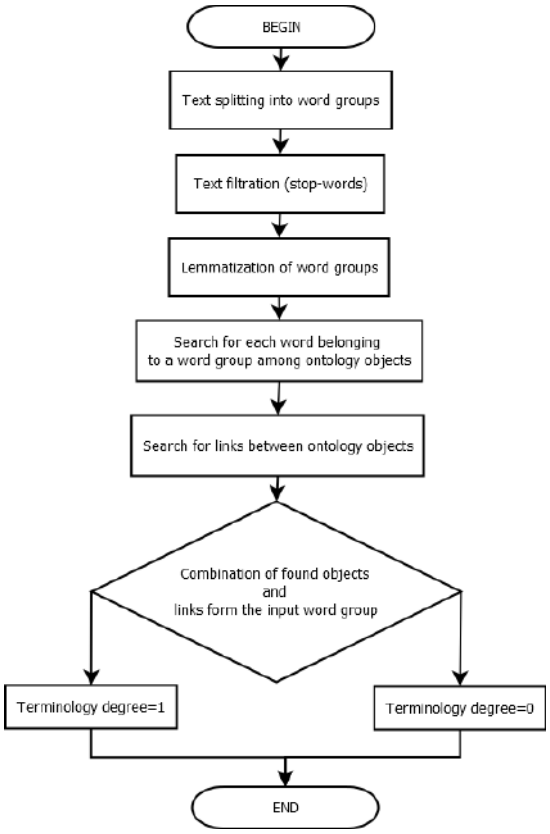


Figure 4. Diagram of the nested link algorithm of terminology extraction.

The scheme of the given algorithm is shown in Fig. 5.

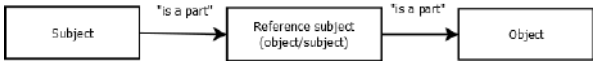


Figure 5. Nested link algorithm.

In this case, the extracted terms, which in its turn belong to terms consisting of a greater number of words, are not considered as terms to avoid redundancy [14].

6. Ontology development

In the course of the research, the linguistic ontology "Cardiovascular diseases" was developed for further recognition of the continuous speech, namely dictated medical diagnoses for this research area.

The ontology has 4 levels of hierarchy, and contains 743 terms belonging to the subject domain. Most of these words were automatically extracted from the texts using the nested link algorithm.

Figure 6 shows a fragment of the developed ontology that contains a description of all the phonemes used in the pronunciation of the extracted terminology.

Moreover, the probabilities of phoneme occurrence, which were found in the extracted terms of the subject domain under consideration, were calculated according to (3). These values will limit the set of selected phonemes while reconstructing the phonetic composition of the recognized words, the speech signal is divided into.

7. Conclusion

The use of lexical ontology implies the possibility to determine the basic set of terms for the analysed subject domain and, as a result, to increase the probability of accurate determination of certain phonetic combinations sequence in the process of speech recognition.

Within the framework of this research, it is planned to carry out a number of experiments to reconstruct the phonetic composition of recognized words using the developed model of the lexical ontology "Cardiovascular diseases" in order to validate the approach effectiveness.

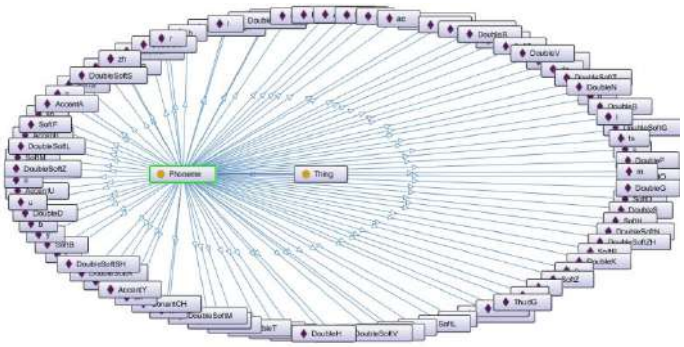


Figure 6. Fragment of lexical ontology. Phonemes.

8. References

[1] Andreev I A, Armer A I, Krasheninnikova N A and Moshkin V S 2017 Attacking the problem of continuous speech segmentation into basic units *III International conference Information Technology and Nanotechnology* 473-476

[2] Andreev I A, Armer A I, Krasheninnikova N A and Moshkin V S 2017 Attacking the problem of continuous speech segmentation into basic units *III International conference Information Technology and Nanotechnology* 6-9

[3] Pienado A and Segura J C 2006 *Speech recognition over digital channels: robustness and standards* (John Wiley & Sons Ltd.) p 257

[4] Keshet J and Bengio S 2009 *Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods* (John Wiley & Sons Ltd.) p 253

[5] Gelbart D 2008 Ensemble Feature Selection for Multi-Stream Automatic Speech Recognition *Technical Report No. UCB/EECS-2008-160* (University of California at Berkeley) p 117

[6] Solov'ev V D, Dobrov B V, Ivanov V V and Lukashevich NV 2006 *Ontologies and thesauri* (Kazan', Mosow)

[7] Fellbaum C 1998 *WordNet: an Electronic Lexical Database* (MIT Press, Cambridge)

[8] Moshkin V S and Yarushkina N 2015 Methods for constructing fuzzy ontologies of complex subject domains *Open Semantic Technologies for the Design of Intelligent Systems* (Minsk: BGUIR) 401-406

[9] Khoroshevskiy V F 2008 Knowledge domains on the Internet and Semantic Web *Artificial Intelligence and Decision Making* **1**

[10] Mikhaylov D V, Kozlov A P and Emelyanov G M 2016 Extraction the knowledge and relevant linguistic means with efficiency estimation for formation of subject-oriented text sets *Computer Optics* **40(4)** 572-582 DOI: 10.18287/2412-6179-2016-40-4-572-582

[11] Mikhaylov D V, Kozlov A P and Emelyanov G M 2015 An approach based on tf-idf metrics to extract the knowledge and relevant linguistic means on subject-oriented text sets *Computer Optics* **39(3)** 429-435 DOI: 10.18287/0134-2452-2015-39-3-429-435

[12] Galunov V I and Solov'ev A N 2004 Modern issues in speech recognition *Information Technology and Computer Systems* 41-45

[13] Andreev I A, Bashaev V A, Kleyn V V, Moshkin V S and Yarushkina N G 2015 Estimation of the terminology of lexical units on the domain ontology basis *Open semantic technologies for the design of intelligent systems* (Minsk: BGUIR) 395-400

[14] Yarushkina N, Moshkin V, Klein V, Andreev I and Beksaeva E 2016 Hybridization of Fuzzy Inference and Self-learning Fuzzy Ontology-Based Semantic Data Analysis *Proceedings of the First International Scientific Conference "Intelligent Information Technologies for Industry"* 277-285

Acknowledgements

This work was supported by RFBR. Projects № 16-48-732046 and № 18-37-00450.

Methods of RFID data processing in intelligent systems for the identification and movement control of industrial products

A V Astafiev[1], A A Orlov[1], D P Popov[1] and M V Pshenichkin[1]

[1]Vladimir State University, Gorkij str. 87, Vladimir, Russia, 600000

Abstract. The article describes the development and research of methods of RFID data processing to build intelligent systems that provide timely and reliable automatic movement control and identification of industrial products. Conducted and presented an analytical review of Russian and foreign scientific-technical base for the development of methods and algorithms for movement control systems. The structure of the hardware-software complex of the system developed. The presented method of movement control products. Experimental studies of the developed system and methods.

1. Introduction

One of the most important elements of the quality management system of industrial production, which largely determines the efficiency of its functioning, is the mechanism of identification, which makes it possible to ensure traceability of products throughout the technological production cycle. Traceability in production helps to ensure compliance with government requirements and international quality standards, to perform a fast and targeted tracking of the entire technological cycle of manufacturing products, which, in turn, minimizes financial consequences. Especially relevant is the question of tracing products in the enterprise, if the production cycle consists of many stages, implemented in large production areas. The organization of the mechanism for tracing products is possible by automating the control of the movement of industrial products.

According to the GOST 18353-79 standard, there are 9 methods of non-destructive testing, of which 2 were widely used in the MCS: radio frequency identification and technical vision [1]. The use of technical vision approaches is complicated by the need for graphic marking of proper quality, which is difficult to realize in real production conditions and requires significant financial and human resources. The use of radio frequency identification is less demanding in the process of marking products. Traffic control, based on radio frequency identification methods, is an advanced information technology for the construction of warehouse accounting systems.

A large number of companies around the world are engaged in developing software and hardware for movement control systems (MCS), but the tasks of developing and implementing automated product tracking systems in production are still unresolved. At present, there are still a number of problems in industrial enterprises, the solution of which is not realized by means of modern MCS. These include:

- not all radio frequency markings may be in the field of view of equipment;
- a number of existing MCS can not be used in view of safety restrictions;
- not all MCS can ensure continuous and correct operation in large industrial areas;
- a number of modern MCS can work only in the presence of positioning systems, which are equipped with transporting devices;

- the presence of interference and signals that make identification difficult.

The listed problems do not allow to organize automatic traffic control at all sections of the production process. To solve these problems, it is necessary to develop new methods that allow more efficient processing of RFID data.

Thus, the development of new methods for processing RFID data for the construction of automatic systems that provide an operative and reliable control over the movement of industrial products is topical.

The aim of the project is to develop and study methods for processing RFID data for the construction of intelligent systems that provide prompt and reliable automatic traffic control and identification of industrial products.

To achieve this goal, it is necessary to solve the following tasks:

1. Analytical review of the Russian and foreign scientific and technical base on development of methods and algorithms of MCS.
2. Development of the structure of the hardware-software complex MCS.
3. Development of an intelligent system and methods for controlling the movement of products.
4. Experimental research and modernization of the developed methods.

2. Analytical review of the Russian and foreign scientific and technical base on the development of methods and algorithms MCS

The PCS software is developed by: PCT-Invent (Russia, Sakt-Petersburg), AiTiProekt (Russia, Moscow), Impinj (USA, Seattle), Motorola (USA, Morrisville), Nordic ID (Finland, Salo), FEIG (Germany, Weilburg) and many others. However, the main activity of these companies is the development and production of the hardware part of radio frequency identification systems and the creation of specialized software for working with it. This implies that for the development of automatic identification systems based on the proposed software and hardware, it is necessary to collect a whole team of technical specialists to write the project "from scratch."

To quickly obtain real-time information on the location of the product and all its movements, systems based on the principles of radio frequency identification of product labels are used. The basis of the systems is the method of remote receipt of product data by transmitting a radio signal from the RFID tag located on its surface to the tags recognition devices (read) and then writing the information received to the database of computers forming the computer network. Such world-famous companies as Wal Mart, METRO Group, Gillette, Procter & Gamble, Tesco, Benetton and others demonstrate on their practical experience the benefits of using RFID technology in the organization of automated warehouse accounting and cargo control.

The automatic identification system (AIS) of objects of rolling stock of railway transport (locomotives, cars, and also large-capacity containers) "Palma" is known, including a corporate computer network, code onboard sensors and reading points equipped at reference points, (RFID tags) the information on the location of the monitored object is automatically removed.

The disadvantage of this AAL Palma is the limited functionality of the system, which consists in the fact that information from the RFID tag is read only when the object passes the specially equipped information reading point (reader). Further location of the facility, for example in the warehouse, is not automatically recorded, which reduces the reliability and efficiency of using AIS when solving the problem of cargo storage.

The closest approach to the proposed approach is the JPL RFID system for warehouse accounting of metal pipes, based on the marking with radio frequency labels of plugs installed at the ends of pipes. This approach allows the inventory of stored products in an automated mode using hand-held portable readers.

The disadvantage of this AIS is that the movement of pipe products is produced by various means of minor mechanization in such a way that the marking installed on the product in most cases does not fall within the field of view of the reading sensors, which makes it impossible to track the products in automatic mode.

A common disadvantage for this kind of AIS is that for the organization of automatic traffic control, readers are installed permanently, and the movement of the tracked products occurs as it

passes along this device. An obligatory requirement is that the marking is in line of sight for the reader, which is practically impossible with the use of such transport devices as bridge cranes, loaders, stackers and other means of minor mechanization. The movement of products for this type of transportation is of a stochastic nature, and the location of the cargo can close the direct visibility of the reader to the marking.

The scientific community is also working on the introduction of RFID technologies in various spheres of human activity. A great contribution to the development of radio frequency identification technology and SKD in various spheres of life was made by Bondarevsky AS, Zolotov RV, Do Zuy Nyat, Kamozin D.Yu., Manish B., Shahram M., Ke-Sheng Wang, Worapot Jakkhupan, Somjit Arch-int, Yuefeng Li, Mahir Oner, Alp Ustundag, Aysenur Budak and many others. The latest works in this area are [2-7].

The application of these knowledge-intensive technologies makes it possible to automate the processes of controlling the movement of industrial products at enterprises of various spheres of life and, ultimately, to improve the efficiency and reliability of the control of transportation and warehouse accounting of manufactured products.

However, they are not without flaws. The use of existing software and hardware solutions is more focused on the organization of automated warehouse accounting and is less suitable for automating traffic control. In confirmation of this at a number of industrial enterprises.

3. Development of the structure of the hardware and software complex MCS

The development of intellectual MCS is aimed at solving the task of organizing and providing control over the movement of products through the territory of industrial enterprises, handling and cargo transportation, including warehousing of goods equipped with multi-code marking. Information on the identification of products can be collected in the database of computers that are part of the computer network that encompasses the points of the cargo transportation route and stores data on the origin and destination of the products and on its location in warehouses.

The proposed MCS consists of 4 main modules: server, client, product marking and marking of the storage area.

Marking of storage areas is performed using Bluetooth-labels, and the current location of the small-scale mechanization means is determined using the IBeacon technology.

Labelling of products at the enterprise is done by attaching an RFID tag to them. In the event that the product contains components shielding the signal from the mark, the marking is applied to several sides of the product.

The client is a stand-alone device equipped with an RFID reader to collect data from RFID tags of products, a Bluetooth adapter for determining the current location, a microcomputer for collecting, processing and sending telegrams to the server and batteries to provide autonomous operation.

The server is a computing device that receives telegrams from client devices, compares the received data with information from the database and generates messages about movements transmitted to the server of the plant's process control system.

The structure of the hardware-software complex of the MCS being developed is shown in Figure 1.

The developed architecture is applicable to various means of minor mechanization. Figure 2 shows an example of placement on a bridge crane.

The proposed system is explained with the following drawing (figure 2), on which it is indicated:

1 - Bridge crane, for moving products;
2 - Bridge crane truck;
3 - Bridge crane beams;
4 - Movable cargo or article;
5 - Customer;
6 - Product identification;
7 - Warehouse area limiter;
8 - Marking of the storage area.

Figure 1. The structure of the hardware and software complex of the developed MCS.

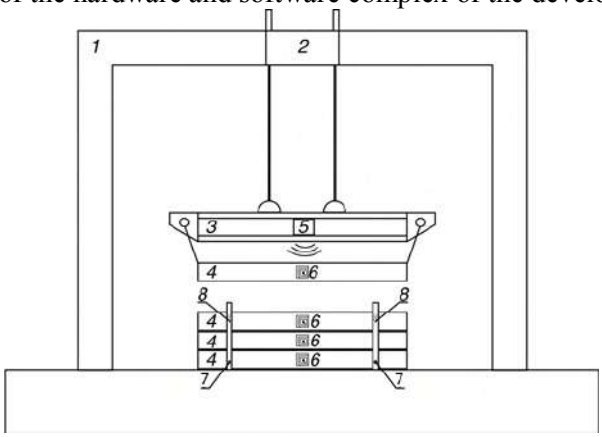


Figure 2. General scheme of means of moving pipes.

4. Development of an algorithm for determining the threshold for clipping markers that are not involved in the movement

The aim of the work is to develop an algorithm for filtering RFID tags, which makes it possible to exclude marks that are not involved in the course of movement and to conduct a series of laboratory tests of the developed system for automatic control over the movement of products along untyped routes.

A disturbance is an external perturbation acting in the transmission system and preventing proper reception of signals. Sources of interference may be either outside or inside the transmission system itself. If the interference is regular and known, then fighting it is not difficult. For example, the background of an alternating current can be eliminated by compensation; interference from a particular radio station with a modulation spectrum of normal width is eliminated by an appropriate filter. The struggle with random interference presents the greatest difficulty.

There are many methods of detecting interference. The following were considered in the work:

1. The arithmetic mean.
2. Determination of emissions in statistics.
3. Determination of the optimum noise level in the flow of spectra.
4. Level of significance.

Two shelves have been created, each of which is marked with an iBeacon-tag. Portable products are marked with RFID tags. Moves three products 1, 3, 4 from the first rack to the second one. Items c labeled 2 and 5 are located between the shelves and are not subjected to transfer. Table 1 presents the primary data.

During the move, 23 signal level values (RSSI) were received from all RFID tags.

To determine the threshold value of p, which will allow us to weed out marks that are not involved in the movement, four methods were tested (Figure 3 a-d).

Table 1. Results of displacement.

Label	1	2	3	4	5
Readings	69	56	71	68	53
	71	52	75	75	-
	71	-	74	76	-
	77	-	77	83	-
	79	-	77	80	-
	78	-	79	81	-
	78	-	81	-	-
Total	523	108	534	463	53

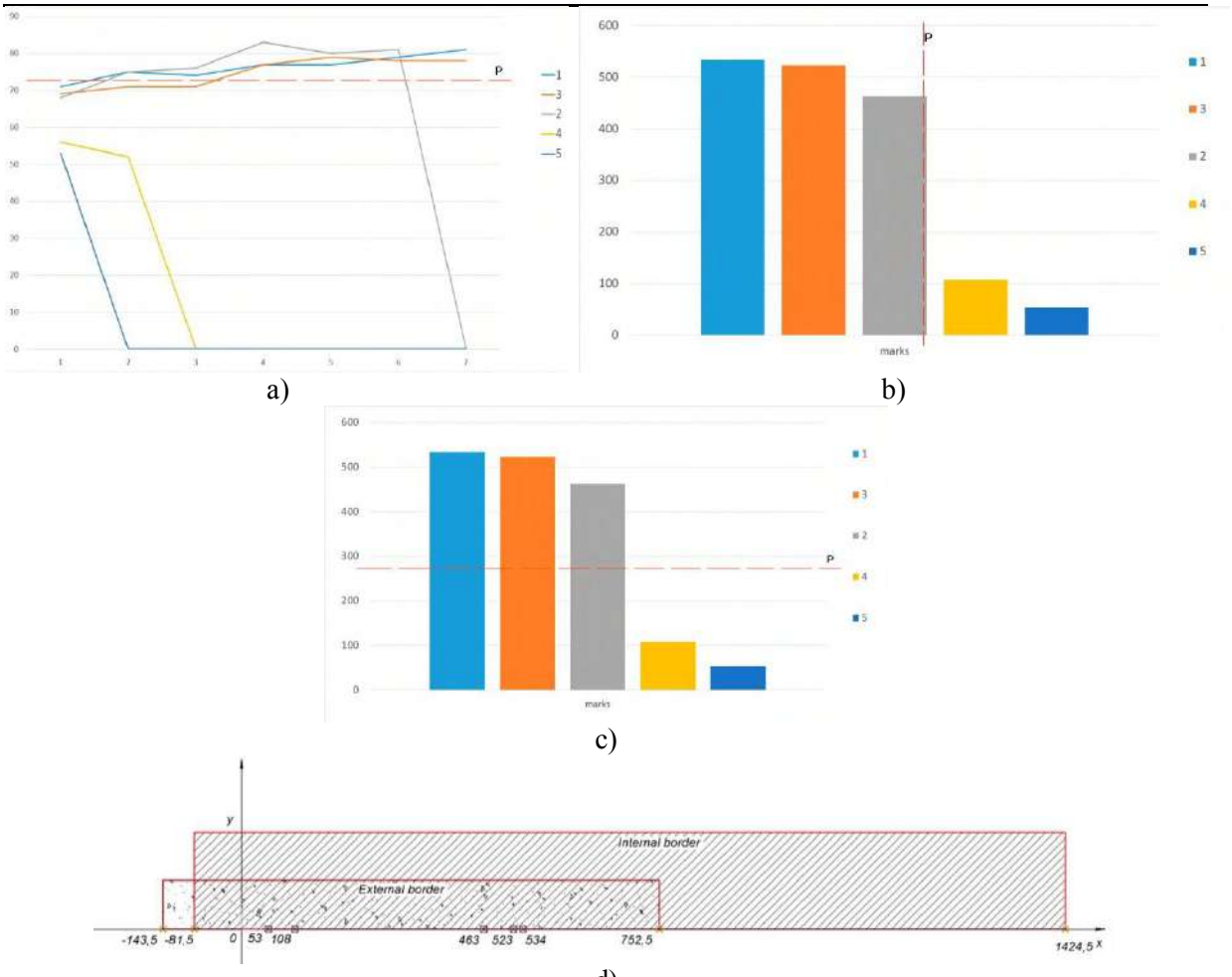


Figure 3. a) Allocation of useful data on the threshold P, based on the arithmetic mean, b) Determination of useful data on the histogram, c) Determination of the threshold by significance level p=0.10, d) Determination of emissions.

Laboratory investigations were carried out, during which it was established that the method based on the level of significance was the most suitable method for testing the tested methods. The determination of noise by percentage of the area or the arithmetic mean does not always yield correct results. The emission detection method can not be used for this task, since false signals are not discarded, due to the small sample that is obtained in most cases.

5. Experimental results

To test the evaluation of the reliability of the results obtained, an experimental study was carried out. The subject of the study is a system for automatic control over the movement of products. The object is a product, an enterprise or a warehouse.

Laboratory experience was conducted with a prototype system in conditions close to real production. To the moving cart was attached a model of the part of the beam, with the reading and processing device mounted on the side. Under the traverse on the cart is mounted a pallet for products in the form of pipes. View of the laboratory setup is shown in figure 4.

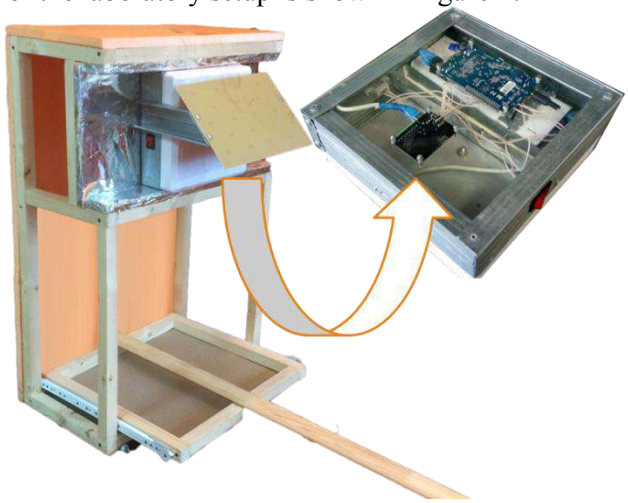


Figure 4. View of the laboratory setup.

For the experiment was created two racks (c1, c2), each of which is labelled iBeacon-tag. The movement of products made from rack 1 to rack 2. During the movement of production readings from their RFID-tags and iBeacon-tags of racks have been processed and are presented in tables (Table 2, 3).

Table 2. Readings from iBeacon-tags of racks.

Time	Rack 1	Rack 2
35:22:00	0	43
35:28:00	0	26
35:34:00	0	0
35:40:00	0	0
35:46:00	3	0
35:52:00	34	0
35:58:00	37	0
36:04:00	27	0
36:10:00	27	0
36:16:00	16	0

Table 3. Readings from RFID tags of products.

Time	Product 1	Product 2	Product 3	Product 4
35:22:00	65	16	91	89
35:28:00	61	5	101	101
35:34:00	75	20	105	105
35:40:00	81	11	108	107
35:46:00	73	9	110	109
35:52:00	57	21	98	95
35:58:00	38	7	103	110
36:04:00	38	21	99	110
36:10:00	61	2	107	110
36:16:00	35	3	78	80

The duration of the experiment was 1 minute. During the experiment with iBeacon-tags of rack has received 212 readings and from RFID tags of products - 2715. Graphic interpretation of the data presented in figures 5, 6.

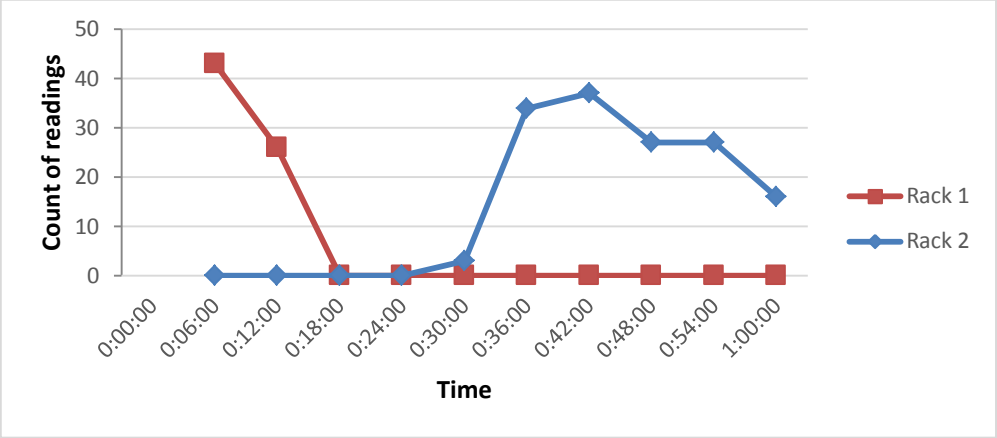


Figure 5. Readings from iBeacon-tags of racks.

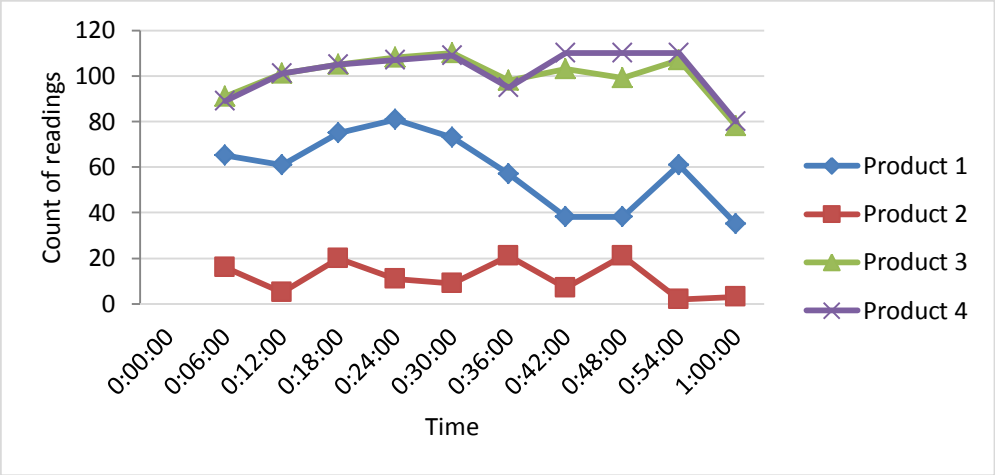


Figure 6. Readings from RFID tags of products.

From the graph in figure 4 shows that at time 00:18:00 readings from iBeacon-tag of the first rack were no longer received by the reader. At time 0:24:00 began to receive readings from iBeacon-tag of the rack number 2. This fact indicates that there was a movement of the transport device from the area of the rack 1 to the area of the rack 2.

Figure 5 shows the readings from the RFID-tags of products, time-spaced with an interval of 6 seconds. During the experiment, 4 items were moved. According to the graph it is clear that a stable signal came from all 4 product tags. This allows us to say that during the movement of the conveying device the products moved along with it.

In the course of the work, more than 150 experiments were conducted in the laboratory. Also, the installation was tested in an industrial plant. The results of the experiment showed the reliability of the movement identifications in the amount of 97.3%. As a result, information was also collected that allowed adjustments to the system to improve its efficiency.

6. Conclusion

A system and method for automatic control of the movement of products based on RFID-identification when moving hoisting-and-transport mechanisms is developed. Considered and analyzed the existing systems of automatic control of the movement of products based on radio frequency identification. The main advantages and disadvantages are revealed. The structure and levels of the hardware and software complex MCS are described, the methodology of its operation is described. Thanks to the presented system and methodology, the problems of moving several units of products simultaneously were solved, the accuracy of detail of the location information was high, the client-server approach and the movement history analysis algorithms were implemented, the problem of authenticity of

identification and the tracking process outside the route was solved. Laboratory investigations were conducted, during which it was possible to obtain a high degree of reliability of the identification of displacements in 97.3%, proving the urgency of the developed system and methodology. These studies bear a high scientific value for providing control over the movement of products.

7. References

[1] *GOST 18353-79 Nondestructive check. Classification of types and methods*
[2] Astafiev A V, Orlov A A and Popov D P 2017 Development the algorithm of positioning industrial wares in-plant based on radio frequency identification for the products tracking systems *CEUR Workshop Proceeding* **1901** 23-27
[3] Astafiev A V, Orlov, A A and Privezencev D G 2016 Method of controlling the movement of large metal products with the use of algorithms for localization and recognition of bar code markings *Dynamics of Systems, Mechanisms and Machines, Dynamics* **7819046** DOI: 10.1109/Dynamics.2016.7818969
[4] Orlov A A, Provotorov A V and Astafiev A V 2016 Methods and algorithms of automated two-stage visual recognition of metal-rolling billets *Automation and Remote Control* 1099-1105 DOI: 10.1134/S000511791606014X
[5] Zhiznyakov A L, Privezentsev D G and Zakharov A A 2015 Using fractal features of digital images for the detection of surface defects *Pattern Recognition and Image Analysis* **25(1)** 122-131
[6] Nhat D D 2015 Researches and application of RFID technology (radio frequency identification) *International Research Journal* **5** 34-37
[7] Kamozin D Y 2013 Comparison of the effectiveness of bar-code technology and RFID technology's application in logistics processes *Bulletin of Baikal State University* **3** 71-75
[8] Parikh D and Jancke G 2008 Localization and segmentation of a 2D high capacity color barcode *Proceeding of IEEE Workshop on Applications of Computer Vision* 1-6 DOI: 10.1109/WACV.2008.4544033
[9] Kazanskiy N L and Popov S B 2012 The distributed vision system of the registration of the railway train *Computer Optics* **36(3)** 419-428
[10] Morozov A A and Sushkova O S 2016 Analysis of real-time video images using the means of the Actor Prologue *Computer Optics* **40(6)** 947-957 DOI: 10.18287 / 2412-6179-2016-40-6-947-957
[11] Epifantsev B N, Pyatkov A A and Kopeikin S A 2016 Multisensory systems for monitoring restricted areas: the capabilities of a video analytics channel for intrusion detection *Computer Optics* **40(1)** 121-129 DOI: 10.18287 / 2412-6179-2016-40-1-121-129

Development of an algorithm for forecasting and preventing emergency situations in industrial traffic control systems based on data analysis of multi-code labels

A V Astafiev[1], A A Orlov[1] and T O Shardin[1]

[1]Vladimir State University, Gorkij str. 87, Vladimir, Russia, 600000

Abstract. This article proposes a method for controlling the movement of industrial products based on data from multi-code labels using the algorithm for forecasting and preventing extraordinary situations. In the course of the work, possible extraordinary situations arising in the process of displacement were analyzed. Also on the basis of the developed algorithms, simulation modeling was carried out, according to the results of which it was shown the effective use of this method for controlling the movement of products in industrial enterprises.

1. Introduction

At present, for the identification of industrial products, product marking is mainly used with the help of barcode [1] or radio frequency tags [7-10]. These solutions do allow you to shorten the time of searching for a product on the territory of the enterprise, however, they are not without shortcomings, since they are not able to forecast possible supernumerary situations.

For example, when moving a load, the label is not always in the sensor's field of view, so that the marking is not visible to the reader or person. Although it is worth noting that for security reasons, people in such jobs are practically not involved. Also, the marking, during transport or movement on the conveyor belt, may be poorly secured, as a result, it may fall off or fall onto another product. Proceeding from this, it can be concluded that for reliable product identification, several markings on one product should be used, the number of which can depend on the geometric parameters of the object. To accurately control the location of products, you should use the algorithm for finding errors in the occurrence of abnormal situations (for example: when several labels identify one label from several possible ones), which will promptly make a decision to the operator and eliminate the violation.

The purpose of this work is the development of an algorithm for forecasting and preventing emergency situations in traffic control systems for industrial products based on the analysis of data of multi-code labels, during which it is necessary to perform the following tasks:

- conduct a comparative analysis of analogue systems;
- develop a simulation model of the process of moving products;
- to formulate freelance situations;
- develop an algorithm for forecasting and preventing extraordinary situations;
- carry out an experimental study using simulation tools.

2. Overview of analogues

At the moment, to implement the goal, there are several similar solutions. We will perform a comparative analysis of these analogs, as a result of which we can conclude which system is more profitable to use in the future:

1. VITRONIC - automatic recognition system is used to read barcodes in various industries. The results of a comparative analysis of this system are given in Table 1:

Table 1. Advantages and disadvantages of using VITRONIC.

VITRONIC	
Advantages of use	Disadvantages of use
Recognition of different types of codes (1D, 2D) Reading OCR (serial numbers, batch numbers, articles, etc.) Multi-code reading	The implementation of an algorithm for monitoring contingencies is required

3. OptiCode is an industrial scanner for high-speed barcode reading. Advantages and disadvantages are presented in Table 2:

Table 2. Advantages and disadvantages of using OptiCode.

OptiCode	
Advantages of use	Disadvantages of use
Tracking the product from the moment of receipt to the warehouse Scanning barcodes from a long distance (15m) Allows you to position the camera on cranes or carts	The implementation of an algorithm for monitoring contingencies is required No multi-code reading support

As a result of the comparative analysis of the presented analogs, it can be concluded that these solutions are not entirely suitable for use, since they basically do not have an algorithm for forecasting and preventing contingencies. In some cases, there is no support for multi-code reading, as a result of which a significant amount of funds will be required for the development, which is unprofitable.

3. Development of a simulation model for the movement of industrial products

At the enterprise every 100 ± 2s there are applications for receiving products. Then the products are moved to storage racks. Processing of such applications takes 180 ± 2s. After the work can move to the neighbouring shelves or the place of shipment. The processing takes 150 ± 2s and 160 ± 4s, respectively. In the case of accepting applications for participation or the remaining applications do not have time to process within a specified period of time, they are automatically lost. It is required to simulate the work by moving industrial goods during one working shift. The scheme of this simulation model is shown in Figure 1.

It should be noted that during the operation of the simulation model, the following contingencies may arise during the transportation of products, which must be taken into account:

1. Supernumerary situation 1 – during the movement of the product, the same marking is read.
2. Supernumerary situation 2 – in the process of identification, a non-existent marking fell into the field of view of the reader.
3. Supernumerary situation 3 – During the transfer to the product, the marking of another object, by mechanical action (marking off) or deliberate action of personnel (intentional re-gluing of the marking) fell on the product.

4. Development of an algorithm for forecasting and preventing emergency situations during the movement of products

Consider the algorithm for predicting and preventing emergency situations by steps:

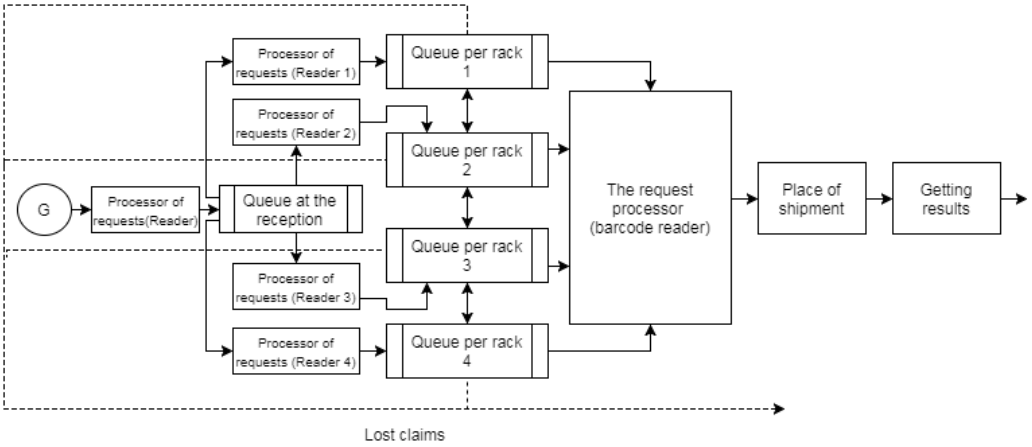


Figure 1. Model outline.

1. When goods arrive at the place of reception, we fix the identified product markings, the time of arrival, and its location.

2. When moving, the time, the current location and location of the rack, to which the product enters, is fixed.

3. Receipt on the rack is the same as in paragraph 1 with the comparison of marks.

4. Moving between the shelves and when entering the place of shipment occurs, in accordance with paragraphs 1-3.

5. If during the movement of the product from the storage areas the same marking is received in the field of view of the reader, the system generates a warning for its verification, which corresponds to an abnormal situation 1.

6. If a non-existent marking has appeared in the field of view of the reader or from another product, we generate a warning, which corresponds to abnormal situations 2, 3.

The flowchart of the algorithm is shown in Figure 2:

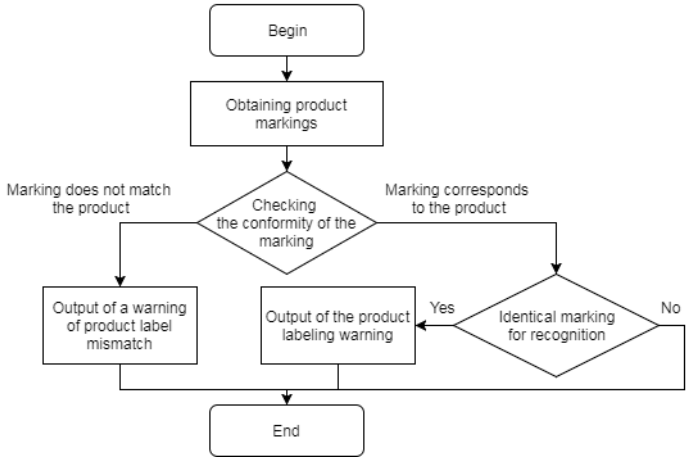


Figure 2. Block diagram of the algorithm being developed.

5. Experimental study with barcodes

5.1. Description of the experiment

Before the implementation of the algorithm and model, an experiment was conducted to collect all the necessary input data.

The experiment consisted of the following:

1. We took an object of a cylindrical shape, placed on it 4 barcodes. There can be several items, since in real production when moving products with a crane or a conveyor belt, several products are simultaneously received in the field of the reading sensor.

2. With the help of any technical device that allows you to capture an image (for example: a mobile phone camera or a camera), you took a few pictures, on which the products were imaged in different positions (different positions are meant).

3. On the basis of the collected images, a representative sample of the read barcodes was made and all the data entered into the table.

4. As a result of the collected information, the probability of identified markings was calculated, which later allowed modeling and implementing the necessary algorithms:

5.2. Image analysis

During the experiment, 3 cylindrical objects were taken, on each of which 4 barcodes were attached. For multi-code labeling, a Code-11 bar code was used, as shown in Figure 3:

Figure 3. Example barcode used for marking products.

In order to get a more accurate result, you need to photograph as many different cases as possible. The result is shown in Figure 4:

Figure 4. Example of images taken during shooting.

As can be seen from Figure 4, with different positioning of the object, in most cases 1 or 2 barcodes are identified. It is worth noting that on some images, there were recognized 3 marks on one product, but this result is extremely rare and can be neglected, since the label is practically not visible. Based on the images received, a sample was taken, the results of which are summarized in Table 3:

Table 3. Total number of identified labels.

Image No.	**Number of identified barcodes**												**Unrecognized barcodes**
	Pipe 1				Pipe 2				Pipe 3				
	Barcode number on the pipe				Barcode number on the pipe				Barcode number on the pipe				
	0000 0001	0000 0002	0000 0003	0000 0004	0000 0005	0000 0006	0000 0007	0000 0008	0000 0009	0000 0010	0000 0011	0000 0012	
1				1			1	1			1		1
2	1				1			1			1		0
3	1			1			1	1			1	1	0
4				1			1					1	2
…	…	…	…	…	…	…	…	…	…	…	…	…	…
1597			1				1	1			1		1

1598			1				1				1	1	1
1599	1			1				1	1				2
1600		1					1				1		1
Total Recognized	680	240	280	720	80	80	760	760	720	280	400	320	1840
Total		1920				1680				1720			1840

According to the information presented in Table 3, you can find out the numeric data on the recognition of only 1 barcode, 2 or more:

Table 4. Number of read tags on each object.

Option	Number of read tags		
	Pipe 1	Pipe 2	Pipe 3
1 barcode recognized	920	960	1020
2 barcodes recognized	680	640	580
3 or more barcodes recognized	0	0	0

As a result, from the presented tables it is possible to calculate the probability of barcode identification on the product:

1. Probability of recognition of 1 barcode: 66,67%
2. The probability of recognizing 2 barcodes: 33.33%

6. Experimental study with RFID technology

6.1 Description of the experiment

This experiment allowed us to determine the angle at which RFID readers should be installed for optimal recognition for different forms of products.

The experiment consisted of the following:

1. We took a cylindrical object.
2. Was pasted with RFID tags from different sides.
3. Next, the object was placed arbitrarily on a flat surface.
4. Using RFID reader determined how many labels are considered at a certain angle, namely: at an angle of 0 degrees, 45 degrees (vertical), and 45+45 degrees (vertical and horizontal), for two seconds.
5. The obtained data were entered in tables for further plotting.

6.2 Results of the experiment

Figure 5 shows the process of reading RFID tags by the reader:

Figure 5. The process of reading RFID tags by the reader.

The data of the read marks from the cylindrical object are shown in table 5:

Table 5. Number of read RFID tags.

№	Angle	Label 1		Label 2		Label 3		Label 4	
		Count	RSSI	Count	RSSI	Count	RSSI	Count	RSSI
1	0	13	54	10	55	8	56	10	66
2		5	60	3	53	7	61	6	59
3		4	61	6	51	4	55	7	58
4		5	52	8	48	4	56	4	59
5		3	58	6	62	7	60	5	59
…		4	55	7	58	7	55	3	60
97		2	69	13	52	13	58	13	61
98		4	58	6	59	6	55	5	65
99		6	57	7	58	5	56	3	64
100		6	54	7	55	3	67	5	60
1	45	5	63	5	58	7	57	4	64
2		5	53	4	55	5	66	7	54
3		5	61	6	49	5	66	5	53
4		6	59	5	54	6	64	4	55
5		7	58	6	52	4	68	4	64
..		4	64	7	57	3	62	7	57
97		6	58	6	58	9	54	0	0
98		6	56	6	57	9	56	0	0
99		3	64	8	51	4	56	6	60
100		5	56	2	51	7	55	7	56
1	45+45	4	61	5	55	7	59	5	62
2		6	54	6	56	4	66	5	60
3		6	64	4	64	6	64	5	61
4		4	52	7	59	5	59	5	59
5		5	51	6	51	6	58	4	64
…		11	50	11	47	10	55	9	72
97		6	49	6	50	5	61	4	63
98		5	52	8	54	6	71	2	59
99		10	61	11	57	9	58	11	58
100		8	59	11	56	12	67	10	61

According to the information presented in table 5, you can find the average number of each read mark, depending on the position of the RFID reader:

Table 6. Average number of tags read.

Average number of tags read				
Angle	Label 1	Label 2	Label 3	Label 4
0	5	7	6	6
45	5	5	6	4
45+45	9	8	10	12

According to the results presented in tables 5, 6, you can see that the read marks are most observed at an angle of 45 + 45 degrees (vertical and horizontal), since the sensor field of view is increased, it allows you to capture more marks than in other versions of its RFID reader placement.

7. Results of the simulation model work

Based on the scheme of the developed simulation model, presented in Figure 1, it can be seen that each movement of the object is controlled by reading the markings with the help of the handler of applications from the place of receipt of the product to its shipment. In this handler there is a comparison of the markings and further decision-making on the output of messages in the event of abnormal situations during transportation (Figure 6, 7):

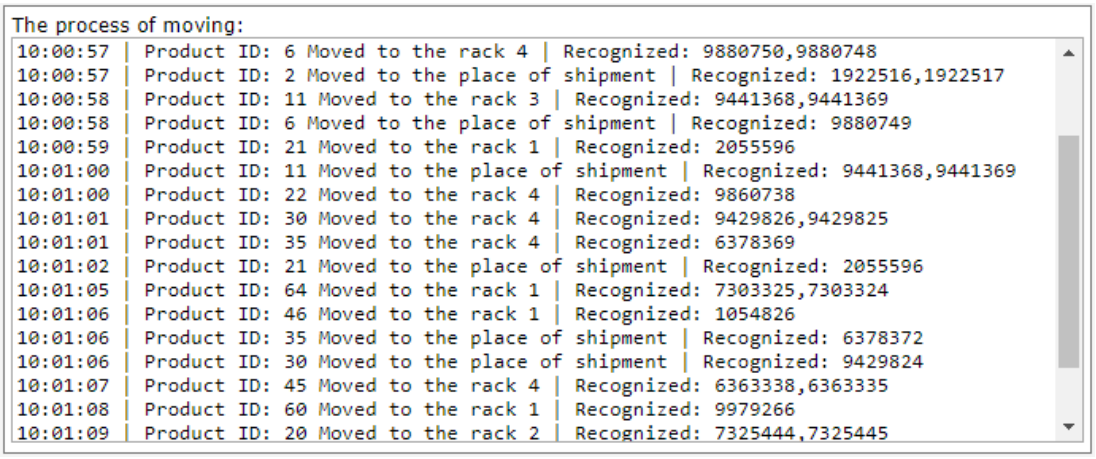


Figure 6. Displaying the progress report.

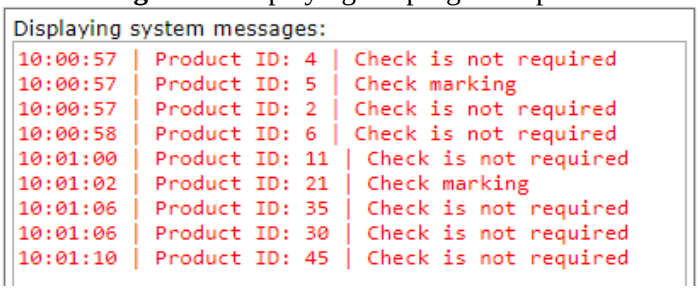


Figure 7. Displaying system messages during movement.

The figures show the work of the algorithm for predicting and preventing extraordinary situations. For example, if you take the product number 21, then during the move it can be noticed that on arrival in the rack 1 the reader identified only one marking (barcode number: 2055596). After receipt of the product at the place of shipment, the same marking is observed, resulting in the generation of a message and notification of a possible violation in the marking of the product for the purpose of checking it. An example of a general simulation result report is shown in Figure 7:

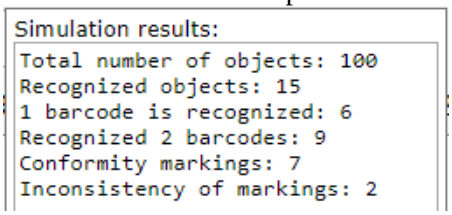


Figure 7. Example of a general simulation results report.

8. Conclusion

On the basis of the work done, it can be concluded that when using several labels for product identification, the probability of recognition increases, since regardless of positioning in any case, 1 mark will be visible to the reader. Based on the results of the input data, an imitation of the operating model was developed using the algorithm for predicting and preventing extraordinary situations. The results of simulation have shown effective use for business and implementation in the enterprise.

9. References

[1] Parikh D and Jancke G 2008 Localization and segmentation of a 2D high capacity color barcode *Proceeding of IEEE Workshop on Applications of Computer Vision* 1-6

[2] Astafiev A V, Orlov A A and Popov D P 2017 Development the algorithm of positioning industrial wares in-plant based on radio frequency identification for the products tracking systems *CEUR Workshop Proceeding* **1901** 23-27

[3] Astafiev A V, Orlov A A and Privezencev D G 2016 Method of controlling the movement of large metal products with the use of algorithms for localization and recognition of bar code markings *Dynamics of Systems, Mechanisms and Machines, Dynamics* **7819046** DOI: 10.1109/Dynamics.2016.7818969
[4] Orlov A A, Provotorov A V and Astafiev A V 2016 Methods and algorithms of automated two-stage visual recognition of metal-rolling billets *Automation and Remote Control* 1099-1105 DOI: 10.1134/S000511791606014X
[5] Zhiznyakov A L, Privezentsev D G and Zakharov A A 2015 Using fractal features of digital images for the detection of surface defects *Pattern Recognition and Image Analysis* **25(1)** 122-131
[6] Nhat D D 2015 Researches and application of RFID technology (radio frequency identification) *International Research Journal* **5** 34-37
[7] Kamozin D Y 2013 Comparison of the effectiveness of bar-code technology and RFID technology's application in logistics processes *Bulletin of Baikal State University* **3** 71-75
[8] Kazanskiy N L and Popov S B 2012 The distributed vision system of the registration of the railway train *Computer Optics* **36(3)** 419-428
[9] Morozov A A and Sushkova O S 2016 Analysis of real-time video images using the means of the Actor Prologue *Computer Optics* **40(6)** 947-957 DOI: 10.18287 / 2412-6179-2016-40-6-947-957
[10] Epifantsev B N, Pyatkov A A and Kopeikin S A 2016 Multisensory systems for monitoring restricted areas: the capabilities of a video analytics channel for intrusion detection *Computer Optics* **40(1)** 121-129 DOI: 10.18287/2412-6179-2016-40-1-121-129

Method monitoring of movement in the task of indoor navigation

M P Osipov[1] and V S Andreev[1]

[1]Lobachevsky State University of Nizhni Novgorod, Prospekt Gagarina 23, Nizhni Novgorod, Russia, 603950

Abstract. The paper considers ways to minimize positioning errors by using information obtained with the help of MEMS sensors available to the user. An approach for correcting a positioning error by using information about a floor plan of a building is proposed. A set of methods is presented to ensure monitoring of the user's movement in the task of navigation.

1. Inroduction

Inner structure of airports, railway stations, shopping and recreational centers have complex configuration. Present-day buildings represent a huge complex of floors with staircases, corridors and rooms. Navigation in such buildings often appears a challenging task. For instance passengers can encounter difficulties trying to find way to the needed office in a huge business or shopping center, determine evacuation routes and arrange salvage actions in case of emergency etc. Movement monitoring is essential for providing users with various kinds of notifications. For example, informing about nearest shops purchase incentives or mobile guide service in a museum.

Usage of global positioning navigation systems such as GPS-GLONASS inside buildings is rather problematic subject to concrete walls and partitions which bring signal level to nothing. A range of visual landmarks on the walls are traditional navigation solution for such buildings. But the draw-back of such an approach is information overload and complexity of analysis of the received information. Technology companies all over the world work over automatization of user location search in enclosed spaces.

There are several approaches to solving this task based on using: different data communication networks (Bluetooth, Wi-Fi, FM-ether waves, UWB); the Earth's magnetic field; RFID-tags; QR codes; MEMS sensors etc. [1] However, the mentioned approaches either lack accuracy and reliability or involve complexity of setting the corresponding infrastructure and its maintenance or require significant investments.

2. Problem Statement

The aim of this paper is to develop an approach of a standalone inertial navigation based on using (MEMS) and information about building layout. This approach is the "cheapest" among the existing navigation technologies, but positioning accuracy index is not so high. MEMS installed on mobile devices do not show high accuracy. Such MEMS positioning error value eventually increases and comes out at

10% of the distance covered, that makes them unreliable for navigation purposes. Positioning errors in complex configuration of corridors and rooms can bring about incorrect evacuation routes determination in case of emergency. The paper puts forward algorithms ensuring positioning accuracy and reliability increase.

3. Inertial navigation methods

Usage of micromechanical inertial sensors allows to perform movement tracking along the route without additional installations inside the buildings. MEMS include 3 types of sensors: accelerometer (measures acceleration along the axis), gyroscope (measures angular rate along the axis), gaussmeter (device for earth magnetic field characteristics measuring). Accelerometer is used for counting number of steps made by the user. Gyroscope and gaussmeter are used to determine the step direction.

However, the algorithms based on movement sensors show some drawbacks which need to be levelled down in course of algorithm execution. The low-cost gyroscopes show high displacement errors which constantly accumulate during turning angle. Geomagnetic disturbances can cause compass errors.

It appears impossible to build a universal model of human movement due to individual peculiarities (pace, parameters of a person) which lead to glitches of step detection and consequently to incorrect user's position determination.

Traditional approach to step detection according to accelerometer readings include 2 stages. Primarily accelerometer readings filtration is performed by means of a low-pass filter in oreder to decrease the noise impact. Further the task reduces to finding signal extreme values (maximum and minimum) within a certain time frame. Various methods of glitches excluding are used therewith.

Algorithms executing on this principle are described in the works [2-5]. It was experimentally determined that the three-stage algorithm shows the most reliable results in step detection among the described algorithms.

In this case MEMS sensors are placed onto the user’s feet. The algorithm essentials are in catching the user's foot motion path in the course of making a step. Human step cycle comprises three phases: phase of a stroke, phase of heel contact, resting phase. Hence step detecting can be performed by means of the phase determination. A certain number of step detecting glitches occurred during the three-phase algorithm testing.

Therefore it was decided to improve the existing algorithm towards increasing severity of the step detecting criterion. Additionally, the algorithm should take into account physical peculiarities of the user and adjust the threshold of detection in the course of execution.

Modification of the three-phase algorithm consists in the usage of a configurable phase pattern on each stage of step phase determination, formed on the grounds of information about previous user’s step. At the moment of the first step phase determination additional verification is started, specifically each new point is compared with the existing pattern with account for possible error considered as input parameter.

If the point does not fall within the predefined locality of the pattern point the step is considered not detected and flags of the determined phases are dropped. After the third phase is determined the step is considered detected.

It is worth noting that the predefined locality of the acceleration pattern varies for each phase, as the behavior of acceleration function on each phase is different. The pattern is formed as arithmetic average of each point of the current step with the already existing step; the pattern is initiated by means of the first detected step.

The algorithm execution is depicted on figure 1.

The conducted experiments showed that the proposed modification of the three phase algorithm makes it possible to increase the accuracy of the step determination by 20% in comparison with the classical approach.

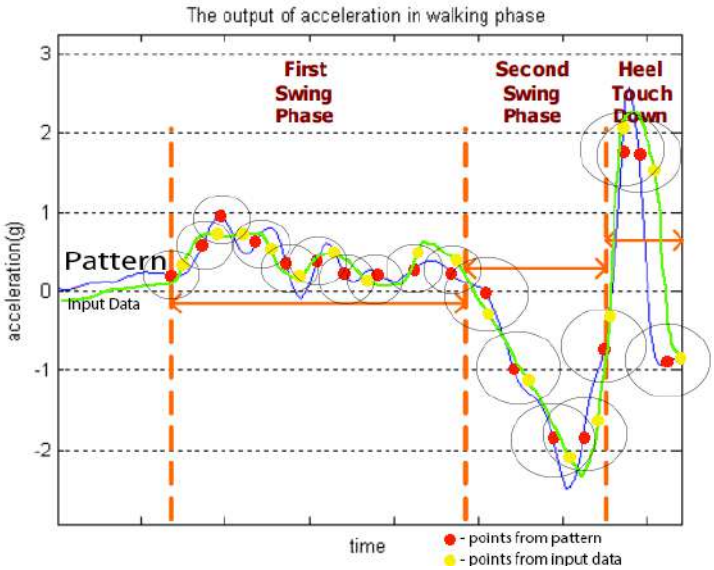


Figure 1. Three-phase algorithm modification.

4. Movement monitoring with the use of building floor plan

The major drawback of navigation methods based on MEMS sensors readings is the accumulation of measuring error in the course of movement. Using the information about digital floor plan of the building allows compensating the error accumulation. Distributed database management system is used for storing a digital floor plan and is based on an object-oriented approach to the representation and processing of complex structured spatially distributed information [6]. Floor plan contains information about the building structure in the form of connectivity graph of rooms, corridors, staircases, elevators etc. [7]. Connectivity graph allows predefining all the expectable paths of user's movement, out of which the optimal path [8] is selected with account of specified restrictions [9]. In the course of project implementation the algorithm was developed that allows making correction of MEMS sensors positioning error based on the information about expectable paths of user's movement. After sensors detect users turning movement the search of the nearest turn on the floor plan in the specified direction is performed [10]. The found turn makes it possible to correct positioning. Therefore the accumulated MEMS sensors measuring error is reduced [11]. The algorithm executions result is displayed on figure 2.

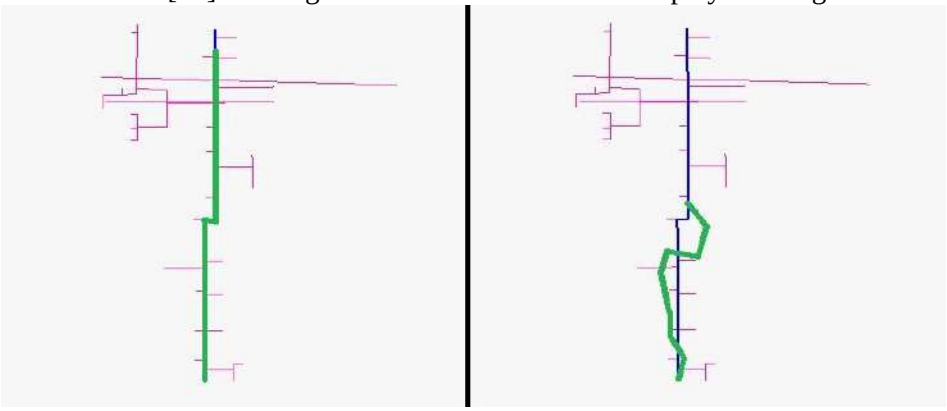


Figure 2. Results of inertial navigation algorithms with movement monitoring and without it.

Due to imperfection of readings of micromechanical inertial sensors some false turns can occur. In such a case user's movement path can deviate considerably. To solve this problem it is recommended to

use not only the main movement path but also alternative ones in the floor plan configuration when user reaches the turn (crossroads). Each alternative movement path has a priority parameter, which is going to be decreased in the course of further guiding along the route, if the determined by inertial sensors expected users position does not correspond to the floor plan configuration. Paths with the lowest priority are eliminated. Path with the highest priority gets current. The algorithm allows determining the actual user's movement path in a complex environment of branched rooms and corridors configuration.

Algorithm of movement paths prioritization is based on two principles:

- Principle of direction coincidence.
- Principle of configuration coincidence.

Principle of direction coincidence is based on the following statement: if the current path is true, the current vector of path should be co-directed with the current path vector, received by means of standalone navigation.

Principle of configuration affirms: if the current step made by a user does not belong to any of the current path sections, the route is considered incorrect as the points not belonging to path sections belong to different physical barriers in the process of movement (wall sections).

Let us consider how the algorithm uses the mentioned principles. The algorithm is based on the idea of random process, basically on the Markov Chain model. A number is assigned to each movement path. Every such number is referred to a probability which shows that at a certain time point t user is on the current section (or current movement path), the probability is selected in such a way that the norm axiom should be true for a random value, namely the sum of the probabilities for each number should be equal to 1. And the conclusion from this is that at each time point t the system can be described as a discrete random value. It should be noted that the maximum amount of paths M that are examined at each time point t equals to 6. The number was chosen based on the two heuristics:

- Crossings dividing the path of movement by two usually prevail in each building configuration.
- Difference in any two paths configuration is possible to detect after the first crossing, as lengths of corridors are widely diverging.

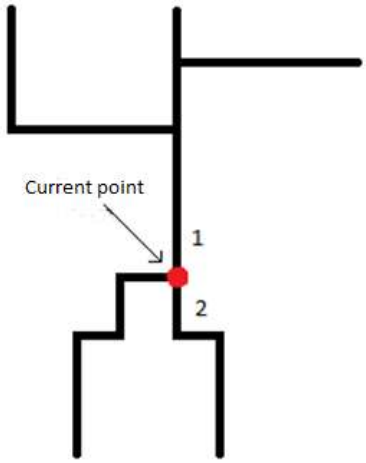


Figure 3. Diagram of movement paths configuration.

Figure 3 shows the standard situation, the path is divided by two and path 1 is much different from the path 2 according to its configuration (at least according to the length of segments next to the crossing)

By this means at each time point there is a discrete random parameter which takes a value of $i \in \overline{\{1,6\}}$, and shows a degree of confidence in the i-th path. Probabilities for such random value are calculated according to the formula:

$$P_i^t = aF\left(f(x_i^{t-1} + x_i^t, v_i^t)\right) + bG(g\left(\overrightarrow{(x_i^t - x_i^{t-1})}, \overrightarrow{v_i^t}\right), \quad (1)$$

where f – function determiming the distance from a point to a vector

$x_i^t, v_i^t -$ current point and segment, $a, b -$ selectable invariables,

$F, G -$ normal distribution densities, $g -$ function of vector direction difference,

Threat the total probability at each examined path or the probability on basis of which the decision about the chosen path is made is calculated according to the formula (in accordance with the definition of random process):

$$\tilde{P}_i = \prod_{t=1}^{N} P_i^t, \quad (2)$$

$$P_i = {}^{1}/_{\sum_{i=1}^{M} \tilde{P}_i} \tilde{P}_i \quad , \quad (3)$$

The equation (3) is introduced for normalization of the obtained values in order to keep the property of normalization of random value.

It is necessary do define the parameters of the indicated above formula, namely *a*, *b*, as parameters of weighted sum of the two detached principles and parameters of random allocation of functions *F*, *G*. As far as the functions *f*, *g* show the deviation from the actual value mathematical expectation were considered equal to 0. Values of quadratic averages were picked based on the heuristic ideas specifically for the function *G* the quadratic average equals 0.6, as if the angle deviation more than 45 degrees it may be concluded that the selected direction is incorrect and the value of distribution density should be less than 0.1.

Judgments of that kind are applicable to the function of configuration error as well, it is accepted that the average corridor size of a large building is more than 2 meters, thus every two points the error for which is more than two meters should take the function value less than 0.1. Based on the above stated judgments the following distribution densities were chosen $F = N(0,2),\ G = N(0, 0.6)$. As the principle of configuration coincidence is more prominent in comparison with the principle of directions coincidence it was concluded that parameter a = 3, and parameter b = 1, that does not come into conflict with the probability definition $P_i^t,$ because in case of the selected parameters $0 < P_i^t < 1$.

Consequently the algorithm of prioritization sets the priority to each examined path thereafter a decision about the current path is made. In case of the path change the path is recalculated from the current point to the endpoint.

5. Conclusion

The paper introduces modification of three-phase algorithm that allows increasing the step detection accuracy in the task of navigation. It considers the method to improve user's positioning reliability by using information about a floor plan of a building. The executed experiments proved the efficiency of proposed algorithms.

6. References

[1] Kim J, Jang H, Hwang D and Park C A 2004 Step, Stride and Heading Determination for the Pedestrian Navigation System *Journal of Global Positioning Systems* **3(1-2)** 273-279 DOI: 10.5081/jgps.3.1.273

[2] Muset B and Emerich S 2012 Distance Measuring using Accelerometer and Gyroscope Sensors Carpathian *Journal of Electronic and Computer Engineering* **5(1)** 83-86

[3] Khan M I 2013 *Design and Development of Indoor Positioning System: For Portable Devices* (LAP LAMBERT Academic Publishing) p 108

[4] Lukianto C and Sternberg H 2011 *Overview of Current Indoor Navigation Techniques and Implementation Studies* (Access mode: http://www.fig.net/resources/proceedings/fig_proceedings fig2011/papers/ts09a/ts09a_lukianto_sternberg_5102.pdf)
[5] Lan K-C and Shih W-Y 2012 Estimating Step Distance Using Simple Harmonic Motion *IEEE 75th Vehicular Technology Conference (VTC Spring)* 1-5 DOI: 10.1109/vetecs.2012.6239903
[6] Vasin Yu G and Yasakov Yu V 2016 Distributed database management system for integrated processing of spatial data in a GIS *Computer Optics* **40(6)** 919-928 DOI: 10.18287/24126179-2016-40-6-919-928
[7] Vasin Yu G, Osipov M P, Muntyan S V and Kustov E A 2015 Procedural Modeling and Interactive 3D Visualization of Objects of the Internal Structure of Buildings and Facilities *Pattern Recognition and Image Analysis* **25(2)** 278-280 DOI: 10.1134/s105466181502025x
[8] Kaixu L, Gianmario M, Tianyi M and Tao G 2016 Multi-floor Indoor Navigation with Geomagnetic Field Positioning and Ant Colony Optimization Algorithm *IEEE Symposium on Service-Oriented System Engineering (SOSE)* 314-323 DOI: 10.1109/SOSE.2016.18
[9] Vasin Yu G, Osipov M P, Egorov A A and Yasakov Yu V 2015 Autonomous Indoor 3D Navigation *Pattern Recognition and Image Analysis* **25(3)** 373-377 DOI: 10.1134/s1054661815030256
[10] Aggarwal P, Thomas D, Ojeda L and Borenstein J 2011 Map matching and heuristic elimination of gyro drift for personal navigation systems in GPS-denied conditions *Measurement Science and Technology* **22** 12 DOI: 10.1088/0957-0233/22/2/02520
[11] Osipov M P and Patrushev A O 2016 Methods of location correction in the task of Autonomous navigation in indoor spaces *GraphiCon conference proceedings* 417-419

Acknowledgments
This research was supported by the Russian Foundation for Basic Research, project № 16-07-01214 A.

Distributed Infrastructure for Big Data Processing in the Transaction Monitoring Systems

M U Sapozhnikova[1], M M Gayanova[1], A M Vulfin[1,2], A V Nikonov[1] and A V Chuykov[1]

[1]Ufa State Aviation Technical University, K. Marks St. 12, Ufa, Russia, 450077

Abstract. To increase the effectiveness of detecting fraudulent bank transactions, the structure of the system is proposed to analyze data of user environment in order to identify potential fraudulent activities. The system for collecting and analyzing information about the user environment allows to accumulate data about the user environment, to mark precedents in manual and automatic modes and build a database of images for classifiers training. It is necessary to implement data collection, storage and access interface for the application of data mining tools. Operation of significant amount of accumulated data requires the use of special tools (frameworks and hardware platforms) for processing large data. In this paper the analysis of the existing software and hardware tools for distributed processing of indefinitely structured data of bank transactions (frameworks: Hadoop, Apache Spark) is presented. The structure and recommendations for the deployment of a hardware and software stand for testing algorithms for detecting financial fraud on the basis of data mining analysis as part of a distributed data processing system for bank transactions based on the selected framework are developed.

1. Introduction

The penetration of information technologies into all spheres of human life creates the basis for the formation of new conditions for the functioning of the market. In these conditions it became possible to develop the digital economy. The key factors of this economy are electronic technologies and services and digital representation of large volume multi-branch data [1, 2]. E-commerce is a significant institution in this branch of economy, penetrating into an increasing number of legal relationships that are emerging in the field of trade in electronic form. There is a rapid growth in the sphere of financial technologies: the introduction of artificial intelligence technologies, machine learning, analysis of large data to improve the efficiency of interaction of all participants in legal relations [3, 4, 5].

An important aspect of the functioning of the digital economy is the provision of information and economic security of business, personal data protection. As a result of the rapid development of financial technologies in the whole world, there has been an increase in fraudulent activities in the electronic environment. According to the Central Bank of Russia for the year 2014, the share of fraudulent transactions in Internet banking was 63%, and in the last 2 years – increased 5.5-fold and accounted for 93% of all crimes related to embezzlement of funds from cardholders' accounts [6, 7].
Nowadays, the application of big data processing technologies and data mining methods (DM) is an important element of the anti-fraud system. For example, introduction of Big Data by HSBC has

increased the efficiency of fraudulent incident detection by 10 times [8]. VISA anti-fraud system helps prevent fraudulent payments amounting to $ 2 billion annually [9].

2. Analysis of user environment data as a part of anti-fraud system

The transaction monitoring system (TMS) or anti-fraud system (AFS) is specialized software or hardware-software complex that monitors, detects fraudulent activities, and provides support for decision-making on the detected illegal operation.

The most promising solution for today is the use of technologies to define the user environment in combination with the methods of machine learning. The use of machine learning is a necessary metrics, since a large amount of information about the user environment is collected and the application of rules to this data becomes impossible. Classic methods for detecting fraudulent actions can not sufficiently accurately answer the question: was this action actually performed by the user? There are many ways to obtain illegal access to a user's account: phishing, vishing, pharming, mobile fraud, and other methods related to social engineering techniques [10, 11, 12].

To analyze the large volumes of data collected about the user environment, it is advisable to use approaches based on data mining. The application of the data mining algorithms for solving the problem of fraudulent transaction recognition and analysis of the user environment data is given in [11, 12, 13, 14]. The greatest efficiency is achieved using a combination of various algorithms (stacking-bagging) and the use of large data technologies (Hadoop Processing) [11]. This is explained by the fact that in pure form these algorithms are no longer capable of solving existing problems in view of the increasing volumes of processed data. There is a need to modify these algorithms, combine them to obtain an acceptable result, and also apply technologies capable of processing huge amounts of accumulated information.

The goal of this research is development of the infrastructure for collection and analysis of user environment as a part of anti-fraud system on the basis of big data processing technologies.

To achieve this goal, the following tasks were formulated:

- Development of the structure of the system for collecting and analyzing information about the user environment based on DM techniques.
- Development of a structural and functional scheme of processing user environment data for testing algorithms of detecting financial fraud on the basis of DM as a part of the system of distributed processing of banking transaction data.

3. System structure for collecting and analyzing user environment information

Technologies of remote banking service (internet banking) for accessing accounts and operations through a web browser do not require installation of the client part software and have become very widespread [15, 16]. The user makes certain manipulations in a web browser that interacts with the Frontend server interface. Frontend server generates a set of data about the user environment and transfers data about the user's actions about transaction initializing to the Backend server of remote banking system (RB) and then to the automated banking system (ABS) for the calculations [15, 16, 17]. Backend server transfers transaction data and collected data about the user's environment for analysis to the anti-fraud system. If the legitimacy of the transaction is recognized, data is transferred to the ABS server, otherwise the Backend server refuses user to perform the transaction.

Anti-fraud system evaluates risk of the current transaction and, in case of exceeding a certain threshold value, triggers additional mechanisms for verifying the legitimacy of the transaction [18]:

- automated ways of additional transaction authentication
  - SMS / push-notification;
  - request to answer the test questions;
- manual ways of additional transaction authentication
  - a phone call from a security specialist to the user.

In this architecture, user session management module (USMM) is the main element of the anti-fraud system. The module analyzes the transaction data and the user environment data (UED)

collected by the client-side script. The implemented script forms a set of data about the user environment [19]. The generalized structure of collected UEDs is the following:

- Color depth;
- Document size;
- Screen size;
- Time zone offset;
- Fonts;
- Plugins;
- IP-address;
- Number of processor cores;
- UserAgent.

The module's main task is to classify the current session and its transaction (legitimate transaction or fraudulent actions) based on the composition of the analysis methods: signature and automatic. If the module's estimated legitimacy of the transaction is below the threshold, an additional mechanism for authenticating the user is triggered.

The signature analysis module allows the use an expert knowledge and their formalization in the form of a system of production rules "IF-TO". The main task of this module is to classify UED and/or existing data about banking transaction in order to detect fraudulent activities. A special feature of the module is the use of a unified signature database based on the system of production rules, which allows to integrate mechanism for explaining the decision in an understandable form for security expert into the system. Initially, the database of the signature analysis module contains typical templates of fraudulent and legitimate transactions and UEDs in appropriate cases. Replenishment of the signature database is possible both in manual mode through the interface of the manual analysis module, and with the help of the automatic analysis modules which detects signatures in the new data processing mode. Analysis of the accumulated data under control of the analyst makes it possible to identify new production rules in an automatic mode based on the DM technologies.

Task of the automatic analysis module is automatic data classification based on the DM methods, for example, using a neural network classifier trained on a database of committed transactions marked by an expert [11]. At the stage of data preparation for DM algorithms application, the following tasks must be solved:

- features selection – selection of the most significant features for the adoption of a classification decision.
- features transformation – filling gaps in data, removing emissions and filtering noise components.
- features extraction – transformation of the selected features into a new feature space for feeding to the classifier input;

In the previous works [11, 12] the original feature space included 40 parameters. After evaluating available parameter values and its distributions, 12 parameters were deleted from the original data. As a result of expert analysis of the feature space transformation results, new nonlinear features were added – combinations of the initial features, characterizing possible combinations of some available parameters.

The core of automatic analysis module is the neural network – the "black box" – which allows to assign the current vector of features extracted from the data collected about the user environment to one of the previously defined classes. The following classes are proposed:

- User system is under remote control;
- The user system uses action anonymization mechanisms;
- User system does not contain any suspicious elements.

Results of the signature and automatic analysis modules are comparing. If the verdicts do not coincide, then an anti-fraud system expert can be brought in for manual analysis of the precedent. While the accumulation of data on transactions and UED occurs, the neural network classifier is after-trained. While group of use cases, the parameters of which do not fit into the current scheme of "IF-TO" signature rules, are being formed, new rules are being extracted, replenishing the existing

signature database. If new signatures are added by the expert, then the current base of marked use cases is analyzed in order to update the class labels and retrain the neural network classifier.

Manual analysis module is designed to correct the markup of the existing database of the user environment and transactions in order to form a training sample for the neural network classifier. The module allows expert to analyze the system solution for each of the use cases and correct it in case of erroneous operations.

The AFS control modules allow to monitor the basic performance of AFS, analyze the log of the system and debug the interaction of the signature and automatic modules.

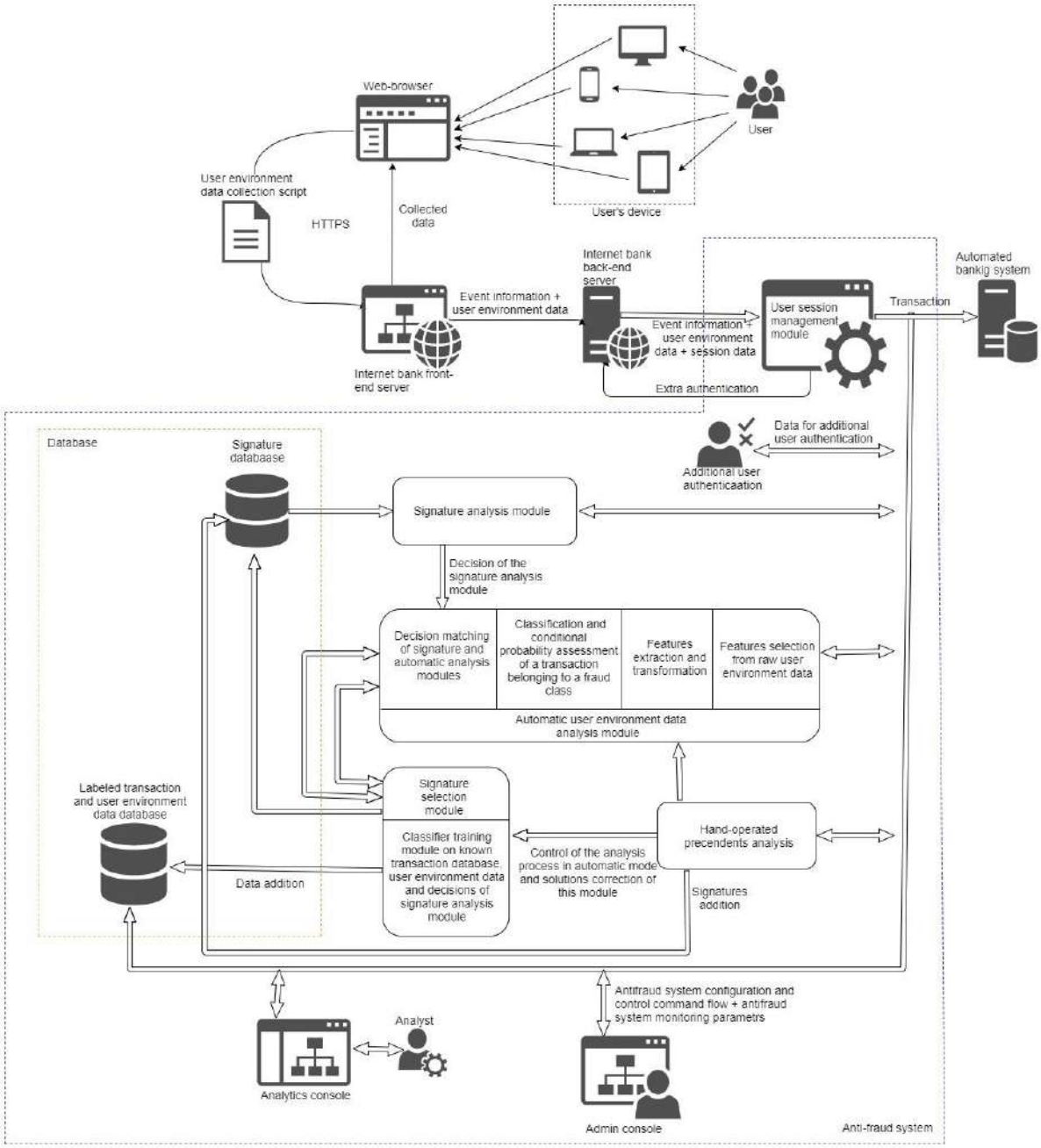


Figure 1. Structure of the user environment data and transaction analysis system within the AFS.

Thus, structure of the system for collection and analysis of user environment as a part of anti-fraud system is presented in figure 1. The key element of the system is data mining module. Analysis algorithms should be applicable in the context of "Big Data" (a set of approaches, tools and methods for processing structured and unstructured data of huge volumes and significant diversity for obtaining

human-readable results that are effective under conditions of continuous growth, distribution over multiple nodes of the computer network) [20].

4. Designing a structural and functional scheme for processing "big data" of the user environment as a part of a distributed data processing system for bank transactions

The implementation of algorithms for detecting financial fraud based on data mining techniques of banking transactions as part of a distributed processing system of banking transactions requires the solution of a few tasks related to the design and deployment of an appropriate infrastructure for storing and processing accumulated data.

To date, there are many tools for the distributed processing of banking transaction data (frameworks: Hadoop, Apache Spark, ClickHouse, ElasticSearch, Splunk Free) [21, 22, 23, 24, 25]. The proposed structure of the distributed processing system of banking transaction data is presented in the figure 2.

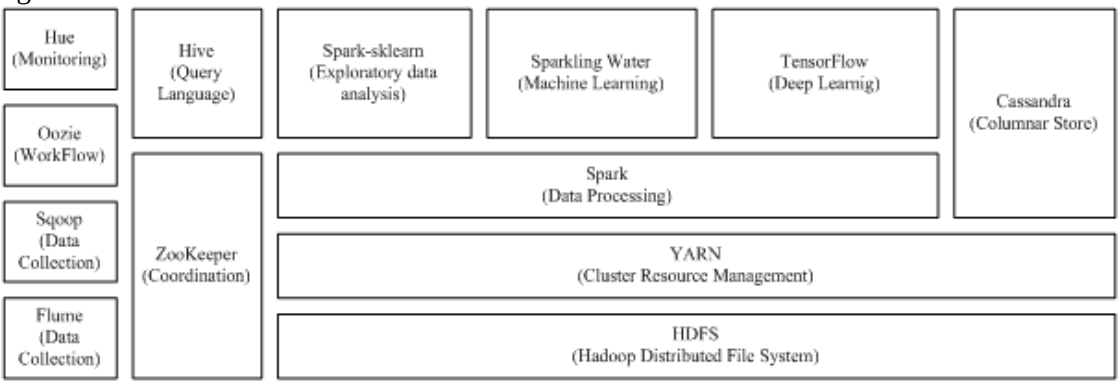


Figure 2. Hadoop cluster structure

Main element of the distributed processing system of banking transaction data is a distributed file system. Nowadays the most popular distributed file system is HDFS [21].

Next element of the big data processing system is the distributed programming and machine learning infrastructure. The core of this element is Spark, a cluster computing infrastructure similar in MapReduce [21]. The structure of this infrastructure includes the machine learning tool MLLib, which allows to implement the DM tools of the accumulated data.

To directly store the accumulated data, it is proposed to use solutions from the New-SQL family [26].

A detailed description of the additional elements and their functions in the distributed data processing system for bank transactions is presented in the table 1.

Table 1. Hadoop ecosystem applications

Function	Tools name	Short description
Store	HDFS[19]	Distributed file system
	Cassandra [24]	NoSql database management system
Cluster resource management	YARN [19]	Operating system for big data application
Data processing	Spark [20]	Engine for big data processing
Machine learning	Spark-sklearn [25]	Scikit-learn python library integrated in Apache Spark for exploratory data analysis
	Sparkling Water [26]	H2O library integrated in Apache Spark for machine learning in Hadoop system
	TensorFlow [27]	TensorFlow library integrated in Apache Spark for deep learning in Hadoop system
Coordination	Zookeeper [28]	Application for maintaining configuration information, naming and etc
Data Access	Hive [29]	Application for data summarization, SQL-like query, and analysis
Data Collection	Sqoop [30]	Application for transferring data between relational

		databases and Hadoop
	Flume [31]	Application for transferring data between relational databases and Hadoop
WorkFlow	Oozie [32]	Application for collecting, aggregating, and moving of unstructured data
Monitoring	Hue [33]	Web interface to monitor Hadoop system

The use of three machine learning libraries is due to the need of rapid prototyping of the developed algorithms, debugging on small amounts of data and the possibility of importing the developed models.

The hardware-software stand structure for testing the algorithms of detecting financial fraud based on DM techniques in the distributed data processing system for bank transactions based on the selected big data processing stack is shown on figure 3.

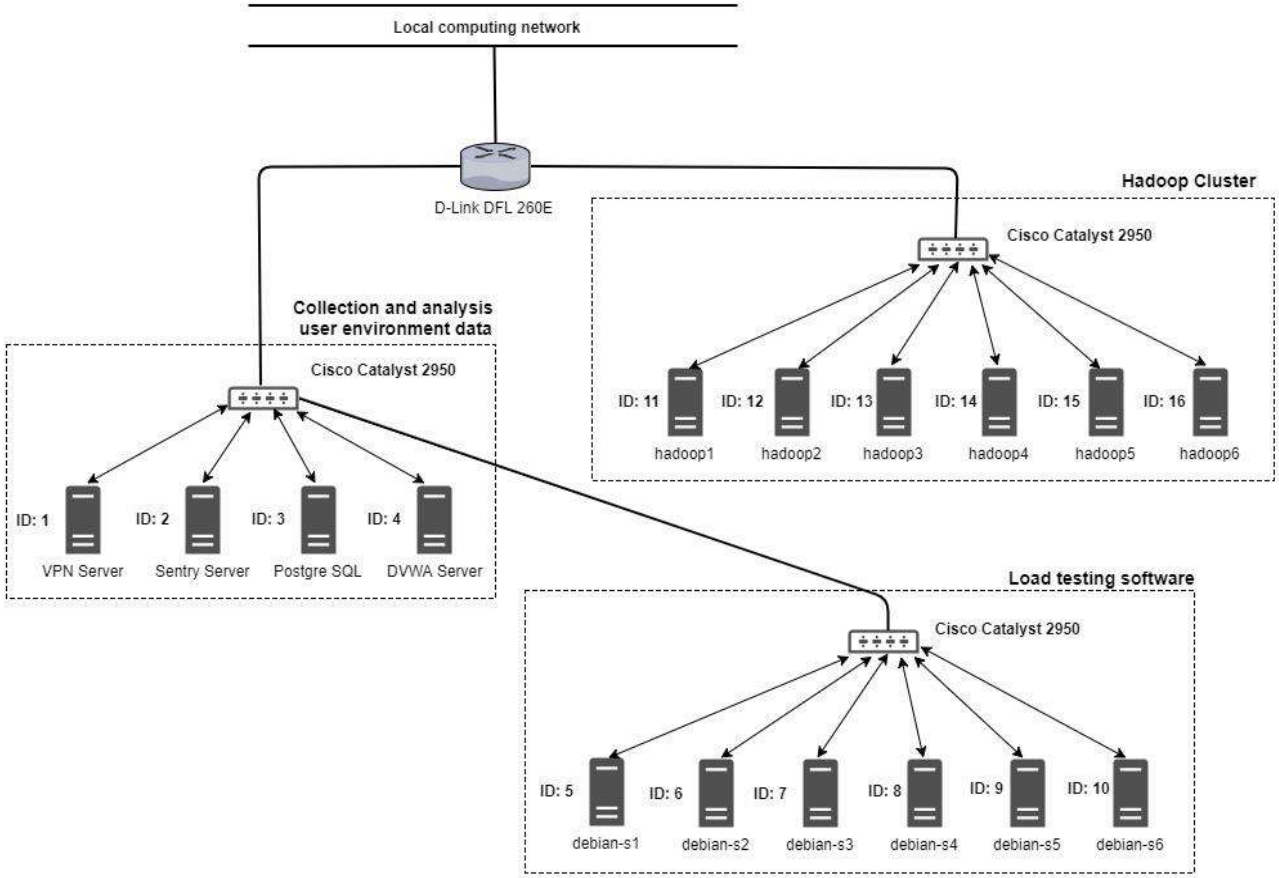


Figure 3. The hardware-software stand structure testing algorithms for detecting financial fraud.

The data collection and analysis module consists of:

- Sentry software suite [36] for collection client-side script logs;
- Gitlab service [37] for organization joint work on the implemented analysis algorithms source code;
- DVWA (Damn Vulnerable Web Application) for testing the script of user environment data collection.

The load testing module is designed to automate collection of the user environment database.

For distributed processing of bank transaction data, the Hadoop cluster is used, on which the software is deployed from Table 1. Typical configuration of the used server machine pool is shown in Table 2.

Table 2. Server settings.

ID	Sector	Configuration	OS
1	Collection and analysis user environment data	2x3.4 GHz / 4GB	debian 8.2
2		2x3.4 GHz / 2GB	debian 8.2
3		2x3.0 GHz / 1GB D	ubuntu 16.04
4		1x2.4 GHz / 1GB	ubuntu 16.04
5	Load testing software	2x3.0 GHz / 3GB	debian 9.2
6		2x3.0 GHz / 3GB	debian 9.2
7		2x3.0 GHz / 4GB	debian 9.2
8		2x3.0 GHz / 3GB	debian 9.2
9		2x3.0 GHz / 4GB	debian 9.2
10		2x3.0 GHz / 3GB	debian 9.2
11	Hadoop Cluster	2x3.4 GHz / 12GB	ubuntu 14.04
12		2x3.4 GHz / 12GB	ubuntu 14.04
13		2x3.4 GHz / 8GB	ubuntu 14.04
14		2x3.4 GHz / 6GB	ubuntu 14.04
15		2x3.2 GHz / 6GB	ubuntu 14.04
16		2x3.0 GHz / 6GB	ubuntu 14.04

5. Conclusion

The main problem of improving the TMS efficiency is the insufficient amount of registered parameters transferred from an online banking client side to a processing center, and the imperfection of signature analysis methods and algorithms because of low adaptability and configuration flexibility.

Nowadays the most promising solution is the use of technologies for determining user environment in combination with the methods of machine learning within TMS. The use of machine learning is an indispensable criterion, because big amount of user environment information being collected and application of the rules to this data becomes difficult. Algorithms for analysis should be applicable in the context of “big data”.

In the paper the infrastructure for collection and analysis of user environment as a part of anti-fraud system on the basis of big data processing technologies is proposed.

6. References

[1] Teoh C S and Mahmood A K 2017 National cyber security strategies for digital economy *Int. Conf. on Research and Innovation in Information Systems* 1-6

[2] Crabtree A 2016 Enabling the new economic actor: Personal data regulation and the digital economy *Proc. IEEE Int. Conf. on Cloud Engineering Workshops* 124-129

[3] Tung H H, Cheng C C, Chen Y Y and Chen Y F 2016 Binary Classification and Data Analysisfor Modeling Calendar Anomalies in Financial Markets *7th Int. Conf. on Cloud Computing and Big Data* 116-121

[4] Trelewicz J Q 2017 Big Data and Big Money: The Role of Data in the Financial Sector *IT Prof.* **19(3)** 8-10

[5] Luvizan S S, Nascimento P T and Yu A 2017 Big Data for innovation: The case of credit evaluation using mobile data analyzed by innovation ecosystem lens *Proc. Portland Int. Conf. on Management of Engineering and Technology: Technology Management For Social Innovation* 925-936

[6] Sapozhnikova M U, Gayanova M M, Nikonov A V and Vulfin A M 2017 Data mining algorithms of bank transaction as a part of antifraud system *Proc. Conf. Information Technologies for Intelligent Decision Making Support* **2** 143-149

[7] Lopez-Rojas E A and Axelsson S 2016 A review of computer simulation for fraud detection research in financial datasets *Proc. of Future Technologies Conf.* 932-935

[8] Dodds L S 2014 Big ideas are coming from using big data (Access mode: https://www.raconteur.net/technology/big-ideas-are-coming-from-using-big-data)

[9] Rosenbush S 2014 Visa Says Big Data Identifies Billions of Dollars in Fraud *CIO J.* (Access mode: https://blogs.wsj.com/cio/2013/03/11/visa-says-big-data-identifies-billions-of-dollars-in-fraud)
[10] Piskunov I 2017 Anti-fraud systems and how it works *J. Securitylab* (Access mode: https://www.securitylab.ru/blog/personal/Informacionnaya_bezopasnost_v_detalyah/ 339929)
[11] Sapozhnikova M U, Gayanova M M, Nikonov A V and Vulfin A M 2017 Data mining technologies in the problem of designing the bank transaction monitoring system *ComputerScience and Information Technologies* 45-56
[12] Sapozhnikova M U, Gayanova M M, Nikonov A V, Vulfin A M and Kurrenov D V 2017 Anti-fraud system on the basis of data mining technologies Int. *Symp. on Signal Processing and Information Technology* 1-5
[13] Vizilter, Yu V, Gorbatsevich V S, Vorotnikov A V and Kostromov N A 2017 Real-time face identification via CNN and boosted hashing forest *Computer Optics* **41(2)** 254-265 DOI: 10.18287/2412-6179-2017-41-2-254-265
[14] Ivanov A I, Lozhnikov P S and Sulavko A E 2017 Evaluation of signature verification reliability based on artificial neural networks, Bayesian multivariate functional and quadratic forms *Computer Optics* **41(5)** 765-774 DOI: 10.18287/2412-6179-2017-41-5-765-774
[15] Abbad M, Abed J M and Abbad M 2012 The Development of E-Banking in Developing Countries in the Middle East. *J. Financ. Account. Mana*g. **3** 107-123
[16] Jarrett J E 2016 Internet Banking Development *J. Entrep. Organ. Manag*. **5** 2-5
[17] *Global Mass Payments, AP Software, B2B Payments* (Access mode: https://tipalti.com/)
[18] Fedotenko M 2017 How banks are protected: explaining the structure and the principles of a bank antifraud system *J. Hacker* (Access mode: https://xakep.ru/2017/04/21/antifrod-1/)
[19] Cao Y, Li S and Wijmans E 2017 (Cross-) Browser Fingerprinting via OS and Hardware Level Features *Proc. Network and Distributed System Security Symp*. 1-15
[20] *Big Data* (Access mode: https://en.wikipedia.org/wiki/Big_data/)
[21] *Apache Hadoop* (Access mode: http://hadoop.apache.org/)
[22] *Apache Spark* (Access mode: https://spark.apache.org/)
[23] *Click House* (Access mode: https://clickhouse.yandex/)
[24] *Elastic search* (Access mode: https://www.elastic.co/products/elasticsearch/)
[25] *Splunk* (Access mode: https://www.splunk.com/)
[26] *Apache Cassandra* (Access mode: http://cassandra.apache.org/)
[27] *Spark-sklearn* (Access mode: https://github.com/databricks/spark-sklearn/)
[28] *Sparkling Water* (Access mode: https://www.h2o.ai/sparkling-water/)
[29] *Tensor Flow On Spark* (Access mode: github.com/yahoo/TensorFlowOnSpark/)
[30] *Apache Zookeeper* (Access mode: https://zookeeper.apache.org/)
[31] *Apache Hive* (Access mode: https://hive.apache.org/)
[32] *Apache Sqoop* (Access mode: http://sqoop.apache.org/)
[33] *Apache Flume* (Access mode: https://flume.apache.org/)
[34] *Apache Oozie* (Access mode: http://oozie.apache.org/)
[35] *Cloudera Hue* (Access mode: https://github.com/cloudera/hue/)
[36] *Sentry* (Access mode: https://sentry.io/welcome/)
[37] *GitLab* (Access mode: https://about.gitlab.com/)

Acknowledgments
Work was supported by the Russian Foundation for Basic Research, research № 17–48–020095.

Neuro-Fuzzy Model for Arrhythmia Diagnostic System

A V Nikonov[1], A M Vulfin[1], M M Gayanova[1] and M U Sapozhnikova[1]

[1]Ufa State Aviation Technical University, K.Marks St. 12, Ufa, Russia, 450077

Abstract. Cardiovascular diseases are one of the leading causes of death worldwide. People suffering from or at high risk of such diseases need constant supervision, early diagnosis and timely assistance. It is shown that the achievement of high accuracy performance in real-time arrhythmias recognition is associated with significant hardware costs. Detection accuracy of arrhythmias recognition does not exceed 80%. An approach, which is offered to solve the problem of high-precision arrhythmia diagnosis on the basis of electrocardiosignal is based on the data mining methods. Application of such methods is necessary for processing large amounts of data with complex structure of the features. Determination of the arrhythmia type with the use of fuzzy inference tools needs to specify the technique of the original data preprocessing. Feature selection, formalization and coding is considered in this paper.
The issue of the knowledge base construction – coding, generation and selection of the features (database) as well as the construction of the rules base ¬ as the part of the neuro-fuzzy diagnostic system is also considered. The research goal is to improve the intelligent systems of arrhythmia diagnostics on the basis of neural network classifiers by developing a solution explanation subsystem based on neuro-fuzzy models.

1. Introduction

Cardiac arrhythmia as the most common disease is an irregular heart rhythm caused by improper operation of electrical impulses that regulate the heartbeat [1,2].

Arrhythmia can be detected by cardiac specialist using electrocardiogram (ECG) strip. Even for skillful cardiologist it may take several minutes to make a diagnosis; in some severe cases, this could be fatal for the patient. ECG analysis is too hard for the beginner or inexperience staff [3,4].

The electrocardiosignal (ECS) is an electrical manifestation of the heart contractile activity and it characterizes the state of the cardiovascular system (CVS). Electrocardiosignal on the ECG looks like a set of alternating waves, flat intervals and segments located on the baseline. Each element contains information of the state of the heart and its components [2]. Parameters of the cardiac contraction pattern and individual features of the ECS are considered in [36].

Intelligent analysis of ECS data using neural network nonlinear models shows the highest results of recognition accuracy on model tasks [7,14-18]. Nevertheless, the distinctive feature of intelligent diagnostic systems based on the neural network technologies is the complexity of explaining the decision made. The reason is that signal processing algorithms and intermediate calculation results of neural network models are uninterpreted by both the medical technique and the knowledge engineer [2,19].

Among the many models of expert knowledge representation, the production rules are widely accepted nowadays due to their flexibility and simplicity of describing the laws of the subject domain. With the help of products it is easier to model the thinking activity of an expert, his reasoning, conclusions, decision-making process [37].

The problem of diagnosing cardiac arrhythmias have to be solved in conditions of indistinctness, incompleteness and unreliability of the initial data. It is impossible to solve such problem using ordinary production rules. It is necessary to use fuzzy products and systems of fuzzy logic inference for this task.

Fuzzy neural networks are hybrid models that combine the main advantages of neural networks (possibility of adaptive learning) and fuzzy systems (interpretability of the result obtained with their help). The fuzzy neural network is analogous to the system of fuzzy logic inference. The system of fuzzy logic inference is able to make decisions only on the basis of a ready (expert-generated) knowledge base, while a fuzzy neural network as a result of training itself forms a system of rules.

Therefore, it is relevant to choose a flexible fuzzy-productive model of knowledge representation in the problem of arrhythmia diagnosis. Neuro-fuzzy system based on a fuzzy neural network allows, as a result of training, to form a knowledge base of the expert system like a set of rules in the form of fuzzy products.

The research goal is to improve the intelligent systems of arrhythmia diagnostics on the basis of neural network classifiers by developing a solution explanation subsystem based on neuro-fuzzy models.

To achieve the goal, the following tasks were set:

- Development of the structure of an intelligent system for diagnosing arrhythmias with a subsystem of explaining the solution based on neuro-fuzzy models;
- Construction of the knowledge base as well as construction of the rules base;
- Development of the solution explanation subsystem model.

2. Analysis of automatic arrhythmia diagnosis methods

Nowadays there are several methods currently used in automatic arrhythmia diagnosis systems [2-5,9,17,23,25,30]. General groups of methods are usually distinguished as linear, nonlinear and methods based on data mining (DM) [2, 10-13].

The introduction and widening application of automatic external defibrillators (AEDs) make strong demands for ECS analysis. It is required highly accurate discrimination between shockable and non-shockable rhythms, with sensitivity and specificity aimed to achieve the maximum values [23]. Also, a false positive detection will initiate a defibrillator to give improper therapeutic intervention [24,25].

Modern diagnostic systems are a compromise between algorithms that require significant computational costs to achieve high accuracy of diagnostic, and hardware with limited performance. It has been established that the accuracy of arrhythmia recognition in existing solutions does not often exceed 80% [2,5,10-13,17,20,25,38].

Other problems that common systems of arrhythmia diagnosis meet are the following:

1) The problem of analyzing large amounts of ECG data

As some arrhythmias appear irregularly, it is necessary recording ECG activity using, e.g., Holter monitor up to a week to successfully capture them [3]. The total number of such cardiac abnormalities reach up to 96 different categories [7]. Each category may contain about 28,800 beats, if 48 hours of single-channel ambulatory recording is considered, assuming an average heart-rate of about 60 beats-per-minute (BPM). The size of the database can be further increased if data is accumulated from multiple channels (up to 12) [7]. Considering the large computational costs of algorithms for processing ECS, processing large amounts of data is tending to take a long time.

2) The problem of generalization possibility

Existing approaches tend to suffer from high sensitivity to noise included in the ECG, and unreliability in dealing with new or ambiguous patterns from other patients. Classifiers of CAD systems must be developed that are able to deal with nonlinear discrimination between classes, incomplete or ambiguous input patterns, and suppression of false alarms [8].

The main difficulty in the task of automatic ECG analysis lies in the large scatter of the morphology of the ECG curves. Different patients have significantly changing parameters of the sequence of heartbeats (due to movements that produce high-frequency noises, or breathing, causing a

baseline walk, or differences in the electrical characteristics of the body) [20]. Creation of a data set that would cover all possible ECG morphologies collected from different patients with various abnormalities in the functioning of the CVS is impossible. As a result, diagnostic systems based on classifiers of different types and trained on some collected data, show unsatisfactory diagnostic results of the ECG analysis of different patients [6,21,22] under real conditions of use.

Thus, for the task of automatic arrhythmia diagnosis it is proposed to use data mining (DM) methods. The DM technologies are designed to search in large amounts of data of non-obvious, objective and useful regularities that reflect the objective internal data structure in comparison with the subjective opinion of the expert. The main purposes of the application of intelligent methods are:

- analysis of signals characterized by a high degree of uncertainty, e.g., "non-stochastic" type, which includes most biomedical signals, including ECS;
- increasing the level of intelligent assistance of medical specialists;
- revealing hidden regularities and extracting new knowledge from the accumulated data, which will allow to build production systems of explaining the diagnostic solutions.

To apply modern diagnostic methods to real ECS data, it is necessary to adapt existing methods of data mining, which allow to study complex nonlinear ECS more fully. CRISP-DM [27] (The Cross Industries Standard Process for Data Mining) is the most popular and widely used methodology. In accordance to the CRISP standard, DM is a continuous process with many cycles and feedbacks.

3. Development of the structure of an intelligent system for arrhythmia diagnosis

To analyze complex nonlinear ECS signals it is proposed to use the diagnosis methods based on DM methods and algorithms. The structure of the proposed intelligent diagnostic system combining neural network models and fuzzy logic models is shown in figure 1.

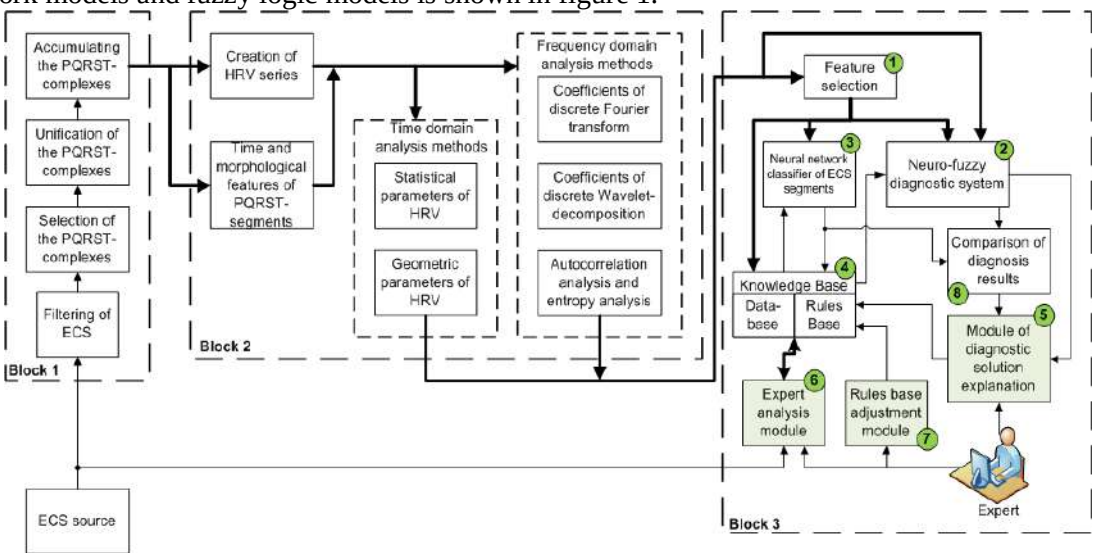


Figure 1. Structure of the neuro-fuzzy diagnostic system.

The system is divided into three blocks. The block 1 performs preprocessing of the raw ECS signal such as filtering the original signal and selecting the main segments describing the stages of cardiac contraction [13,18] – PQRST complex.

The procedure of generating and pre-selecting features is implemented in block 2. Time-stamps – durations of individual segments of the PQRST complex – are extracted from the signal and a time series of heart rate variability (HRV) is constructed, the analysis of which is widely used for the diagnosis of arrhythmias [32]. Feature generation methods perform signal processing both in the time and frequency domains. Statistical methods of analysis compute the basic statistical parameters of a HRV series [32], while geometric methods based on the analysis of the form of distribution function of the HRV series [33].

Since changes in the heart rate are difficult to analyze in the time domain [32], spectral analysis methods [2, 34-36] are used to investigate the frequency properties of HRV and the marked initial ECS such as discrete Fourier transform, wavelet-decomposition of the signal, and also autocorrelation analysis [33].

Most of arrhythmias can be analyzed due to the differences in the QRS part of the signal. The parameters of QRS complex can help to determine the differences in three of the five above-mentioned arrhythmias [4]. The most important function of the diagnostic system is the operative recognition of arrhythmias that directly threaten the patient's life.

It is proposed to consider the following types of arrhythmias and their characteristic features:

1. Asystolia and expressed bradycardia. Lack of or too low repetition rate QRS-complexes (less than 35 min-1). Effects cardiac arrest. For reliable recognition enough information about the durations of RR-intervals. RR-intervals are measured by QRS-complex detection algorithm.

2. Ventricular fibrillation. It begins with ventricular flutter (VF). Wave form similar to sine wave with frequency of up to 200 - 300 min^{-1}; QRS and T waves are merged into a single wave of large amplitude without isoelectric intervals; Continuous wave form and amplitude varying with frequency of 400 to 600 min^{-1} (ventricular fibrillation wave with small amplitude); small number of chaotic waves (150 - 300 min^{-1}) of greater amplitude (ventricular fibrillation wave with large amplitude). There is no visible P wave. QRS complex and T wave are merged in regularly occurring undulatory waves with a frequency between 180 and 250 beats per minute. Effects termination of blood circulation and cardiac arrest. It requires analysis of wave structure of the ECS, featuring a variety of forms submitted from regular (sine) or irregular (variable amplitude and frequency) waves to chaotic oscillations expressed.

3. Ventricular tachycardia. Enhanced the QRS-complexes (longer than 120 ms), having the same or spindle-shaped, which is followed by a frequency of 120 - 220 min^{-1}. Effects ventricular flutter, ventricular fibrillation. It may have three or more beats in a row that originate from the ventricle. A way that VT can be diagnosed is noticing that the P wave in the PQRST wave is no longer visible because the T wave from the previous beat has overtaken.

Based on this description, it is proposed to use the following ECS characteristics as the initial data for diagnostics:

1) Statistical:
   - HR – heart rate;
   - P-wave – moment of atrial contraction;
   - PQ-segment – isoelectric interval, important for the recognition of the baseline;
   - QRS-Complex – moment of ventricles contraction;
   - ST-interval – isoelectric interval;
   - T-wave – moment of the ventricles relaxation;
   - RR-interval – defines HRV;
   - PNN50 (%) – the proportion of adjacent R-R intervals that differ by more than 50 ms;
   - RMSSD – the square root of the mean square of the difference in the values of adjacent R-R intervals. It is a measure of HRV with a short duration of cycles.
   - SDANN –standard deviation from the mean of the R-R intervals computed for each 5-minute ECG recording section. Characterizes HRV with a long cycle duration.
2) Geometric (in time-domain):
   - Mode (Mo) – the most common value of the cardio interval in this dynamic series. Under normal distribution and high stationarity of the process under study, Mo is slightly different from the mathematical expectation;
   - Amplitude of the mode – number of cardiointervals corresponding to the value of the mode;
   - Variational range – reflects the degree of variability in the cardiointervals values in dynamic series. It is calculated from the difference between the maximum (Mx) and the

minimum (Mn) values of the cardiointervals and, therefore, may be distorted in arrhythmias or artifacts.

3) Geometric (in frequency-domain):

The frequency composition of the heart rhythm is usually represented in the form of a graph of power distribution in frequency, according to which one can judge the expression of frequency components in different ranges:

- Total power in the range 0.15–0.40 Hz (period 6.5s–2.5s);
- Total power in the range 0.04–0.15 Hz (period 25s–6.5s);
- Total power in the range less than 0.04 Hz (period more than 25s);
- Total power in the range from 0.003 Hz to 0.40 Hz;
- Power Ratio.

4) Autocorrelation:

The calculation and construction of the autocorrelation function of the dynamic series of RR-intervals is aimed at studying the internal structure of this series as a random process. The autocorrelation function is a graph of the dynamics of the changes in the correlation coefficients obtained by sequentially shifting the analyzed dynamic series by one number relative to its own series:

- C1 – value of the correlation coefficient after the first shift;
- C0 – number of shifts, as a result of which the correlation coefficient becomes negative.

Block 3 performs the diagnosis of arrhythmia. It consists of two complementary systems: a neural network classifier (NNC) and a neuro-fuzzy diagnostic solution (NFS). The neural network classifier was described in [10,13,32,36] and showed the results of a diagnostic accuracy of about 86% on the field data.

Features of marked signal go to modules 1 and 2. Module 1 implements feature selection procedure for NNC. Based on a pair of Hebb filters, this module isolates the most significant features of the time and frequency domains [10,13]. Then the set of selected features and the marked PQRST-segments are fed to the input of NNC (module 3) and NFS (module 2), and are also recorded in the database (module 4) [2,10,11,13,19].

The result of the comparison of the diagnostic solutions of the NNC and NFS, as well as the features identified by the NFS, are verified by the expert in module 5. The expert has access to the analysis of the original signal and the selected features through the interface module of expert analysis (module 6), as well as access to the rule base on which the NFS module operates via the interface module of the rule base adjustment (module 7). Analyzing the data at all stages of processing, the expert can make changes by manually marking PQRST segments, changing rules in the rule base or determining the type of arrhythmia if necessary.

The processed record is stored in the knowledge base. In the subsequent processing, this data will be used for the further NNC training.

4. Development of the solution explanation subsystem model

It is proposed to construct an explanation subsystem on the basis of neuro-fuzzy models. The advantage of such models is that it is a neural network implementation of fuzzy inference. On the one hand, parameters of membership functions act as weights and can be adjusted via training the network. On the other hand, each layer is a phase of fuzzy inference that makes transparent the functioning of network – it becomes easy to correlate outputs of certain layer with the corresponding phase of fuzzy inference.

Nowadays, such structures as the Takagi-Sugeno-Kang model (TSK) or the Mamdani model are distinguished as classifiers based on a fuzzy inference with the possibility of adjusting the parameters of membership functions [12,15,29,31-33, 39]. However, these models in the classical version use a full rules base, which limits their application for problems with a large number of input variables and terms. So, it is proposed to use a structure based on the Mamdani model to classify the arrhythmia types based on the fuzzy inference, adapting the parameters of the membership functions of the terms of input and output linguistic variables and constructing a base of fuzzy rules as a result of training

based on self-organization. Structure of the neuron used in the neural network interpretation of the fuzzy inference model is shown in figure 2, structure of the neuro-fuzzy model is shown in figure 3. Notations of layers of the network are given in Table 1 for the description.

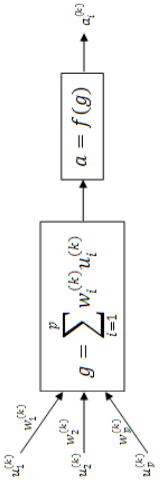


Figure 2. Neuron structure.

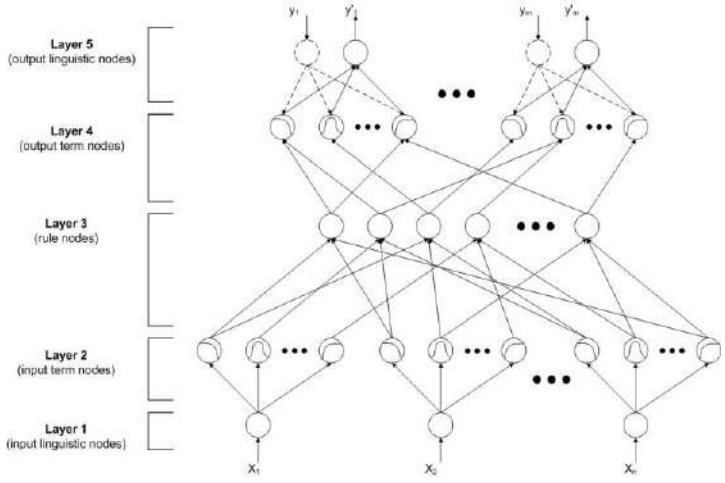


Figure 3. Structure of the neuro-fuzzy module.

Table 1. Notations of layers of the network.

Notation	Description
k	– number of the layer
p	– amount of node's input links
$u_i^{(k)}$	– i-th input signal of k-th layer
$w_i^{(k)}$	– i-th link weigth k-th layer
g	– integration function
a	– neuron activation function
n	– amount of input variables
m	– amount of output variables
σ_{ij}, m_{ij}	– width and center of the Gauss act. function of j-th term of i-th linguistic variable
x_i	– input variable
y_i	– training sample
y'_i	– output variable
h_i	– amount of terms of i-th linguistic variable

From the point of view of neural network architecture, rules base is a set of links and neurons of layer 3 in figure 3. Links are antecedents and neurons outputs are consequents of the rules. While constructing the network, each neuron of layer 2 connects to each neuron of layer 3 – this means building a full rules base. During the network training, most of the rules will be united or excluded.

Model of neural network functioning is the following. Nodes of first layer transmit input values directly to the next layer:

$$u_i^{(1)} = x_i, g = u_i^{(1)}, a = g, w_i^{(1)} = 1, i = \overline{1,n}. \quad (1)$$

Second layer performs a fuzzification operation. Activation function of the neurons of this layer is a membership function of the corresponding term of the linguistic variable. In this case, the Gaussian function is selected as the activation function:

$$g = \mathrm{M}_{x_i}^{j}(m_{ij}, \sigma_{ij})^2 = -\frac{(u_i^{(2)} - m_{ij})^2}{\sigma_{ij}^2},\ a = \mathrm{e}^g,\ j = \overline{1,h_i}. \quad (2)$$

Number of neurons of the third layer is equal to the number of fuzzy rules, defined as

$$\text{numRules} = \prod_{i=1}^{n} h_i. \quad (3)$$

At the outputs of this layer, rules degree of activity is formed, defined as the minimum of the degrees of belonging calculated on the previous layer – the operation of the fuzzy "AND":

$$g = \min_{i=1,\overline{p}} u_i^{(3)}, a = g, w_i^{(3)} = 1. \quad (4)$$

Neurons of the fourth layer operate in two modes: when transmitting a signal from the bottom to the top (operating mode) and vise versa (training mode). In the operating mode, neurons implement the composition of fuzzy subsets assigned to each output variable, using the sum of input signals with saturation – a fuzzy "OR" operation [28,33]:

$$g = \sum_{i=1}^{p} u_i^{(4)}, a = \min(1,\ g), w_i^{(4)} = 1. \quad (5)$$

In the training mode, the neurons of this layer operate similarly to the neurons of the second layer.

As shown in the figure 3, the fifth layer consists of two types of neurons. In the training mode, a training sample is fed to the input of the first type of neurons on the basis of which the parameters of the activation function of neurons of the fourth layer are adjusted. Thus, neurons of this type (marked with dashed lines) operate similarly to the neurons of the first layer:

$$g = y_i,\ a = g. \quad (6)$$

In the operating mode, second type of neurons operates (marked by a solid line). Neurons of this type perform the defuzzification operation using centroid method:

$$g = \sum w_{ij}^{(5)} u_i^{(5)} = \sum (m_{ij}\sigma_{ij}) u_i^{(5)},\ a = \frac{f}{\sum \sigma_{ij} u_i^{(5)}}. \quad (7)$$

Learning algorithm [28,33] is divided into two phases:

- Training on the basis of self-organization to initialize the membership functions of terms of input and output linguistic variables and to construct fuzzy rule base;
- Supervised learning to adjust the parameters of the created membership functions and minimize the network error.

At the first stage of phase 1expert manually determines the structure of the fuzzy system such as the number of terms of each input and output linguistic variable and initiates parameters of membership functions.

Second stage is responsible for primary selection of membership functions and their parameters: center and width. It is proposed to use the method of competitive training [28,33]. It involves independent iterative training of each neuron that implements the membership functions to the corresponding terms of linguistic variables. This training involves finding the minimum distance from the center of the membership function to the current input sample from the training sample. To initialize the value of the width of membership functions, first nearest neighbor heuristic can be used, since in the second phase of learning the optimal values of the parameters will be determined.

At the third stage, fuzzy rules are forming. First, a complete rules base is created, which is a combination of all conditions with all conclusions. Then, the competitive learning algorithm is used to determine the required number of rules from the full database. Here the fourth layer neurons operate in the training mode and receive the training sample signals from the fifth layer. Denoting the output of the j-th neuron of the fourth layer as $o_j^{(4)}(t)$, and the production part of the i-th rule (the weighted sum of the signals entering the neurons of the 4th layer) as $o_i^{(3)}(t)$, calculate the links weights w_{ij} as:

$$w_{ij}(t+1) = o_i^{(4)}(t) \cdot \left[w_{ij}(t) - o_i^{(3)}(t) \right]. \quad (8)$$

At the fourth stage, some rules are deleted. If there is more than one output connection for the neuron of the third layer, the connection with the largest weight is selected, and the remaining ones are deleted. It is also possible to delete all links and the neuron itself if the weights of all the links are less than a given threshold $\theta = 0.4$.

At the fifth stage, rules are combined to reduce their number. In figure 3, the preconditional part of the rules is displayed by neurons of the third layer, and part of the consequent – by the connections of the fourth layer.

The sixth stage involves building a neural network with a complete structure and a set of elements and links.

At phase 2, optimization of activation function parameters values is done by supervised learning of the network. For example, a back propagation algorithm for can be used for training [19]. Thus, the reverse pass, which calculates the error for the parameters of the center and the width of the Gaussian activation functions of layers 5 and 2, is determined similarly as in [28,33].

5. Experiment

Figure 4 shows a generalized scheme of the diagnostic model of signal preprocessing, detection and identification the events on the basis of ECS analysis with the indicating of algorithms proposed for the implementation of each stage.

Thus, after filtering and detection the PQRST-complex, wavelet decomposition is made on appropriate section of the original data vector and vector $\left[sa_j, sd_j, sa_{j-1}, sd_{j-1}, \ldots, sd_1\right]$ is forming, where *sa*, *sd* are the approximation and detailing coefficients of j, j-1, …,1 levels respectively. Then, the 8 major components are allocated from resulting vector [10]. As shown in [32], a given number of main components is sufficient enough to save 95% of the variance of the original signal.

In experiments [10, 32, 36], the resulting vector was fed to the input of a multilayer perceptron. The accuracy of arrhythmia classification was 86%. In this research, the vector is fed to the input of the fuzzy model committee described above. Each model determines the probability evaluation of having an appropriate arrhythmia class, as shown in figure 5.

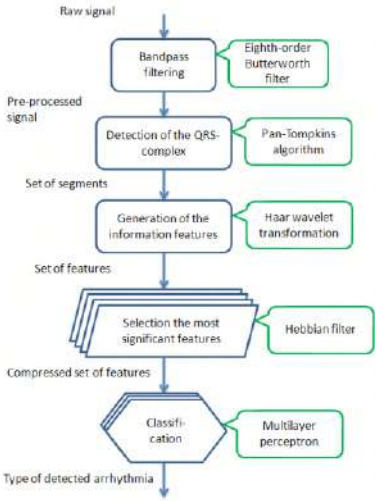


Figure 4. Generalized scheme of the diagnostic model of signal preprocessing.

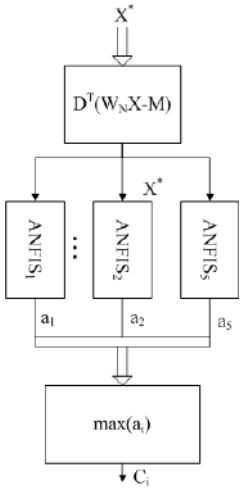


Figure 5. Fuzzy model committe of ANFIS networks.

In figure 5: X – source vector – 256…64 points of PQRTS-complex taken at regular time intervals; $W_N = \left\lfloor \frac{H_N}{G_N} \right\rfloor$ – matrix recording of successively applying the Haar decomposition, where: H_N – low-frequency components matrix allocation (approximation coefficients); G_N – high-frequency components matrix allocation (detailing coefficients); $X^* = W_N X$ – direct Haar transformation; $X = W_N^T X^*$ – reverse Haar transformation; $Y = D^T\left(W_N X - M\right)$ – vector of extracted main components; $D^T = (V_1, V_2, \ldots, V_K)$ – eigenvectors corresponding to the first k-principal components,

forming the projection matrix; M – the mathematical expectation of each component of all observation in vector X.

For each sample x_j from the source data after the training of the network committee, a module for explaining the decision can be obtained:

1. For the *i*-th ANFIS network, which demonstrates the greatest response to the input sample, a vector of activities of fuzzy rules can be constructed.
2. Sorting the rules in order of decreasing its activity, the contribution of the first M rules to the summary conclusions can be evaluated and the types of their conditional part can be traced: the corresponding values of linguistic variables terms.
3. With the help of the transformation (9) it is possible to move from the reduced feature space to the initial basis and to interpret the values of the initial characteristics typical for each type of arrhythmia:

$$\overset{\circ}{\mathrm{X}} = \mathrm{W}_N^T \left[\mathrm{D}\left(\mathrm{D}^T \left(\mathrm{W}_N \mathrm{X}' - \mathrm{M} \right) \right) + \mathrm{M} \right] \quad (9)$$

X- values of the initial characteristics typical for i-th type of arrhythmia; X' – vector obtained as a result of the reverse output in $ANFIS_i$.

Described procedure of ECS processing was implemented in MATLAB and tested on a sample from the database of arrhythmias "MIT-BIH Arrhythmia Database" [2]. From this base records of diseases of 5 classes were selected, total amount of records that have been classified is 29.537. The choice of these classes is due to a sufficient number of records required for a complete training of the classifier. Selected classes are given in the Table 2.

Table 2. Notations of layers of the network.

Class (type of arrhythmia)	Class code	Amount of records
Paced beat	C1	6977
Atrial premature beat	C2	2452
Left bundle branch block beat	C3	8018
Right bundle branch block beat	C4	7091
Normal beat	C5	4999

Arrhythmias classification was carried out by using a multilayer perceptron and a neuro-fuzzy system committee implemented as neural networks. Each network in the committee determines the probability evaluation of having an appropriate arrhythmia class in the sample of input data. The network structure is presented in the table 3:

Table 3. Structure of the neuro-fuzzy system networks committee.

Input vector dimension	(8.1)
Number of networks in the committee	5
Number of rules in one network	15
Number of terms of a fuzzy variable	15
Membership function	Gaussian

Table 4 contains the comparative results of the classification of the multilayer perceptron and the neuro-fuzzy system performed in the form of a neural network committee.

Table 4. Computation results.

Parameter	Value MLP	Value NFS
Amount of records	29537	29537
Correctly recognized	25565 (86.55%)	25704 (87.023%)
Prediction error	3972 (13.45%)	3833 (12.976%)

Judging by the number of correctly classified records, it can be argued that the neuro-fuzzy model is not inferior to the neural network in terms of the classification quality. At the same time, the neuro-fuzzy model allows to rank the fuzzy production rules according to the level of their activation, and to

demonstrate by which elements of source feature vector a decision was made on whether the sample belongs to appropriate class of arrhythmia.

6. Discussions

When using a neural network classifier, various parameters can be used as characteristics of the sample of the analyzed area of the ECS, related to the frequency, time, and parametric forms. The key step in the analysis of the ECS in the diagnostic task is the construction of a classifier capable of deciding that the signal section belongs to a particular class of state in terms of the totality of the allocated features. Source space of characteristics can be transformed into one of the subspaces in which the separating hyperplane will have a simpler description. In this approach to the construction of an intelligent diagnostic system, it is difficult to implement a decision explanation subsystem.

With the use of a neuro-fuzzy classifier, it becomes possible to use the experience accumulated by cardiologists in the tasks of diagnosing certain types of arrhythmias. The existing medical methods can be adapted to a representation in a set of product rules of the form "IF-THEN", taking into account the adjustable weights that allow us to assess the degree of importance of both the individual feature and the rule as a whole.

Formalized expert knowledge in this case is transferred to the structure of the neuro-fuzzy module and becomes the core of the classifier. At the next stages of the diagnostic system construction, with DM techniques it becomes possible to identify hidden and unobvious patterns that complement expert knowledge, which will improve the accuracy of diagnosis. A feature of the use of neuro-fuzzy networks is the possibility of supervised learning on the existing base of marked samples assigned to different classes, which will allow to expand the base of diagnostic rules.

7. Conclusions

In this paper an approach to improve the intelligent systems of arrhythmia diagnostics on the basis of neural network classifiers by developing a solution explaining subsystem based on Mamdani neuro-fuzzy model is proposed.

The analysis of modern methods of automatic diagnosis of cardiac arrhythmias is carried out. It is established that the correctness of diagnostics of the most common algorithms in automated diagnostic systems does not exceed 80%.

The structure of a diagnostic system for the detection of cardiac arrhythmias based on intellectual data mining technology is proposed.

The issue of the knowledge base construction – coding, generation and selection of the features (database) as well as the construction of the rules base – as the part of the neuro-fuzzy diagnostic system is considered.

It is shown that using the neuro-fuzzy classifier it becomes possible to use the experience accumulated by cardiologists in the tasks of diagnosing certain types of arrhythmias. A model of the neuro-fuzzy solution explaining module is proposed to formalize the expert knowledge.

8. References

[1] Cardiovascular diseases (CVDs) *Fact sheet* (Access mode: http://www.who.int/mediacentre/factsheets/fs317/en/)
[2] Rangayyan R M 2015 *Biomedical signal analysis. Second edition* (Wiley – IEEE Press) p 720
[3] deChazal P, O'Dwyer M and Reilly R B 2004 Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features *IEEE Trans. Biomed. Eng.* **51(7)** 1196-1206
[4] Boonperm P, Supakasemwong D and Naiyanetr P 2015 ECG analyzing program for arrhythmia detection *7th Biomedical Engineering Int. Conf.* 1-4
[5] Chakroborty S 2013 Arrhythmia classification using auto-associative neural network *Proc. of the Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society* 4247-4250
[6] Ceylan R and Özbay Y 2007 Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network *Exp. Syst. with Appl.* **33(2)** 286-295
[7] Chakroborty S and Patil M A 2014 Real-time arrhythmia classification for large databases *36th*

Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society 1448-1451
[8] Gao D 2005 Bayesian ANN classifier for ECG arrhythmia diagnostic system: A comparison study *Proc. of the Int. Joint Conf. on Neural Networks* **4** 2383-2388
[9] Amann A, Tratnig R and Unterkofler K 2005 Reliability of old and new ventricular fibrillation detection algorithms for automated external defibrillators *Biomed. Eng. Onl.* **4** 60
[10] Nikonov A V, Vulfin A M and Gayanova M M 2017 Data mining algorithms in the problem of cardiac arrhythmia diagnosis *Proc. of a VI Scientific Conf. with Int. Participation "Information Technologies and Systems"* 200-205
[11] Rodriguez R 2014 Adaptive threshold and principal component analysis for features extraction of electrocardiogram signals *Proc. Int. Symp. on Computer, Consumer and Control* 1253-1258
[12] Kar S and Majumder D D 2016 An Investigative Study on Early Diagnosis of Breast Cancer Using a New Approach of Mathematical Shape Theory and Neuro-Fuzzy Classification System *Int. J. Fuzzy Syst.* **18(3)** 349-366
[13] Nikonov A V and Vulfin A M 2016 Neural network hardware-software complex for the cardiac arrhythmias diagnosis *Proc. of the X All-Russian Youth Scientific Conf. "Mavlyutovskie Readings"* 1593-1597
[14] Ahmed A F, Owis M I and Yassine I A 2014 Novel Bayesian classifier discriminant function optimization strategies for arrhythmia classification *Int. Conf. on Biomedical and Health Informatics* 693-696
[15] Holimchayachotikul P and Leksakul K 2017 Predictive performance measurement system for retail industry using neuro-fuzzy system based on swarm intelligence *Soft Comput* **21(7)** 1895-1912
[16] Dilmac S and Korurek M 2013 A new ECG arrhythmia clustering method based on Modified Artificial Bee Colony algorithm, comparison with GA and PSO classifiers *Int. Symp. On Innovations in Intelligent Systems and Applications* 1-5
[17] Faziludeen S and Sabiq P V 2013 ECG beat classification using wavelets and SVM *Conf. On Information and Communication Technologies* 815-818
[18] Chetan A, Tripathy R K and Dandapat S 2015 Cardiac arrhythmia classification from multilead ECG using multiscale non-linear analysis *Up Section Conf. on Electrical Computer and Electronics* 1-4
[19] Haykin S 2008 *Neural Networks and Learning Machines* (New Jersey: Pearson Prentice Hall) p 906
[20] Barro S 1998 Classifying multichannel ECG patterns with an adaptive neural network *IEEE Eng. in Med. and Bio. Mag.* **17(1)** 45-55
[21] Ghanwat T B 2015 Arrhythmia and neural network *Int. Conf. on Communication and Signal Processing* 1361-1365
[22] Osowski S, Markiewicz T and Hoai L T 2008 Recognition and classification system of arrhythmia using ensemble of neural networks *J. Int. Meas. Confed.* **41(6)** 610-617
[23] Jekova I 2000 Comparison of five algorithms for the detection of ventricular fibrillation from the surface ECG *Physiol. Meas.* **21(4)** 429-439
[24] Arumugam S S and Gurusamy S G 2009 Wavelet based detection of ventricular arrhythmias with neural network classifier *J. Biomed. Sci. Eng.* **2(6)** 439-444
[25] Uus A and Liatsis P 2011 Singular spectrum analysis for detection of abnormalities in periodic biosignals *Int. Conf. Syst. Signals, Image Process* 375-378
[26] Hadjem M and Nait-Abdesselam F 2015 An ECG T-wave anomalies detection using a lightweight classification model for wireless body sensors *Int. Conf. on Communication Workshop* 278-283
[27] Chapman P 2000 *Crisp-Dm 1.0* (Cris. Consort) p 76
[28] Rutkovskaya D, Pilihjskij M and Rutkovskij L 2008 *Neural Networks, Genetic Algorithms and Fuzzy Systems* (Moscow: Hot line – Telecom) p 452
[29] Coy C G 2010 A hybrid-genetic algorithm for training a Sugeno-type Fuzzy Inference System

with a mutable rule base *Theses and Dissertations* (Access mode: http://utdr.utoledo.edu/theses-dissertations/814)
[30] Pan J and Tompkins W J 1985 A Real-Time QRS Detection Algorithm *IEEE Trans. Biomed. Eng.* **3** 230-236
[31] Wang L and Jin Y 2005 *Int. Conf. Fuzzy Systems and Knowledge Discovery* (Changsha: Springer) p 1354
[32] Nikonov A V, Vulfin A M, Gayanova M M and Sapozhnikova M U 2017 Neuro-fuzzy extractor in the problem of cardiac arrhythmia diagnosis *Proc. of the XI All-Russian Youth Scientific Conf. "Mavlyutovskie Readings"*
[33] Lin C 1995 A Neural Fuzzy Control System with Structure and Parameter Learning *Fuzzy Sets Syst.* **70** 9-13
[34] Amin A-H N 2014 System of processing and neural network analysis of bioelectrical signals for solving problems of medical diagnostics *M.S. thesis Dept. 05.12.04 Electron. Eng.* 421-424
[35] Fomasa E 2014 HRV spectral and fractal analysis in heart failure patients with different aetiologies *J. Comput. Cardiol.* **41** 421-424
[36] Nikonov A V 2017 Cardiovascular diseases diagnosis on the basis of neural network analysis of the biomedical signals *Int. Symp. on Signal Processing and Information Technology*
[37] Ilyasov B G and Vasilyev V I 2009 *Intelligent control systems. Theory and practice* (Moscow: Radiotekhnica)
[38] Gaidel A V, Khramov A G, Kapishnikov A V, Kolsanov A V and Pyshkina Yu S 2017 A method for digital renal scintigram analysis based on brightness and geometric features *Computer Optics* **41(1)** 103-109 DOI: 10.18287/2412-6179-2017-41-1-103-109
[39] Soldatova O P, Lyozin I A, Lyozina I V, Kupriyanov A V and Kirsh D V 2015 Application of fuzzy neural networks for defining crystal lattice types in nanoscale images *Computer Optics* **39(5)** 787-794 DOI: 10.18287/0134-2452-2015-39-5-787-794

Acknowledgments
This work was supported by the Russian Foundation for Basic Research, research № 17-08-01569.

Integration Issues of Big Data Analysis on Social Networks

A V Ivaschenko[1], N Yu Ilyasova[1,2], A A Khorina[1], V A Isayko[1], D N Krupin[3], V A Bolotsky[3], P V Sitnikov[4]

[1]Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086
[2]Image Processing Systems Institute - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001
[3]IPSI SEC "Open Code", Yarmarochnaya Str. 55, Samara, Russia, 443001
[4]ITMO University, Birzhevaya liniya 14 lit. A, Saint-Petersburg, Russia, 199034

Abstract. Nowadays Social Media becomes one of the major providers of Big Data for analysis of users' behaviour, focus, trends, and deviations. One user can be presented in several social networks by various avatars. Most users have different dynamics of data processing and generation. In order to provide a solution capable to deal with this, there was developed and implemented a software library for integration with a number of social networks. This paper describes the problem, solution architecture and technical details of its implementation supported by the results of simulation and real data analysis for a number of popular social networks.

1. Introduction

The way we deal with the advent of the era of Big Data is crucial. Although this phenomenon has the right take place in conditions of uncertainty form the future, but with increasing automation of data collection and analysis - the number of algorithms that can extract and illustrate large-scale models of human behavior also increases. How do systems conduct this practice, and how do they regulate the flow of data?

The market sees big data like net opportunity: marketers optimizing their proposals, based on market analysis, Wall Street bankers process tons of information about the dynamics of changing rates. Legislation has already been suggested to limit the collection and storage of data, as a rule, about the inviolability of private life.

In recent years, the amount of information formed by business, science and social networks increases in geometric progression. This phenomenon is also called phenomenon known as a data stream.

In business, Valmart's valuation transaction databases estimate the amount of data currently stored in more than 2.5 petabytes of data, including: information on customer behavior and preferences, data about network activity and devices, information on market trends.

As for science, for example, the Large Hadron Collider (LHC) in The European Organization for Nuclear Research produced 13 petabytes of data in 2010.

In addition, the sensor, social networks, mobile data, subscriber data and the location data grow at a frenzied pace. Simultaneously with this growth in the volume of information, data also become more

interconnected. Facebook, for example, is almost completely connected, from 99.91% network to one, large connection component.

Modern social media can be treated as a major source of Big Data that describes the process of users' interaction and various information exchanges. Analysis of this data turns out to become a complicated technical problem: it is required to integrate with multiple social media for data import, associate separate profiles of the same users in different networks, match the facts of their interaction across the real events and derive basic trends and deviations.

To solve this problem there was developed a model of social media user behavior and a based on it software solution that provides capabilities for social networks analysis and simulation.

2. State of the Art

New opportunities of interaction in virtual environments allow Internet users to exchange the ideas immediately. At the same time everybody needs to obtain and process lots of incoming events. Under such informational pressure individuals start prioritizing the most important data, filtering and rejecting everything that is not currently interesting. Such a focus on the current interest instead of importance leads to various imperfections, including the creativity constraints. This process can be described by the modern principles of distributed simulation and decision-making support powered by multi-agent technology [1].

The virtual world of social media should be treated as a complex network of continuously running and co-evolving intelligent agents. Such solutions are based on holons paradigm and bio-inspired approach [2], which requires development of new methods and tools for supporting fundamental mechanisms of self-organization and evolution similar to living organisms (colonies of ants, swarms of bees, etc) [3]. As for the human beings represented by actors or agents, social network user should consider a combination of human and time factors. Interaction of customers and service providers powered by intermediary services generate and can be characterized by a big number of events that form Big Data and require modern technologies for its analysis [4].

Modeling the Internet users' behavior can be based on the modern principles of knowledge representation in the form of Ontologies [5]. These concepts allow formalizing self-organization and semantics, which is advantageous for abstract description of social concepts and their interaction in technical applications [6]. The papers discusses the detection process of uncharacteristic behavior of users [7] and methods classify users [8].

In the context of this paper there should be mentioned the papers on Internet development strategies [9], virtual communities and social networks studies [10-11]. Despite the successful application of mathematical statistics used to cluster and generalize the user's behavior the problem of Big Data analysis of social networks remains open. This happens due to a necessity to personalize user activity models and understand individual features of human behavior.

Our experience in the area of integrated information space development and its users' behavior analysis [12-15] can be used to build a software solution to derive basic trends in social media and provide intelligent functionality for social media big data analysis. The proposed abstract model and solution vision are given below.

3. Abstract Model

Let us present a community of Internet users by u_i, where $i = 1\backslash..N_u$ – a number of users. The activity of users information exchange can be presented by posts, comments or messages p_j, where $j = 1\backslash..N_w$ – an absolute number of an informational object. Post generation is an event

$$g_{i,j} = \left(u_i, p_j, t^0_{i,j}\right). \quad (1)$$

Issue or processing of an information object can be presented by an event $e_{i,j,k}$ that can be characterized by the combination of user, focus, and time:

$$e_{i,j,k} = e\left(p_j, \left(u_i, f_{i.k}, t_{i,j,k}\right)\right) = \{0,1\}, \quad (2)$$

where focus $f_{i,k}$ presents the current user interest and can be described by a tag cloud, which is a set of pairs:

$$f_{i,k} = \left\{ \left(\tau_n, w_{n,k} \right)_{i,k} \right\}, \tag{3}$$

where τ_n is a tag (keyword) with weight $w_{n,k}$.

The sequence of interdependent user focuses represents the evolution of the user's interest.

Each user has own ontology that forms the basis of his perception. It changes with time under the influence of learning and forgetting the information (presented by posts, comments or messages) and can be presented by a chain of contexts:

$$c_{i,m} = \left\{ \left(\tau'_l, w'_{l,m} \right)_{i,m} \right\}. \tag{4}$$

This change is correlated with user focus.

The focus cannot be considered new in order to provide positive perception, and at the same time it is not equal to the context to be able to excite interest.

Considering this correlation let us synchronize the context and focus changes:

$$e'_{i,j,m} = e' \left(p_k, \left(u_i, c_{i.m}, t_{i,j,m} \right) \right) = \{0,1\}. \tag{5}$$

The statements (2) and (5) are Boolean variables, which mean that appearance or perception of a post, comment or message does not guarantee changes in focus and context.

Events (2) and (5) can be used for analysis. One of the possible implementations is presented below. Study of the user's focus and context trends allows identifying tendencies, variations and iterations that form the patterns of user creativity.

In case new informational proposal remain suspended and does not make any effect over the user focus, this means that the user does not see any interest.

Possible reasons are concerned with context: additional education is needed to provoke such interest. On the other side, lots of changes of the user context indicate the search for a stable interest that should be proposed for the user at a certain time.

Context and focus can be also influenced by negative intervention.

In order to manage the user focus there can be generated a series of repeated affections partially covering the actual context and the targeted interest. Such patterns can also be identified applying cross-correlation analysis to the proposed model, which helps identification and resistance to negative informational influence.

4. Solution Architecture

The proposed approach is based on simulation of focus and context. It was implemented in multi agent architecture, which is presented in Figure 1.

Under the bounds of our proposed architecture we provide profile descriptor, post generator and navigator.

These are methods generated and used to simulate real activity of users in the social networks. Post generator used to create posts according to predefined logic.

Navigator is used to process incoming data which is described by network and can be presented like a sorted graph there the nodes are informational objects, for example web sites, documents, posts, comments, and the links are references between these objects.

Each object can refer to several other objects and documents; and the navigator according to all predefined logic decides which link to go.

In addition to navigator and post generator the informational frames are provided under the multi-agent architecture that correspond to a predefined above focus and context concepts. Focus is used to represent the current interest of Internet users.

Context is used to formalize informational space in which the agent performs its negotiating activity. Based on the provided model an algorithm has been developed for social media big data analysis.

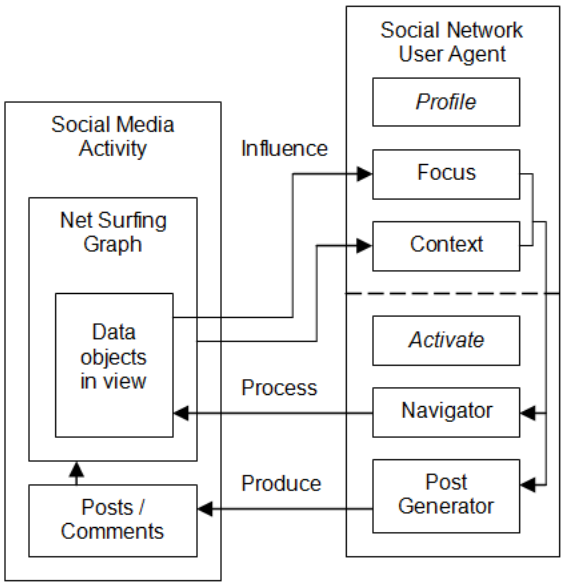


Figure. 1. Solution multi-agent architecture.

The model is used to formalize the social media user and integrate the analytical software with various social media open for data import and analysis.

This algorithm consists of 2 stages:

1. Calculation of the sample frequency vector for all users and development of the standard deviation vector for a variety of users.

You need to select topics and convert them to a view

$$\left(T_i, \frac{1}{N_{i,p}} \right), \tag{6}$$

where p is an hour of publication, $N_{i,t}$ is a number of users that posted on a certain topic in a time period p, T_i is an identificator of a topic.

Then process the received pairs and calculate the amount

$$\left(T_i, \sum_{k=0}^{i} \frac{1}{N_{k,p}} \right) \tag{7}$$

and calculate the standard deviation for each pair σ_j^k.

The obtained values are divided by the period.

2. Calculating the deviation metric for a particular user. You need to select topics and convert them to a key-value view. After that you need to process the data pairs and count the sum of topics with the same key:

$$\begin{cases} 0, \Delta_i \le 3\sigma_i, \\ 1, \Delta_i > 3\sigma_i \end{cases} \tag{8}$$

Count the deviation of a particular user, summarize the deviations of a particular user, divide the sum of the deviations by the number of topics (n) for a particular user and generate the resulting CSV-file in the form of a table with user data and information of standard deviation of this user.

One of the main features of Internet users' activity online that should be considered in the explored scope is mutual influence of contexts and focuses of communicating peers. This factor makes it possible to introduce the control loop: in addition to web content semantic analysis the platform starts to manage the users interest based on focus identification and context feedback.

This information is being collected in social networks and has all necessary details to get actual estimations. Still in this case it is required to provide integration with social networks and the data being processed contain tons of subjective assessments and perceptions.

Online libraries and professional communities are more neutral. For example, Wikipedia enforces various groups of authors to update the articles targeting maximum objectivity. Analysis of this data can help adequate identification of significant trends of consumers' focus identification that can be practically used e.g. for marketing and product placement.

Activation method is used to simulate multi agent activities in real time. The special agent dispatcher will call all the agents by using this activation method and after be activated each time an agent generates the time series period according to some distribution rule. Agent generates time series of navigation calls and post generation methods.

At these stages, we solve the proposed navigation and generation in such a way that we can model post-writers or readers and introduce some specific patterns of online activity, for example, the agent can be more active at night, or we can use some time frame for high / low activity.

Focus and contexts update the results of real agent behavior based on the influents of informational objects. We can generate focus and contexts according to our goals and in case we want agent to behave in a sort of specific way we introduce this control directly and formulate focus and context the agent will do that you want.

This approach allows simulating this influence and is introduced in the system. In this case we need to analyze the focus and context changes of the agent during the period of time. On the basis of analysis we introduce changes in focus by generating informational objects inside this network. This can be done in real systems using the contexts based advertisements. We can generate just the objects with certain informational context, which can be described by tag clouds.

The introduced architecture can be used to simulate online users in social networks and model realistic Internet behavior. In the area of simulation, practical application is generation of cognitive patterns of collective behavior based on self-organization. In this case the agent should be simple and the logic of focus and context should be close to very simple but generic behavior. This logic can correspond to know real users of social networks but it can represent some generalized behavior and the community of agent. Such behavior can be used to study and develop some visual cases.

In another case it is implemented as a sort of a frame, using which the algorithms of syntactical analysis or other large data analysis can look onto the real world of social networks and filter the data for intelligent study.

5. Implementation

To implement the proposed approach there was developed a software solution for social media focus identification based on knowledge discovery and Big Data analysis.

The solution can integrate with various data sources, pick out concepts, generate tag clouds for contexts and focuses and process their changes in time. Solution implementation architecture is presented in Fig. 2. The data imported from social networks is captured in database and can be processed either in real time or in batch mode.

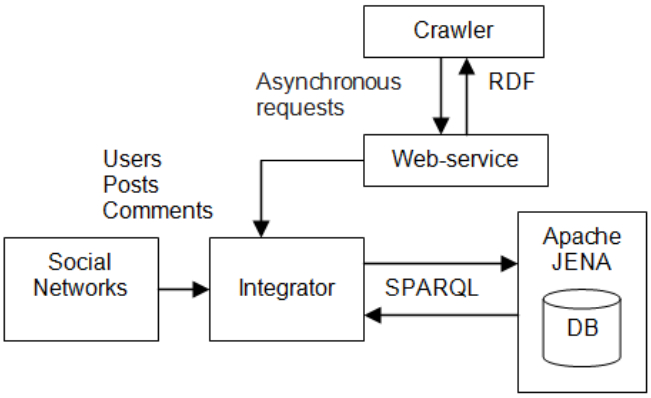


Figure 2. Integration model.

Crawler addresses asynchronously to a web service with requests for data from social networks. After receiving the request, the web service starts processing it. Next, the web service accesses the integrator, which starts downloading the requested data in the form of RDF / XML files, storing the intermediate data received from the single request of the crawler to receive the data by the single block to transfer the already downloaded ones.

Then in the background, i.e. in a mode where there is no need to control the data unloading process, the integrator automatically continues the embedded process and uploads the data to the database and uses Apache JENA to generate RDF / XML files that will be transferred to the first crawler address. The described model, software solution and its implementation was probated and tested using a typical data set derived from a number of social networks. In addition to a real regular result set of social media users' negotiation there was introduced a peak batch of posts generated by an online bot.

Apart from the social media (getting no a prior knowledge of a data structure) the big data analysis algorithms was able to identify the online bot influence. The results are presented in Fig. 3. Gray lines represent the annual trends of users' activity. The peak identified on Aug 15 corresponds the Bot activity and can be easily identified by the agent comparing the behavior of previous periods. The described research results show that the proposed model can be used for online behavior analysis and identification of negative informational influence.

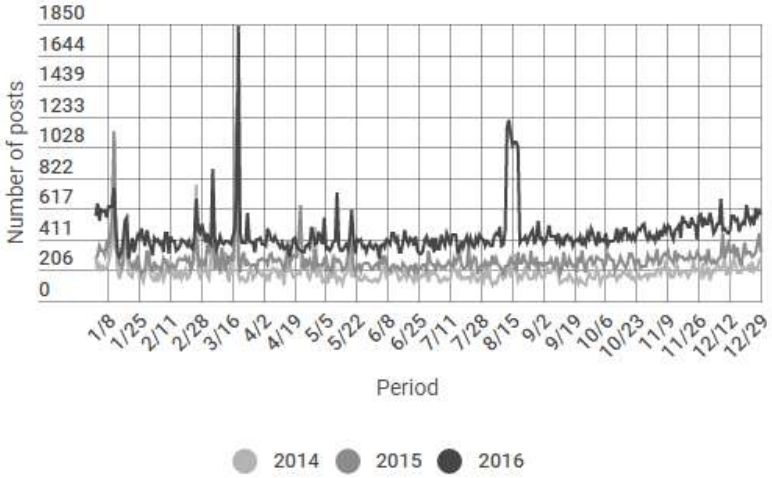


Figure 3. Bot activity identification.

There were processed the data of 32,000 users and their posts for the period of 2014 – 2017. To simulate the intervention there were modeled 50 bot users that automatically perform actions through interfaces intended for people. The given statistics show the distribution of posts throughout the considered time for each year. The horizontal axis, respectively, is temporary, contains the values of t recalculated in step 6 of the abovementioned algorithm.

Each line has 2 similar peaks at the beginning of the year (a detailed analysis showed that such emissions fall on holidays), they are similar to each other throughout the rest of the time. But the curve of 2016 has an unusual outburst (see 8/15), which characterizes the appearance of users' unusual behavior.

This example illustrates that the model and statistically developed patterns of users creativity can be used to identify negative deviations and the attempts to influence using repeated affections.

6. Conclusion

As shown above, the proposed model allows capturing the process of Internet user's activity considering a combination of human and time factors.

7. References

[1] Wooldridge M 2002 *An introduction to multi-agent systems* (Chichester: John Wiley and Sons) p 340

[2] Leitao P 2009 Holonic rationale and self-organization on design of complex evolvable systems *HoloMAS* **LNAI 5696** 1-12
[3] Gorodetskii V 2012 Self-organization and multiagent systems: I. Models of multiagent self-organization *Journal of Computer and Systems Sciences International* **51(2)** 256-281
[4] Bessis N and Dobre C 2014 *Big Data and Internet of Things: A roadmap for smart environments* (Berlin: Springer) p 450
[5] Mouromtsev D, Pavlov D, Emelyanov Y, Morozov A, Razdyakonov D and Galkin M 2015 The simple, web-based tool for visualization and sharing of semantic data and ontologies *CEUR Workshop Proceedings* **1486** 77
[6] *One Internet. Global commission on Internet Governance* 2016 (Access mode: https://www.cigionline.org/initiatives/global-commission-internet-governance) (01.11.2017)
[7] Shatalin R, Fidelman V and Ovchinnikov P 2017 Abnormal behavior detection method for video surveillance applications *Computer Optics* **41(1)** 37-45 DOI: 10.18287/2412-6179-2017-41-1-37-45
[8] Rybintsev A, Konushin V and Konushin A 2015 Consecutive gender and age classification from facial images based on ranked local binary patterns *Computer Optics* **39(5)** 762-769 DOI: 10.18287/0134-2452-2015-39-5-762-769
[9] Balakrishnan H and Deo N 2006 Discovering communities in complex networks *Proceedings of the 44th Annual Southeast Regional Conference* 280-285
[10] Wei W, Joseph K, Liu H and Carley K 2016 Exploring Characteristics of Suspended Users and Network Stability on Twitter *Social Network Analysis and Mining* 6-51
[11] Kadushin C 2012 *Understanding social networks: theories, concepts, and findings* (Oxford: Oxford University Press) p 264
[12] Ivaschenko A 2014 Multi-agent solution for business processes management of 5PL transportation provider *Lecture Notes in Business Information Processing* **170** 110-120
[13] Ivaschenko A, Minaev A and Spodobaev M 2015 Self-mediator software for sensor networks *Proceedings of the 2015 International Siberian Conference on Control and Communications (SIBCON)* 1-4
[14] Ivaschenko A, Lednev A, Diyazitdinova A and Sitnikov P 2016 Agent-based outsourcing solution for agency service management *Lecture Notes in Networks and Systems* **16** 204-215
[15] Protsenko V, Kazanskiy N and Serafimovich P 2015 Real-time analysis of parameters of multiple object detection systems *Computer Optics* **39(4)** 582-591 DOI: 10.18287/0134-2452-2015-39-4-582-591

Storage subsystem for intelligent oil wells control system

K F Tagirova[1], A R Ramazanov[1], A M Vulfin[1]and A V Nikonov[1]

[1]Ufa State Aviation Technical University, K.Marks St. 12, Ufa, Russia, 450077

Abstract. The purpose of developing data warehouse architecture is improving the storage subsystem as part of an intelligent oil well control system. To achieve this goal, it is necessary to develop the structure and information model of the storage subsystem. It is proposed to expand the exploration and exploitation data model of the Epicentre oil fields by adding new entities and a hierarchy of links. Highly effective architectural templates, taking into account the specifics of the research object and the accumulated experience of industry experts, are used. This will reduce the costs for development of the data warehouse architecture and will provide an opportunity for integration with existing information systems of a higher level of control and management – MES, ERP – and with other software and hardware complexes of the current level.

1. Introduction

Nowadays, the oil industry in Russia has entered to the stage of oil field development and oil reservoirs and wells exploitation on a scientific basis with fitting out oil fields and wells with modern equipment. Various “smart” production facilities are being actively implemented – intelligent wells and intelligent fields that allow reducing capital costs and reducing production costs by saving energy and improving control efficiency [4]. At the moment, the second generation of smart fields is being implemented, the main difference of which is realization of the field development remote control [5]. However, the implementation of such systems is impossible without the use of modern measuring instrumentation as well as control of technological equipment.

On the Russian Federation territory about 50 thousand wells are operated by installations of sucker rod pumps (SRP), which provide to 20% of all produced oil [1]. However, most of the wells from the SRP fund – low-yield, so it is economically inefficient to equip them with modern expensive controllers with complex intelligent control algorithms. It is necessary to use an intelligent control algorithms realized as additional modules in the control station of the SRP. Therefore, intelligent monitoring and control systems are proposed to solve the problem of control the SRP equipped wells. This system allows to solve the following tasks [6]:

1. Control of the technical condition of the well equipment;
2. Calculation of the well production rate to select control mode;
3. Control of the balance of pumping rocking machine.

During operation of such system the problem of processing and analyzing large volumes of technological data arises. Therefore is becomes an actual problem to organize a data warehouse for efficient access to the accumulated data and to apply an intelligent algorithms for well control. Appropriate requirements for such storage are [7]:

1. Integration with existing information systems of higher level of control and monitoring – MES, ERP [8];

2. The ability to integrate with other software and hardware complexes of the current control level, for example, with a variety of equipment monitoring systems;
3. Ensuring the logical integrity and consistency of the stored data as well as the coherence of data in distributed storage systems.

The purpose of data warehouse development is improving storage subsystem as a part of the intellectual oil well control system.

To achieve this goal it is necessary to solve the following tasks:

- Develop a structure of the storage subsystem,
- Develop an information model of the storage subsystem.

2. Structure development of a data warehouse subsystem as part of an intelligent oil well control system

Proposed information storage is a part of the intelligent oil wells control system which structural diagram is shown in Figure 1.

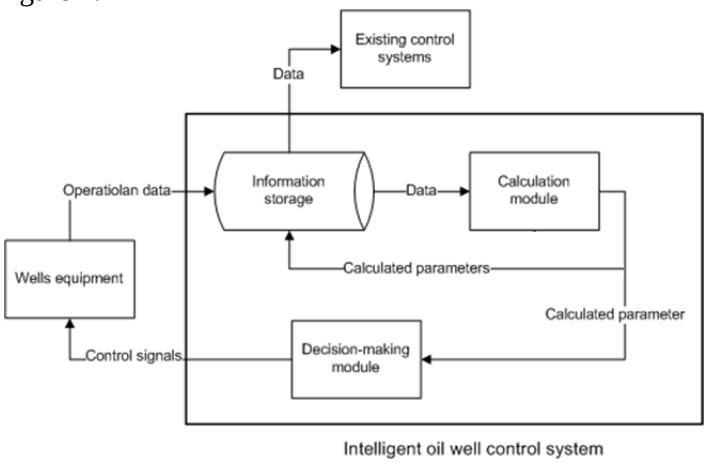


Figure 1. Structural diagram of the intelligent oil wells control system.

Equipment receives wattmetrograms and dynamograms of SRP operating, which are stored in the information storage. From this storage, the data enters the calculation module, which determines the state of equipment, calculates the flow rate and determines the balance of the rocking machine. These data are returned to the information storage and are also transferred to the decision module for determining control actions.

3. Information model development of a storage subsystem based on object-oriented decomposition

In order to use existing high-performance architectural templates for a proposed storage subsystem that takes into account the specifics of the research object and the experience gained by industry specialists embodied in existing information systems for processing and storing data of oil producing enterprises, the exploration and exploitation data model of the Epicentre deposits is taken as the basis. It is developed by the corporation POSC (Petrotechnical Open Software Corporation). Expanding the existing POSC model to implement intelligent management and diagnostic functions, it is possible to build an effective information model of the data warehouse taking into account design features [9, 22, 23, 24, 29].

Epicentre model is created independently of the applications and at a logical level, with no matter of the type of final data warehouse, what allows integration of existing solutions in case they are built with respect account the proposed model. This simplifies the task of communication of modules and systems from different manufacturers and following to the basic principles of open systems.

More than 1,000 existing technical and business objects associated with the oil exploration and production are defined in the Epicentre data model. In the terminology of POSC data modelling, these objects are named entities. Characteristics than may be contained in entities, defines in the model as

entity attributes. The most important are the attributes that define the relationship between the entities [2].

When designing data model for operating the technological parameters of intelligent control system of oil wells stations, it is necessary to enter new entity to expand Epicentre model, namely the entity of wattmetrogramm (*wattmetrogramm*), dynamogramm (*dynamogramm*), pretreated dynamogramms (*prepared_dynamogramm*) and classified dynamometer (*classified_dynamogramm*), which are necessary for solving the problem of diagnostics of equipment condition according to dynamometer data and for control tasks [10, 11, 13-21].

Wattmetrograms and dynamograms can be considered as part of technological time series (TTS) [26, 27]. These entities must be integrated with existing entities by building a hierarchy of links. The Epicentre model for providing communication instead of repeating the identifier, an attribute (attribute), pointing to another entity, is added in the logically integral element (entity). Thus, to organize relations with the rest of data, the entities of *watmmetrogramm* and *dynamogramm* contain a reference to the well (well), and the entities *prepared_dynamogramm*, *classified_dynamogramm* contain a reference to the dynamometer. For example, the entity of *wattmetrogramm* will be as shown in figure 2:

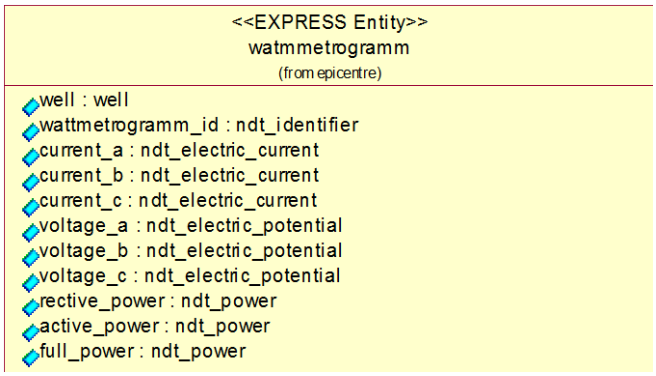


Figure 2. *Wattmetrogramm* entity.

In the Epicentre data model attributes presented by named types, having precise definition (*named defined types*), which consist of one or several basic data types. In the Epicentre data model documentation, named types with a precise definition have their names indicated by the prefix "*Ndt*". Therefore, while creating the data warehouse model, already existing types were used, but for complete description of the model two own types to storage dynamograms were implemented: plunger stroke counts array – *Ndt_position_array* and counts of loads per rod array – *Ndt_load_array*. Figure 3 shows the basic types used during development of warehouse:

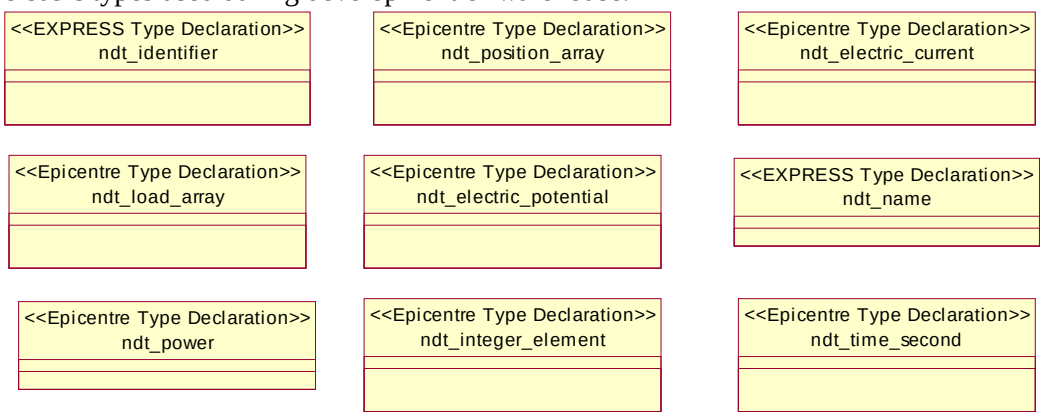


Figure 3. Main named types used in the development of warehouse.

Thus, the overall structure of the information model is as follows (Figure 4) [22-25]:

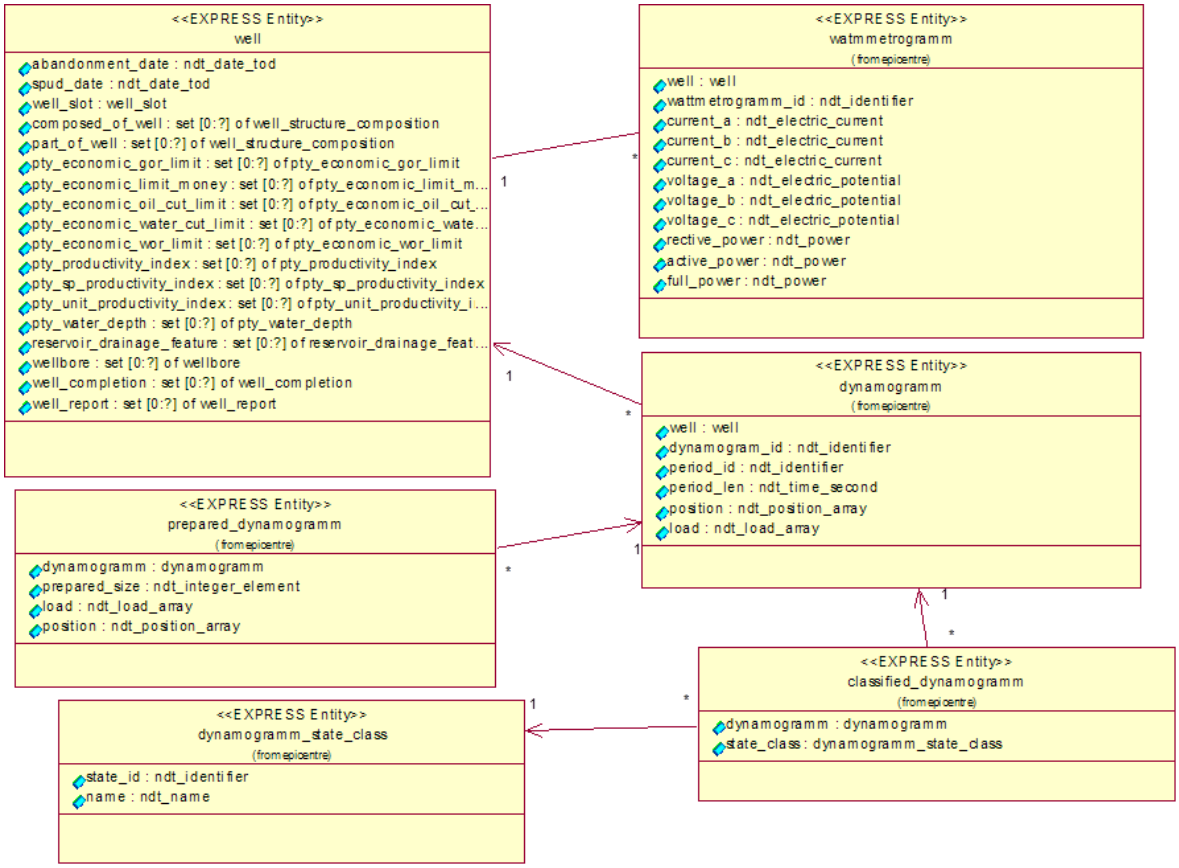


Figure 4. General structure of warehouse model.

Epicentre data model is described in EXPRESS language, but EXPRESS is not the equivalent of the data definition language (DDL), such as the Standard Query Language (SQL) DDL [3]. Therefore, for development of data warehouse it is necessary to convert expressions in EXPRESS language into a set of expressions on DDL language. This conversion POSC corporation called projection. So the creation of the physical warehouse is presented in Figure 5.

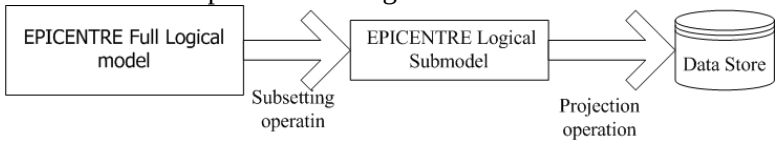


Figure 5. Creating a data warehouse model of EPICENTRE.

In order to project Epicentre model into a set of expressions on DDL language, POSC Corporation proposes to use the tool to project Epicentre data model to the four types of Relational Database Management System (RDBMS) DDL: ANSI, Informix, Oracle, and Sybase.

Based on the Epicentre 3.0 model and the subsets of the Epicentre “AC3 data model: Process Data”, “DM1: Document Specification”, “DM3: Document Schema”, “DM4: Document Content”, it is forming an information model of data processed in the intelligent oil well control system – time series – and in the subsystem for storing the parameters of the intelligent analysis module (Figure 6).

Detailed disclosure of attributes of key entities and basic links is given in the section devoted to the development of the Intelligent Control System (ICS) PC (Figure 7).

From Epicentre data model it was selected three groups of entities characterizing the different stages of processing the technological time series as the main data source for implementing the methods of intellectual analysis, additional elements that store and process the accumulated data.

In [28] the application of intelligent analysis technologies in diagnostics and control tasks is shown: committees of neural network models and preprocessing algorithms for technological time series based on the application of wavelet analysis.

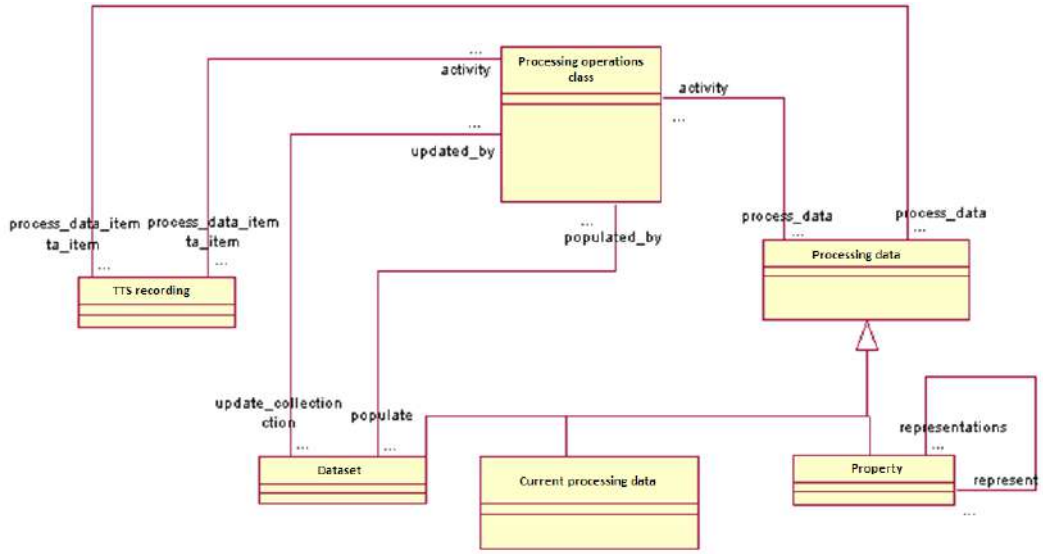


Figure 6. Truncated and augmented subset AC3: Process Data.

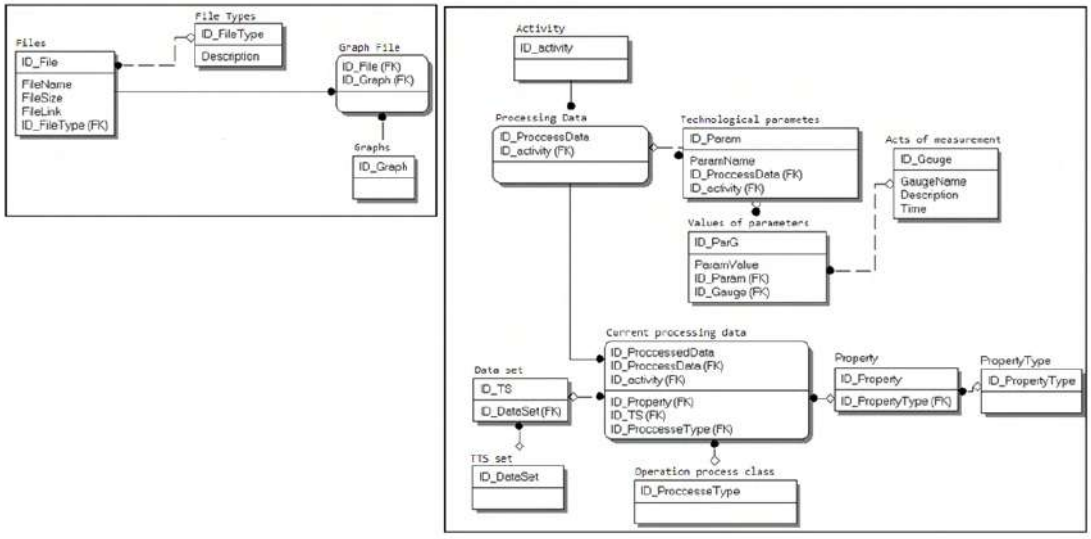


Figure 7. Prototype of the logical model of technological time series processing in the ICS.

The process of life cycle (LC) management of the developed software complex, which implements the modules for intellectual analysis and management of oil producing wells, corresponds to the PDCA cycle and the basic requirements of the process approach, formulated in MS ISO 9000: 2000, as a functional model. ISO / IEC 12207: 1995 “Information Technology - Software Life Cycle Processes” is the main normative document regulating the composition of the software life cycle processes (Figure 8).

Therefore, an information model of TTS processing as part of an intelligent control system for a group of oil producing wells was developed as a subset of the POSC Epicentre model, which allows the use standard interfaces between application programs, database management systems, workstations and users.

4. Conclusions

Thus, it is proposed to expand the Epicenter model by adding new entities and hierarchy of links in the task of the information model construction to improve the subsystem of data storage as part of an intelligent oil well control system, which will allow:

1. Reduce the labour costs for the development of the data warehouse by using existing high-performance architectural templates that take into account the specifics of the research object and the accumulated experience of industry experts.

2. Provide an opportunity for integration with existing information systems of a higher level of control and management – MES, ERP – and with other software and hardware complexes of the current control level through the use of the principle of open architecture.

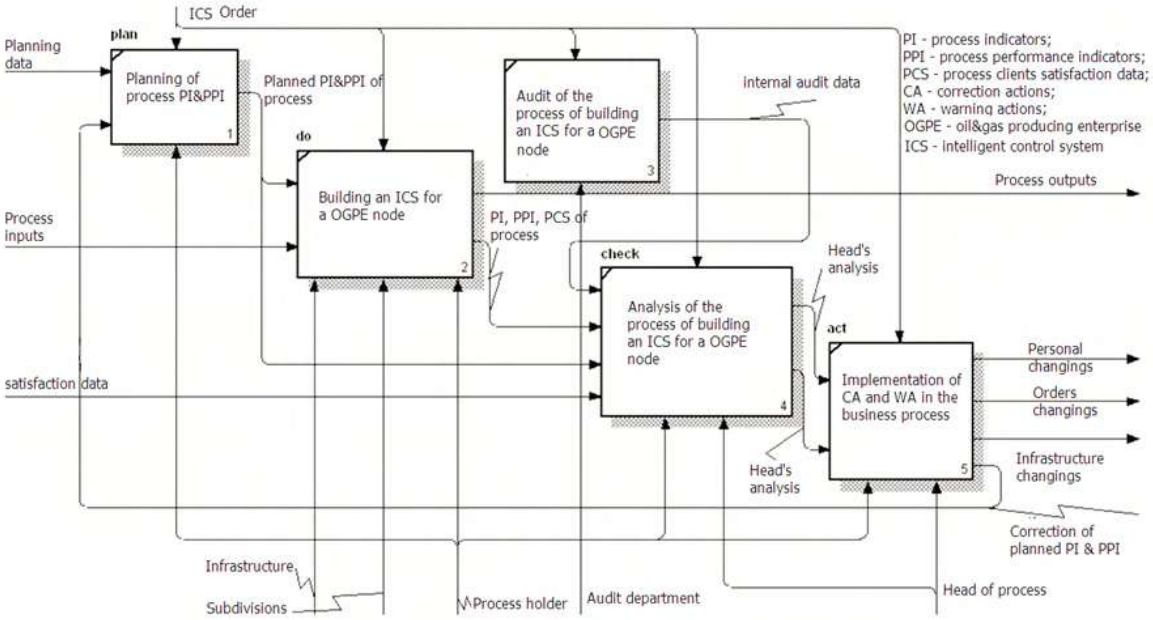


Figure 8. Management of the software development process.

5. References

[1] *Market research facilities sucker rod pumps (SRP). The analytical report* (Access mode: http://www.techart.ru/files/research/ walking-beam-pumping-unit.pdf) (20.11.2017)

[2] Almuhametov A A, Giniatullin V M, Zozula U I and Filippov V M 2003 *Learning objects and relationships in Epicentre information-logical model ofexploration and exploitation oilfields* (Ufa)

[3] *Epicentre Sample Relational Implementation* (Access mode: http://w3.energistics.org/archive/ Epicentre/Relational_Implementation_Methodology_v2.3/main) (20.11.2017)

[4] Gubarev M I 2016 Opportunities and perspectives for the construction of “smart” wells *Materials of XX international scientific symposium of students, postgraduates, and young scientists devoted to the 120 Anniversary of the founding of Tomsk Polytechnic University* **2** 723-726

[5] Dmitrievskii A N and Eremin N A 2017 *Resource-innovation model and the solution of actual problems of development of oil and gas* (Access mode: http://oilconference. ru/d/304647/d/) (20.11.2017)

[6] Mazakov E B 2014 Representation and processing of knowledge in information automated systems of intelligent field *Notes of the Mining Institute* **208**

[7] Yurev A M 2011 Unification of control systems by the information exchange on mes production managements *Automation, Telemechanization and Communication in Oil Industry* **4** 17-20

[8] Kizina I D and Sanarova K A 2011Methodology of development and implementation of oil and gas industry companies energy menegement support information system *Automation, Telemechanization and Communication in Oil Industry* **4** 6-1

[9] Chengfang L 2014 Construction and Application of the Data Model in Petroleum Exploration and Production Fields *International Geophysical Conference & ExpositionSociety of Exploration Geophysicists and Chinese Petroleum Society* 1319-1322

[10] Tagirova K F and Vulfin A M 2013 Neural network algorithms of information processing in the tasks of diagnosing downhole pumping equipment of oil company *Automation, Telemechanization and Communication in Oil Industry* **12** 28-32

[11] Vulfin AM and Tagirova K F 2015 Enhancement of accuracy of deep-pumping equipment based on data minning *Optical Memory and Neural Networks* **24** 28-35
[12] Ayvazyan S A, Buhtshtaber VM and Enyukov I S 1989 *Applied Statistics: Classification and reduction of dimension* (Moscow: Finance and Statistics)
[13] Vasilyev V I and Ilyasov B G 2009 Intelligent control systems. Theory and Practice *Training material* 33-62
[14] Vorontsov K V 2017 *Machine learning. Lecture courses* (Access mode: http://www. machinelearning.ru/wiki/index.php) (20.11.2017)
[15] Belov I G 1960 *Investigation works well pumps dynamography* (Moscow: Gostoptekhizdat)
[16] Mansafov R Y 2010 A new approach to the diagnosis of sucker rod pumps work on the dynamometer *Engineering Practice* **9** 92-99
[17] Vulfin A M and Frid A I 2014 Diagnosing the state oil company equipment based on data mining techniques *II international conference "Information technology intellectual support of decision-making"*
[18] Vulfin A M and Frid A I 2014 Safety Increasing of Oil Companies Engineering Networks Operation with Use of Artificial Intelligence Systems *16 International Science Conference CSIT* **3** 53-58
[19] Tagirova K F, Vulfin A M, Ramazanov A R and Fathulov A A 2015 Modified algorithm determination of current paramters of SRP by dynamometric card data *Automation, Telemechanization and Communication in Oil Industry* **12** 37-41
[20] Tagirova K F, Vulfin A M and Bulgakova G T 2016 Coordinated control system for group of production wells on the basis of a hierarchical system of dynamic models *IXscientific-practical conference "Mathematical modeling and computer technologies in field development processes"*
[21] Vulfin A M, Frid A I, Giniyatullin V M, Shuvalov S A and Salikhova M A 2012 The information processing system for diagnosing oil engineering networks based on data mining techniques *Bashkirsky Chemical Journal* **4** 72-78
[22] Fowler M 2004 *UML distilled: a brief guide to the standard object modeling language* (Addison-Wesley Professional)
[23] Alexandrescu A 2001 *Modern C++ design: generic programming and design patterns applied* (Addison-Wesley Professional)
[24] Gamma E, Helm R, Johnson R and Vlissides J 1994 *Design Patterns. Elements of reusable object-oriented software* (St. Petersburg: "Piter") p 417
[25] Levykin V M, Evlanov M V and Kernosov M A 2014 *Design Patterns requirements for information systems: modeling and application*
[26] Vulfin A M and Frid A I 2011 Neural network model of time series analysis within the framework of the Data Mining methodology *Information-control systems* **5**
[27] Vulfin A M and Frid A I 2011 Intelligent automated decision support system for the technological complex of oil acceptance and delivery *Mechatronics, automation, control* **5** 29-34
[28] Tagirova K F 2014 Improvement of the USSN diagnosis based on data mining dynamometer *Automation, Telemechanization and Communication in the Oil Industry* **11** 23-28
[29] Vasin Yu G and Yasakov Yu V 2016 Distributed database management system for integrated processing of spatial data in a GIS *Computer Optics* **40(6)** 919-928 DOI: 10.18287/2412-6179-2016-40-6-919-928

Application of the principal component analysis to detect semantic differences during the content analysis of social networks

I A Rytsarev[1,2], D D Kozlov[1], N S Kravtsova[1], A V Kupriyanov[1,2], K S Liseckiy[1], S K Liseckiy[1], R A Paringer[1,2] and N Yu Samykina[1]

[1]Samara National Research University, Moskovskoe shosse 34, Samara, Russia, 443086
[2]Image Processing Systems Institute - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001

Abstract. In this paper, we propose an approach to semantic differences detection in texts presented in the form of frequency dictionaries. The original text data has been obtained by collecting records on various online communities. We have implemented a specialized software module that allows us to analyze and download both posts and comments from the social network VK's open communities. To build our frequency dictionary, we have developed an algorithm that takes into account the peculiarities of the data collected from social networks. In the article, we propose an approach based on the use of methods reducing the dimension of feature spaces to identify keywords based on the analysis of their frequency of usage. The algorithm we present uses the principal component analysis technique. As a result, we have shown that by using the coefficients of the obtained linear transformation, it is possible to estimate the importance of words. With the help of these estimates, we were able to identify not only key words, but also semantic differences in social networks communities. The proposed approach can also be used to form metrics and calculate the social distance between Internet communities.

1. Introduction

Development of the newest informational and industrial technologies, dynamics of its wide implementation influence all the personal and public activities in everyday life. Nowadays we can see rapid transformation of ways of interaction between different persons, groups and organizations.

Virtual network communities can be described as social groups in terms of social psychology (there are common interests and activities in such groups, awareness of membership, sense of collective "we", and possibility of direct personal contact between members). At the same time, these communities have specific features because of its virtual, digital nature. Participation in most of virtual communities is not bounded by age, gender, social status or territory. These communities have free membership (there are no entrance fees, membership cards, formal obligations). Members of the communities are able to enter any of it or leave it at any moment as they wish. It allows people to choose a content according to their interests and motives. Besides, in most cases it is not necessary to be a member of virtual community to see its content.

Communication via internet is so wide-scaled that it requires development of new methods to study and analyze it [1] point out the following features of data from social networking services (SNS) that create certain methodological difficulties:

- subjective selectivity and social desirability of the content published by users;
- possibility to delete the content previously published by users;
- existence of fake accounts and bots producing unpersonal, unified, automatic content;
- features of SNS itself, its possibilities and interface implementations that bound and induce certain activities of users at the same time, thus creating a discrepancy between users' behavior in virtual and real life communities;
- autocompleting of forms and choices "by default", often made without awareness, probably not consistent with real choice or opinion of users;
- ethical issues concerning the use of personal information and the right of a user to delete it permanently.

Nevertheless, ability to quantify and classify digital footprints of nearly endless quantity, high speed of information processing (even online), its verifiability, transparency, and low financial costs are create new opportunities and benefits. However, we have to invent new methods to gather and analyze such data to avoid difficulties mentioned above [2].

Content users publish in SNS, and digital footprints, allow revealing and taking into account personal psychological characteristics (values, emotions, mood, self-regulation strategies and motives) essential to economic, social and political decision making [3]. Moreover, SNS make it possible to monitor these characteristics and its dynamics in real-time. There have been already developed trading algorithms that use as input variables actual social mood calculated upon recent posts in SNS (namely, Twitter, LifeJournal or Facebook) [4, 5], and these algorithms proved to be very effective in a real stock market [6]. Similar psychological characteristics influence not only economic decisions, but political and social decisions and behavior as well [7, 8].

However, prediction of psychological characteristics of SNS users is still much less precise than prediction of social or demographic characteristics. For instance, in one of the largest up-to-date studies by Kosinski et al. [9] social and demographic characteristics (such are, e.x., sex or sexual orientation) of SNS users were predicted with high accuracy, while error rate for personality traits was relatively high. Nevertheless, even such not so precise results can be useful to optimize search requests and personal ads on webpages [10]. If we want to lower the error rate in diagnostics and prognosis, we have to go deeper and analyze digital footprints not as behavior markers only. We have to follow new directions and analyze meaning and semantics of posts being published by users [11], and invent new mathematical and statistical methods to execute such analysis [12, 13].

These studies demonstrate potential benefits of analysis of big data obtained from SNS [14, 15]. However, they are the first steps in a long journey. The models proposed in these and other studies include generalized characteristics of very large samples and communities only [16]. It is necessary to develop more precise models to lower error rates in prognosis and decision making. To do this, we have to take into account subjective semantics and psychological features of particular social communities and groups in SNS. One of the tools to achieve this goal is the «Social Sonar» created at Samara University. It is based on new methods of analysis of group dynamics in SNS. In this case, group dynamics refers to all the processes of social groups' life cycle: its creation, functioning, development, stagnation, involution, and vanishing. The present study propose a new method of modelling semantic similarities and differences between SNS communities. This method allows raising the efficacy and precision of such analysis.

2. Data collection and preprocessing

We selected VK for the analysis of communities the Social network, as it is one of the most popular Russian-speaking one in the Internet. A feature of the social network is that it is available to all and does not have a strictly defined theme. In addition, this social network freely provides the application software interface for writing external applications.

Retrieving data from the social network VK is possible only when using standard tools provided by the developers of the network, it is therefore necessary to access the servers of the social network VK.

To do this, it is needed to create an application on the VK server and get the access keys by following the steps below:

1. In the developer section, follow the instructions in the create applications wizard to create an application.
2. Get the application ID and secret key to quickly connect the application to the servers.
3. In the app settings, give the app permissions to work with communities and records.

The next stage of our work was the development of a software module for data collection. The implementation was carried out in Python, using scripting library for VK. This library is developed by third-party developers and has many methods built on the official program interface's. After installing the module into the system, it is imported and an authorization object is created which stores the login information. All interaction with the social network are then carried out threw the module.

As a matter of understanding, here are some terms used in the upcoming paragraphs:. A record is any text message in the community. A post is a record that carries information about one or several events, most often prompting you to start discussing a topic. Comments are entries below the post that reflects the reaction of a specific user to the post or a comment of another participant.

The stage of collecting records from the wall consists of two procedures: the procedures for collection of posts and the collection of comments.

The procedure for collecting posts initializes a special method of the official application data collection system, which returns a list of posts of a selected user or community and processes its posts. To initialize the method, you must specify a unique user ID or a unique community address from which to download information.

The procedure of collecting comments to the post is carried out the same way: through the initialization of the program interface method's and the process of its answer. The difference is in the method used and its required set of parameters: to receive comments, you must specify a unique identifier not only for the community, but also for the post.

The use of the two procedures described above allows the collection of information from any open VK communities, but there is a software limit on the number of requests per day, which significantly complicates the collection of information. So third-party applications have a daily limit of 2500 requests and may receive no more than 100 records per request.

Also, the properties of records contain the date of publication. To reduce the number of requests, it is possible to limit the number of records received in accordance with its time interval. Comparing the time in the record with the interval of interest, you can get only the records corresponding to the time interval of interest.

3. Algorithm for frequency dictionary compilation

Records are a special information storage structure, containing both text fields (the direct content of the record) and metadata, containing information such as identifiers of the record, its author or time data.

Since the basis of the research carried out in this work is the compilation of a frequency dictionary, the developed algorithms relate primarily to the field containing the text of the records under consideration, but the rest of the information of the records can also potentially be taken into account and used for processing.

After referring to the corresponding field of the record and receiving the data, the resulting text fragment is divided into words. Word, text in this case is called sequence of alphabetic symbols, separated by spaces, numbers or punctuation marks. Since the entries in social networks are full of typos and errors, the spelling check of the data obtained was the necessary step for the correct accounting of the number of words in the preparation of the frequency dictionary. A third-party library was used to perform the spell checking procedure for this work. In parallel, the number of posts and comments is taken into account.

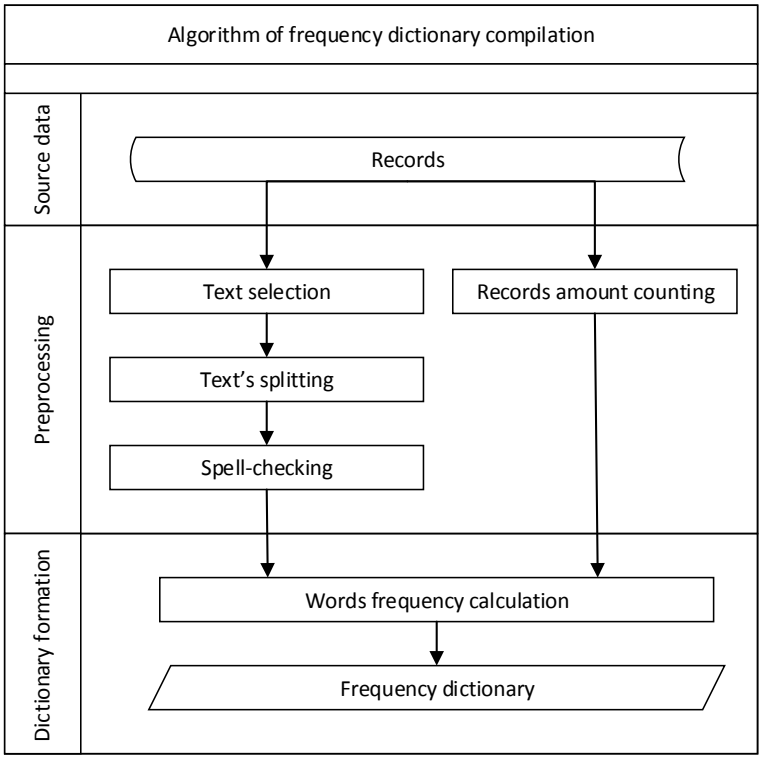


Figure 1. The sheme of algorithm for frequency dictionary compilation.

After performing the described steps, counting the number of occurrences of each unique word w in the whole set S of text data was carried out by the formula:

$$count(w) = \sum_i (w_i \in S)$$

The resulting number of words is divided by the number of entries and thus estimated frequency of use of words in the records – formed frequency dictionary. The scheme of the developed algorithm is presented in Figure 1.

4. Algorithm for semantic differences detection

We developed a matching algorithm for the identification of semantic differences on the basis dictionaries frequency analysis. The concept of this algorithm is to apply existing techniques to reduce the dimensionality of the feature space in order to rank words. The principal component analysis technique [17] was used in this article.

The initial data for the algorithm are frequency dictionaries. Each dictionary can be represented by a vector of words. With two dictionaries, it is possible to form one new attribute that will separate these dictionaries in the best way. A new feature using the principal component analysis is formed by multiplying the values of the initial space by the corresponding vector of coefficients. The values of the vector coefficients used to form a new feature can be used as an estimate of the word's contribution to the formed feature. Thus, it is possible to make a list of words that have made the greatest contribution to the feature's formation. The list of words compiled in this way describes a feature that provides the dictionaries separations and describes their semantic difference.

The first step of the algorithm is data normalization. The values of the frequency of word usage were normalized in order to obtain dictionary values in frequencies in the interval [0;1] according to the following formula:

$$y(x) = \frac{x - x_{min}}{x_{max} - x_{min}},$$

where x_{min} – minimum value among elements in the vector, x_{max} is the maximum value among elements in the vector.

The next feature of the algorithm is the application of the principal component analysis.

Principal component analysis is one of the main ways to reduce the dimension of data with the lesser loss of information. The calculation of the principal components is usually reduced to the calculation of the eigenvectors and eigenvalues of the covariance matrix of the original data. By definition, the covariance of two features X_i and X_j is calculated as follows:

$$cov(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)] = E[X_i X_j] - \mu_i \mu_j,$$

where μ_i — - mean of the i sign. The covariance matrix is a symmetrical matrix, where the diagonals represents the feature's dispersions and outside the diagonals are the covariations of the corresponding pairs of features. The Rayleigh relation [18] implies that the maximum variation of the data set will be achieved along the eigenvector of this matrix corresponding to the maximum eigenvalue. Therefore, the main components on which we project the data are simply eigenvectors of the corresponding eigenvalues of the matrix. The values of eigenvector elements are therefore the desired estimated coefficients for the formation of words describing semantic differences. To obtain a visual interpretation of the relative positions of the dictionaries in the resulting semantic space, we need to multiply the frequency of word use vector's with its corresponding eigenvector.

5. Results

To study the performance of the developed algorithms, we selected communities of similar areas of interest. All the selected communities are communities of Samara and Samara region residents and three of the five communities are open communication platforms for students and teachers of Samara's two largest Universities. Due to the existing restriction on data collection within the framework of this study, we introduced an additional criterion for selection of records based on the time of their publication (from January 1 to May 20, 2018. The following communities have been selected for data collection and processing:

I "Overheard in Samara University".
II "Overheard Samara University" (old name: "Overheard in SamSU").
III "Overheard in SamSTU".
IV "Heard Samara".
V "Overheard Samara".

Table 1. Amount of posts and comments.

Record	Community number				
	I	II	III	IV	V
Post	431	841	1486	2060	5918
Comment	2076	1711	4108	28682	257163

Table 2. A fragment of frequency dictionaries.

Words		**Community number**				
Original	**Translation**	**I**	**II**	**III**	**IV**	**V**
он	**he**	0.038	0.020	0.018	0.038	0.043
очень	**very**	0.037	0.032	0.024	0.032	0.018
вам	**you**	0.021	0.029	0.015	0.030	0.022
мы	**we**	0.029	0.019	0.015	0.025	0.023
нужно	**need**	0.023	0.027	0.021	0.017	0.015
лучше	**better**	0.014	0.014	0.014	0.018	0.015
всем	**to all**	0.013	0.025	0.019	0.017	0.015
потом	**later**	0.016	0.013	0.013	0.019	0.018
еще	**more**	0.013	0.013	0.014	0.022	0.021
время	**time**	0.020	0.026	0.013	0.014	0.014
нас	**us**	0.027	0.017	0.012	0.023	0.019
быть	**be**	0.019	0.014	0.012	0.017	0.017
много	**many**	0.014	0.013	0.014	0.013	0.011
кого	**whom**	0.053	0.018	0.028	0.011	0.011
день	**day**	0.013	0.024	0.010	0.014	0.011
всех	**of all**	0.016	0.013	0.010	0.014	0.016

After collecting records of the selected communities for the given period of time, in accordance with the frequency dictionary algorithm, we selected texts of the various records, before splitting texts into words. Spell-checking was then performed and the number of posts and comments for each of the studied communities was calculated (table 1). The frequency dictionaries obtained as a result of the algorithm are partially presented in table 2.

The graphical representation of the results of the principal component analysis application to frequency dictionaries is shown in figure 2.

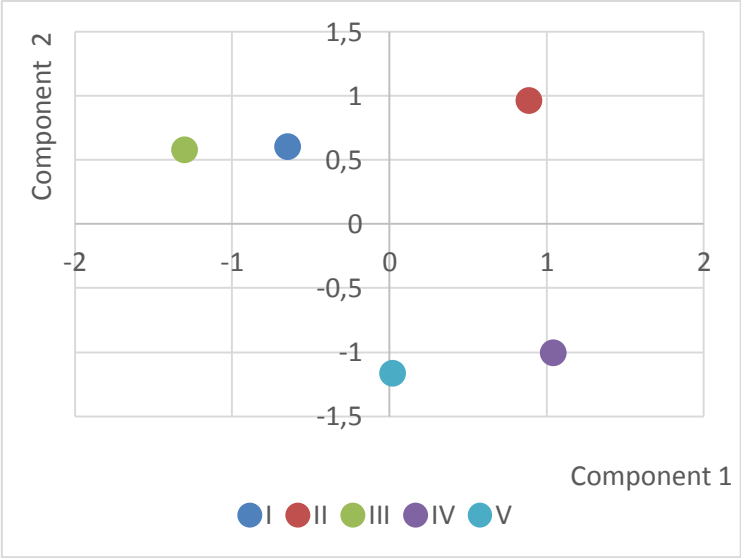


Figure 2. Graphical representation of the principal component analysis.

As can be seen from figure 2, the communities were divided into 3 groups:

- University communities "Overheard in Samara University" and "Overheard in SamSTU".
- City communities "Overheard Samara" and "Heard Samara".
- University community "Overheard Samara University" (old name: "Overheard in SamSU").

The words describing the formed components are presented in table 3.

Table 3. Formed components.

Component number			
1		2	
Original	**Translation**	**Original**	**Translation**
вам	you	имена	First names
нас	us	дома	at home
мы	we	работа	work
всем	to all	интересно	interesting
был	was	могу	can
были	were	Самара	Samara

Meaningful interpretation of the metrics is based on the semantic meaning of its unique components. There are many plural nouns (“you”, “us”, “we”, “to all”) among them. This fact points to the importance of community identity and expressing oneself by belonging to a certain community. It seems that communities with high scores on this metric represents some real or virtual community in virtual space, highlighting social identity of its members, making it more salient. Another peculiarity of the components of the first metric is prevalence of the past term verbs (“was”, “were”). It means that for the members of such communities events that have already happen are more interesting than actual or future events. This is probably because members of such communities are more conservative in their views and more thoughtful and critical about social changes.

Unique components of the second metric consist largely of first names of persons. It possibly implies that conversation in communities scored high on second metric is more personal and direct. Other components ("at home", "work", "can", "interesting") are about mundane and usual activities. It is plausible that communities high on this metric are more about daily routine, and its members are more concerned about their personal circumstances and are not much interested in discussing topics about their community or society.

Scores of the measured communities on the obtained metrics support the interpretation given above. For instance, the second metric clearly separates city communities from university communities. We can say that discussion topics of analyzed university communities are about students' life, not about university events. On the other side, topics of analyzed city communities are more concerned about city events. In other words, university publics are for students mainly (and not for the staff), but city publics are for all the citizens who are active in social networks. These results also shows that the younger members of public are, the more this public should score on the second metric.

Scores of the analyzed publics on the first metric implies that community and social identity issues are not very interesting for students, they are more concerned about everyday routine. Nevertheless, "Overheard Samara University" community gained unusually high scores on the first metric. This is most likely because members of this community still identify themselves as students and alumni of former Samara State University that has been united with Samara National Aerospace University into Samara University a few years ago. We may say that active members of "Overheard Samara University" community feel sorry and nostalgic about the former university, and such feelings are not common among the members of the new community "Overheard in Samara University" of the same university. This fact makes it plausible that there are more alumni in "Overheard Samara University" than in newer "Overheard in Samara University" community.

6. Conclusion

Based on the analysis of the results, we can conclude that the developed algorithms allows us to evaluate the semantic differences of the Internet communities of social networks. The founded components can be used as metrics for analyzing the relative positions of the considered communities. Results can be improved by eliminating filler-words from the frequency dictionaries and grouping words by parts of speech. The task of further research is the development of technology for automatic filtering of frequency dictionaries.

7. References

[1] Kosinski M, Matz S C, Gosling S D, Popov V and Stillwell D 2015 Facebook as a Research Tool for the Social Sciences: Opportunities, Challenges, Ethical Considerations, and Practical Guidelines *American Psychologist* **70(6)** 543-556

[2] Spitsyn V G, Bolotova Yu A, Phan N H and Bui T T T 2016 Using a Haar wavelet transform, principal component analysis and neural networks for OCR in the presence of impulse noise *Computer Optics* **40(2)** 249-257

[3] Raynard R 2017 *Economic Psychology* (Hoboken, NJ: John Wiley & Sons) p 512

[4] Rytsarev I A and Blagov A V 2016 Classification of Text Data from the Social Network Twitter *CEUR Workshop Proceedings* **1638** 851-856

[5] Rytsarev I and Blagov A 2017 Creating the Model of the Activity of Social Network Twitter Users *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)* **9** 27-30

[6] Nofer M, Hinz O 2015 Using Twitter to Predict the Stock Market *Business & Information Systems Engineering* **57(4)** 229-242

[7] Trottier D and Fuchs C 2014 *Social Media, Politics and the State: Protests, Revolutions, Riots, Crime and Policing in the Age of Facebook, Twitter and YouTube* (New York: Routledge) p 252

[8] Housholder E E and LaMarre H L 2014 Facebook Politics: Toward a Process Model for Achieving Political Source Credibility Through Social Media *Journal of Information Technology & Politics* **11(4)** 368-382

[9] Kosinski M, Stillwell D and Graepel T 2013 Private Traits and Attributes are Predictable from Digital Records of Human Behavior *PNAS Proceedings of the National Academy of Sciences of the United States of America* **110** 5802-5805
[10] Kosinski M, Bachrach Y, Kohli P, Stillwell D J and Graepel T 2014 Manifestations of User Personality in Web Site Choice and Behaviour on Online Social Networks *Machine Learning* **95** 357-380
[11] Park G, Schwartz H A, Eichstaedt J C, Kern M L, Kosinski M, Stillwell D J, Ungar L H and Seligman M E P 2015 Automatic personality assessment through social media language *Journal of Personality and Social Psychology* **108(6)** 934-952
[12] Howlader P, Pal K K, Cuzzocrea A and Kumar S D M 2018 Predicting facebook-users' personality based on status and linguistic features via flexible regression analysis techniques *Proceedings of the 33rd Annual ACM Symposium on Applied Computing (SAC '18)* 339-345
[13] Matz S C and Netzer O 2017 Using Big Data as a window into consumers' psychology *Current Opinion in Behavioral Sciences* **18** 7-12
[14] Mikhaylov D V, Kozlov A P and Emelyanov G M 2017 An approach based on analysis of *n*-grams on links of words to extract the knowledge and relevant linguistic means on subject-oriented text sets *Computer Optics* **41(3)** 461-471
[15] Bolotova Yu A, Spitsyn V G, Osina P M 2017 A review of algorithms for text detection in images and videos *Computer Optics* **41(3)** 441-452
[16] Mikhaylov D V, Kozlov A P, Emelyanov G M 2016 Extraction of knowledge and relevant linguistic means with efficiency estimation for the formation of subject-oriented text sets *Computer Optics* **40(4)** 572-582 DOI: 10.18287/2412-6179-2017-41-3-461-471
[17] Jolliffe I T 2002 *Principal Component Analysis* (Springer) p 487
[18] Wasserman L 2005 *All of Statistics: A Concise Course in Statistical Inference* (Springer) p 442

Acknowledgments
This work was partially supported by the Ministry of education and science of the Russian Federation in the framework of the implementation of the Program of increasing the competitiveness of Samara University among the world's leading scientific and educational centers for 2013-2020 years; by the Russian Foundation for Basic Research grants (# 15-29-03823, # 16-41-630761, # 17-01-00972, # 18-37-00418), in the framework of the state task #0026-2018-0102 "Optoinformation technologies for obtaining and processing hyperspectral data".

The problem of fuzzy duplicate detection of large texts

E V Sharapova[1] and R V Sharapov[1]

[1]Vladimir State University, Orlovskaya street 23, Murom, Russia, 602264

Abstract. In the paper, considered the problem of fuzzy duplicate detection. There are given the basic approaches to detection of text duplicates–distance between strings, fuzzy search algorithms without indexing data, fuzzy search algorithms with indexing data. Thereview of existing methods for the fuzzy duplicate detection is given. The algorithm of fuzzy duplicate detection is present. The algorithm of fuzzy duplicate texts detection was implemented in the system AVTOR.NET. The paper presents the results of testing of the system.The use of filtering text, stemming and character replacement, allow the algorithm to found duplicates even in minor modified texts.

1. Introduction

Today computer technology and the World Wide Web have become part of our lives. Billions of users use the Internet every day. The Internet contains a huge number of documents. Many books are transformed in a digital form. Users can read books, articles, newspapers directly from computers and electronic gadgets. Today, in the Internet there are many studying materials: lectures, textbooks, methodical literature, etc. There are big collections of essays, course and diploma projects and scientific dissertations. It leads to problem un-controlling coping of information.

The text of document may be a part of another document, formed from several documents, may be modified part of another document (with a change of words to synonyms, endings, times, etc.). In this way, the document is fuzzy duplicate of other documents.

There is the problem of text checking to presence of fussy duplicates. The purpose of work is the algorithm creation of fuzzy duplicate detection.

2. Methods of the duplicate texts detection

2.1. The distance between strings

The task of comparing texts reduced to comparing their strings. For this reason, assume that the text document is aone string of a large length. There is a requirement to determine the measure of the string similarity. This measure is called the distance between strings.

The distance is the minimum number of string characters changes, which is necessary for the transformation of one string within another. The Hamming and Levenshtein algorithms used to determine the distance [1].

(a) The algorithm of Hamming [2].

The algorithm of Hamming is used to search the distance between strings of equal length, by using the operation "replacement". This algorithm is not suitable to search the distance between the strings, various by length.

(b) The algorithm of Levenshtein [3]

The algorithm of Levenshtein uses operations "replacement", "insert", "delete". They allow to search the distance between strings, different by length. But time of calculation of distance between strings is disproportionately increases with increases of strings size. Therefore, the use of this algorithm is only suitable for comparing multiple pages of documents.

(c) The algorithm of Damerau-Levenshtein [4].

This is modification of the Levenshtein distance. It is also consider the operation of "transposition" the next two letters.The method of calculation of Levenshtein distance was the basis for several algorithms of search for common subsequences.

(d) The algorithm of Wagner and Fisher [5].

This method is based on calculating the Levenshtein distance between the strings prefixes (substrings). The matrix of editorial prescription is made. It contains a summary value of Levenshtein distance (minimum weight operations to change characters). The size of editorial prescription matrix is $(p+1)\cdot(b+1)$, where p and b – compared strings prefixes.

The number of string comparisons is $k\cdot p\cdot b$, where k – coefficient (for natural language $k=0,2$). The complexity of algorithm is $O(p\cdot b)$ [6].This method is the easiest way to create of editorial prescription.

(e) The algorithm of Masek and Paterson [7].

The algorithm is modification of Wagner and Fisher algorithm with applying the approach Arlazorov, Diniz, Kronrod and Faradjev. The distance matrix in it is separated into submatrixes with edges computed in accordance with contiguous to their submatrixes.

The complexity of algorithm is $O(k\cdot(p\cdot b/\log(p)))$.

This algorithm faster than the algorithm of Wagner and Fischer, but, at the direction of the authors themselves [8], it reaches real speed only when comparing very long strings.

(f) The algorithm of Ukkonen [6].

For this algorithm it is required the construction of suffix tree of strings set to minimize search time in substring.

Complexity of algorithm is $O(k\cdot m)$, where m – maximum vertex depth of suffix tree.

Algorithm of Ukkonen is more applicable for search of exact matches strings. If searching for very different texts, then working time is greatly increased.

(g) The algorithm of Hirschberg [9].

The algorithm is modification of Wagner and Fisher algorithm. It is not calculated the distance between the strings but the distance between the longest common subsequence.

The complexity of algorithm is $O(k\cdot(p+b))$.

(h) The algorithm of Hunt and Szymanski [10].

It is based on searching of maximum increasing path on a matching elements graph of compared strings.

The complexity of algorithm is$O(k\cdot(g+b)\cdot\log(b))$, where g – the number of positions in which the string symbols are the equal.

Under certain circumstances, this algorithm shows good results, but in other cases, the comparison time is quadratic.

(i) The suffix tree [11].

This method is proposed of McCreight in 1976. It is similar to the Ukkonen algorithm, but the suffixes are added in reverse order.

The complexity of algorithm is $O(k\cdot m)$.

Disadvantage of algorithms working with suffix trees is use a fixed alphabet. It is suitable only for search of exact match strings.

2.2. Fuzzy search algorithms without indexing data

Let's consider fuzzy search algorithms without data indexing [12]. They are used to search for the previously unknown and unmanufactured texts.

(a) Linear search algorithm [13].

The algorithm uses the distance metric and applies it to the text words. Efficiency of the method depends on the number of errors and mismatches of texts. More numerous they are the more increases the comparison.

The algorithm complexity is $O(s \cdot p)$, where s – the number of errors made when checking, p – the text length.

(b) The Bitap algorithm [14].

This algorithm is applied more than the linear search. In view modifications it calculates the Levenshtein distance between words. At normal conditions, speeds of these two algorithms are the same. The algorithm complexity is $O(s \cdot p)$, but the speed of this algorithm significantly higher on long words than a linear method.

2.3. Fuzzy search algorithms with indexing data

It is necessary to index the text for the application of this group of algorithms. The dictionary is constructed by thetext. It contain words and their position in the text. On base of dictionary creates the index required for further search.

(a) The algorithm of sample expansion.

This algorithm is composedfor the search of many erroneous words and then look for them in the dictionary.

Complexity of the algorithm is $O((b \cdot |A|)^s \cdot b \cdot \log p)$, where A – the dictionary size, b and p – compared words, s – the number of errors in a word.

Disadvantage of the algorithm is that with increasing number of errors, the time of its operation also increase.

(b) The N-gram algorithm [15].

This algorithm is based on words separation to parts (substrings) with the length N (N-grams). It is compiled N-gram list, contain all the words of this text with occurrences of the substring. Next separated similarly inquiry is searched by this list.Most often used division into trigrams.

Complexity of algorithm is$O(w \cdot H)$, where w – the probability of occurrence of N-grams in the word, H – the number of words in the N-gram list.

(c) Hashing by signature [12].

This algorithm represents a word of the text in a code ("hash"), consisting of digits. Indexed words recorded in the table. Query also indexed is make the search in the table.

Complexity of algorithm is $O(|E|^s \cdot \frac{p}{2^{|E|}})$, where E – the hash function, representing the word as a code, p – the text length, s – the number of errors.

(d) The shingles algorithm [16].

The source text is separated into a set of sequences of a defined length (shingles). Documents are similar if the equal some amount of their shingles. Number of shingles is sufficiently large. Therefor are used methods reduce of shingles set.

For example, can be considered only those shingles, whose fingerprint is divided into a defined number n. Other way are selected shingles with the minimal fingerprint, or simply taken a defined amount of shingles.

(e) The algorithm of D. Fetterly [17].

For each document are computed 84 fingerprint by algorithm of Karp-Rabin using a random sample of shingles. These shingles are separated into 6 groups, which are called "super shingle". The combination of 2 super shingle is "mega shingle". 15 mega shingle are computed for the each document. Two documents are similar if they have the same at least one mega shingle.

Algorithm complexity is $O(S1 \cdot S2)$, where $S1$ – number of shingles from document 1, $S2$ – number of shingles from document 2.

(f) The I-Match algorithm [18].

Dictionary A is built for a group of documents. Dictionary consist of words with the average values of a code IDF. For each document is created dictionary B and determine the intersection of A and B.

Words of document, included in the intersection, are used to build I-Match signature of the document. If I-Match signatures are equals, then texts are similar.

(g) The algorithm of key words [19].

From the index are selected "key" words with threshold frequency. For a document constructed binary vector (signature). Value of i-th coordinate of the vector is equal to 1 if the frequency of i-th "key" word in the text is more than the threshold value, or 0 if its frequency in the text is less than the threshold. Texts are similar if they have the same signature.

*(h) The TF*IDF algorithm [20].*

It is a calculated inverse frequency of word in the collection of documents, called as Inverse Document Frequency (IDF). It is a calculated frequency of the word in the documents called as a Term Frequency (TF). A Values TF*IDF helps measure similarity documents to query. This is a very popular method. The disadvantage of this method is that it ignores a relative arrangement of the words in the documents.

3. The algorithm of fuzzy duplicate texts detection

At present, there are a large number of approaches for finding similar strings and small texts [21-25]. For large texts (from 1000 words), the search for coincidences is difficult due to the large time expenditure. Therefore, it is of interest to construct an algorithm for finding coincidences in large texts in a short time.

Let's consider our algorithm of fuzzy duplicate detection.

3.1. A character set conversion.

A checked text document can have different encodings. For a Russian language it is win1251, KOI8-R, UTF-8, ASCII. Therefore, the checked text document is converted to the uniform character set UTF-8.

3.2. Converting a text to the single long length string.

A text document is a set of strings. Formatting of the same documents may be very different. Therefore, it is necessary to simplify the presentation of the text. For this purpose, we used a linear representation of the text as a sequence of characters. In other words, the text appears as a single long length string (an wide string).

3.3. Pre-processing of a text.

It is perform pre-processing of a text: the text of a document is replaced by a filtered copy. For this purpose the following steps are performed:

- Removal of HTML tags.

A text document can to contain HTML tags. Because HTML tags are used for a document formatting, they do not affect to its contents. Therefore presence of HTML tags will just interfere to checking and must be removed from a document.

- Removal of punctuation marks and special characters.

In the text there are punctuation marks and special characters, including tabs, new line, etc. All this characters are removal from the document.

- Case conversion.

In the text of documents many words can be written with a lowercase or uppercase letters, consist entirely of uppercase letters or contain mixed uppercase and lowercase letters in words. Therefore, all characters are converted to lowercase.

- Processing of replacement characters.

Sometimes, in text documents there are replacement characters. For example, in Russian words several letters of the Russian alphabet replaced the similar Latin letters (such as "o", "a", "e", "y", "x", "c", "p", etc.). In this case, it is processed the replacement characters and they transform to an original form.

- Removal of stop words.

Stop words are common words occurring in an almost every document, such as "the", "of", "on", "a", "at", "by" etc. Therefore, these words do not help in the search of text duplicates, but increase the processing time. They must be removed from the document.

- Filtration of the text (not informative words removal).

A filtering of a text is a removal the most frequent words, not informative words etc. Also, filtering words that contain special characters, digits, words of great length, etc. This procedure allows to asignificantly reducing the amount of a computation (the length of the checked text).

- Stemming (processing of word endings).

A stemming is processed closure words. In this case, they are simply deleted. This avoids the effect of such modifications of the text, as a change in the singular and plural forms, masculine and feminine, a present and a past tense etc.

3.4. Detection of fuzzy duplicates for a document.

Our algorithm is based on a method of shingles. The text is divided into a sequence of words. A length of the sequence is fixed (5 words). For the each sequence of words is computed a MD5 code (shingles). The algorithm works with that code (figure 1).

A classical method of shingles used all words to create a sequence of words. We propose to consider not all the text of document but its a processed and filtered copy.

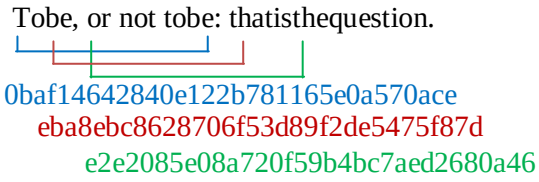


Figure 1. Dividing the text on the shingles.

The number of matching MD5 codes shows the measure of matching documents. If all codes in two documents are match, then documents are full duplicates. If there aren't match codes, then documents are different. If only a part of MD5 codes is match, documents are fuzzy duplicates.

4. The system structure

On the basis of the Vladimir State University, the authors developed the system of fuzzy duplicates detection. The system checks both the sources available on the Internet and on an internal database (databases of articles, course works and examinations etc.). The system generates a report with colouring duplicate parts of texts and viewing found duplicate sources. Let's consider the system structure [26] (figure 2).

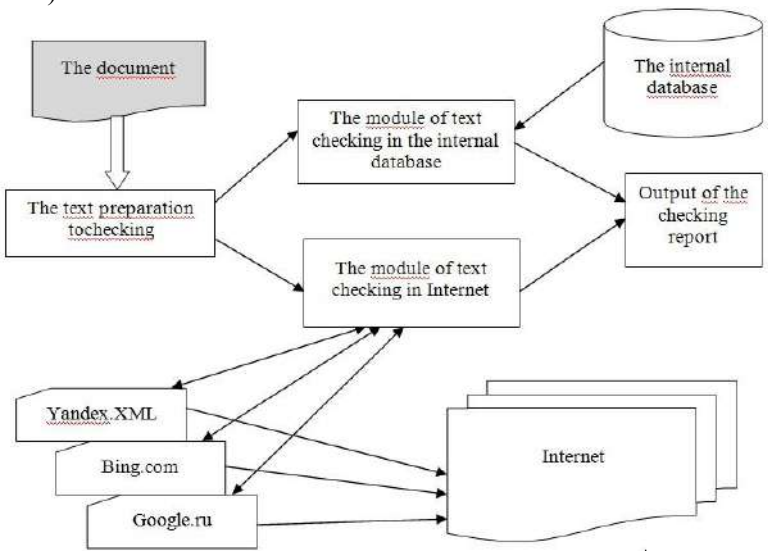


Figure 2. The structure of the system forfuzzy duplicates detection.

A text preparation includes removal of HTML tags, removal of punctuation marks and special characters, case conversion, processing of replacement characters, removal of stop words, filtration of the text, a stemming.

The system of fuzzy duplicates detection includes two modules. An each module operates independently of each other.

The first module checks the internal database sources. Base sources include a database of articles, course works and examinations, diploma projects, textbooks, lecture courses. Sources are stored in the form of a full-text and in the form of specially organized search index. A search index is needed for quick checking duplicates of a text in a database sources. There is no need for each check to a view all the available texts and make them fairly laborious process. All the necessary information for search is already included in a structured search index. A search index generated from the pre-processing text.

The second module checks the documents from the Internet. Text of the document is divided into informative parts. The number of parts depends on the size of the document. Next, using a search engine is searched for a sources containing these informative pieces. A search module use Yandex.XML, Yandex.com, Google.ru, Rambler.ru, Yahoo.com, Poisk.Mail.ru, Nigma.ru etc. Received documents are compared with the original document. To do this, the document format is determined (html-document, txt-file, doc-or rtf-document, pdf-file). From a html-document all markup tags are removed. The files *.doc, *.docx, *.rtf and *.pdf converted in a plain text format with no markup. In a next step documents are pre-processed and computed the similarity with the original document.

The main requirements for the system are the completeness and accuracy of duplicates detection. We did not set thetask of reducing a check time.

The algorithm of fuzzy duplicate texts detection was implemented in a system AVTOR.NET [27]. The system checks both the sources available on the Internet and on internal database (databases of articles, course works and examinations, etc.). The system generates a report with colouring duplicate parts of texts and viewing found duplicate sources.

5. Practical use

In 2016 the system Avtor.NET was used to check for a plagiarism of diploma projects full texts in department of technospheric safety of Vladimir State University. During the 10 days was checked more than 70 diploma projects. An each work has been from 60 to 120 pages. For this purpose, it was necessary to get up a new module to reduce the load on the search engines. A system Avtor.NET coped with the task. All diploma projects have been checked and we have received detailed reports on the presence of a plagiarism.

To test the efficiency of the algorithm we were compiled of three group of tests:

- Fussy duplicates of one document with a reordering of the sentence (T1);
- Fussy duplicates with a change in the text of the times and kinds of words (T2);
- Fussy duplicates of one document with a reordering of the sentence and the addition of the original text between sentences (T3);
- Fussy duplicates of some documents with a reordering of the sentence (T4).

All the tests had a size about 4000 characters and contained an average of 400 words. We used collection of essays, available on the Internet, to creating of tests. 10 tests of each groups was composed.

We compared the results of systems Antiplagiat, AdvegoPlagiatus, AntiPlagiarism.NET, Etxt.ru,Content-watch.ru, Text.Rucont.ru, Unicheck.comwith our algorithm (system Avtor.NET). To assess the quality of detection was used Recall, showing what percentage of duplicate text was really detected. Precision of all the systems was high and trends to 1.

All systems correctly handle test T1. Slightly worse result of Etxt.ru (Recall = 0.75) is due of features comparison algorithm implementation. With the test T2, the systems AdvegoPlagiatus, Text.Rucont.ru and Avtor.NET coped well. Worst results showed systems Unicheck.com (Recall = 0.32), AntiPlagiarism.NET (Recall = 0.58) and Antiplagiat (Recall = 0.63).

Table 1. Test results.

System	T1	T2	T3	T4
Antiplagiat	0.98	0.63	0.89	0.99
AdvegoPlagiatus	0.99	0.92	0.73	0.89
AntiPlagiarism.NET	0.83	0.58	0.72	0.90
Etxt.ru	0.75	0.83	0.65	0.49
Content-watch.ru	0.90	0.76	0.87	0.90
Text.Rucont.ru	0.94	0.95	0.94	0.91
Unicheck.com	0.95	0.32	0.74	0.96
Avtor.NET	0.99	0.97	0.94	0.99

Test T3 was well overcome by Text.Rucont.ru, Avtor.NET (Recall = 0.94), Antiplagiat (Recall = 0.89) and Content-watch.ru (0.87).

Test T4 badly passed only Etxt.ru (Recall = 0.49). All other systems showed very good results.

How can we see the system Avtor.NET coped well with all four tests. The algorithm copes well with the replacement of endings, permutations of pieces of text, compilation from several sources and the alteration of texts with the insertion of new words.

6. Conclusion

The system Avtor.NET correctly handle all texts and show results that are not inferior and sometimes exceed the results of existing systems.Thus, it was created the algorithm of fuzzy duplicates detection. The use of filtering text, stemming and character replacement, allow the algorithm to found duplicates even in minor modified texts.

7. References

[1] Sharapova E V 2014 Analysis of methods and systems for fuzzy duplicate detection *Proc. of 14 International multidisciplinary scientific Geoconference SGEM2014. Informatics, Geoinformatics and Remote Sensing* **1** 27-33

[2] Hamming R W 1950 Error detecting and error correcting codes *Bell System Technical Journal* **29** 147-160

[3] Levenshtein V I 1966 Binary codes capable of correcting deletions, insertions, and reversals *Soviet Physics Doklady* **10** 707-710

[4] Damerau F J 1964 A technique for computer detection and correction of spelling errors *Communications of the ACM* **7** 171-176

[5] Wagner R A and Fischer M J 1974 The string-to-string correction problem *Journal of the ACM* **21** 168-173

[6] Gusfield D 2003 *Strings, trees and sequences in the algorithms: Computer Science and Computational Biology* (Cambridge University Press) 654

[7] Masek W J and Paterson M S 1980 A faster algorithm for computing string-edit distances *Journal of Computer and Systems Sciences* **20** 18-31

[8] Masek W J and Paterson M S 1983 How to compute string-edit distances quickly *Time warps, string edits, and macromolecules: the theory and practice of sequence comparison* (Addison Wesley, Reading, MA) 337-349

[9] Hirschberg D S 1975 A linear space algorithm for computing maximal common subsequences *Communications of the ACM* **18** 341-343

[10] Hunt J W and Szymanski T G 1977 A fast algorithm for computing longest common subsequences *Communications of the ACM* **20** 350-353

[11] McCreight E M 1976 A space-economical suffix tree construction algorithms *Journal of the ACM* **23** 262-272

[12] Boytsov L M 2004 Classification and experimental study of the modern algorithms for fuzzy dictionary search *Proc. of 6-th Russian Scientific Conference Digital Libraries: Advanced Methods and Technologies, Digital Collections* 10
[13] Knuth D 1997 *The Art of Computer Programming* (Addison-Wesley) 396-408
[14] Baeza-Yates R and Navarro G 1996 A faster algorithm for approximate string matching *Proc. Combinatorial Pattern Matching* 1-23
[15] Nagao M and Mori S 1994 A New Method of N-gram Statistics for Large Number of n and Automatic Extraction of Words and Phrases from Large Text Data of Japanese *Proc.of the 15th International Conference on Computational Linguistics* 611-615
[16] Broder A 1998 On the resemblance and containment of documents *Compression and Complexity of Sequences SEQUENCES'97* 21-29
[17] Fetterly D, Manasse M and Najork M 2003 On the Evolution of Clusters of Near-Duplicate Web Pages *Proc. of the First Conference on Latin American Web Congress* 37-45
[18] Kolcz A, Chowdhury A and Alspector J 2004 Improved Robustness of Signature-Based Near-Replica Detection via Lexicon Randomization *Proc. of KDD* 605-610
[19] Ilyinsky S, Kuzmin M, Melkov A and Segalovich I 2002 An efficient method to detect duplicates of Web documents with the use of inverted index *Proc. of WWW Conference*
[20] Salton G and McGill M J 1983 *Introduction to modern information retrieval* (New York: McGraw-Hill) 448
[21] Mikhaylov D V, Kozlov A P and Emelyanov G M 2015 An approach based on TF-IDF metrics to extract the knowledge and relevant linguistic means on subject-oriented text sets *Computer Optics* **39(3)** 429-438 DOI: 10.18287/0134-2452-2015-39-3-429-438
[22] Mikhaylov D V, Kozlov A P and Emelyanov G M 2016 Extraction of knowledge and relevant linguistic means with efficiency estimation for the formation of subject-oriented text sets *Computer Optics* **40(4)** 572-582 DOI: 10.18287/2412-6179-2016-40-4-572-582
[23] Evdokimova N I and Kuznetsov A V 2017 Local patterns in the copy-move detection problem solution *Computer Optics* **41(1)** 79-87 DOI: 10.18287/2412-6179-2017-41-1-79-87
[24] Mikhaylov D V, Kozlov A P and Emelyanov G M 2017 An approach based on analysis of n-grams on links of words to extract the knowledge and relevant linguistic means on subject-oriented text sets *Computer Optics* **41(3)** 461-471 DOI: 10.18287/2412-6179-2017-41-3-461-471
[25] Sharapova E V 2014 One way to fuzzy duplicates detection *Proc. of 14 International multidisciplinary scientific Geoconference Informatics, Geoinformatics and Remote Sensing* **1** 273-277
[26] Zelenkov Y G and Segalovich I V Comparative analysis of methods for fuzzy duplicate detection for Web-documents *Proc. of 9-th Russian Scientific Conference Digital Libraries: Advanced Methods and Technologies, Digital Collections* 9
[27] Sharapova E V 2014 System of fuzzy duplicates detection *Applied Mechanics and Materials* **490-491** 1503-1507

Applying process mining techniques and neural networks to creating and assessment of business process models

K Grigorova[1], K Mironov[1] and E Y Malysheva[2]

[1]University of Ruse, Studentska street 8, Ruse, Bulgaria, 7017
[2]Volga Region State University of Services, Gagarina street 4, Togliatti, Russia, 445677

Abstract. The article presents an approach for automated generation of business process models by applying process mining techniques to event logs created during the operation of information systems used in an organization. Existing algorithms for process mining are discussed. Criteria for performing a comparative analysis of these algorithms are specified. A framework is proposed in which to build and analyze business process models. The framework includes tools for initial analysis of the event log file, extracting elements of a business process model, and composing a new model by applying a trained neural network.

1. Introduction

The rapid development of technology coupled with the proven benefits of implementing a process-oriented management approach are the causes of the growing use of business process models in organizational decision-making. Enhanced efficiency, better resource monitoring, and enhanced competitiveness are one of the reasons why business process management is increasingly involved into medium and small business organizations. Modern marketing methods, combined with software solutions that support and facilitate them, convey the understanding of business beyond the organization.

The need for businesses to get more information on how their processes are performed in a real environment is one of the main reasons for the spreading and use of process mining [1]. Modern business management systems and large volumes of data they generate create a favorable environment for the application of various process mining methods.

In addition to applying the most common algorithms and generating new business process models, analyzing system information allows to validate the integrity of the overall process, as well as to provide valuable metrics about its performance [2]. Another important feature of process mining is the ability to be used as a part of methodology for automating life cycle stages of business processes.

This article aims to present an experimental framework that uses process mining methods and enables automated generating and performance evaluating of identical business process models. A multilayer neural network is used for decision making during the business process generation as well as process evaluation according to the specific use case. In the previous related works the use of artificial neural networks is discussed as a tool for extracting business process models from event logs. The framework described in this paper involves neural network that is used on the stage of decomposing and evaluation of generated process model.

2. Process mining algorithms

The creation of a methodology for retrieving information from system data requires the knowledge of modification and application of the basic approaches of process mining [3]. Following is a review of the most commonly used algorithms for retrieving process models from system data, as well as analysis and classification based on their core characteristics and summarizing the most appropriate algorithms for specific use cases.

2.1. Alpha algorithm

The Alpha algorithm is one of the main in process mining [4]. It is characterized by creating a complete business process model by analyzing and processing event logs. The algorithm works by detecting Workflow nets and merging them into a complete logical structure. Typical for the algorithm is its fast performance, but it is also characterized by some limitations. The presence of a complete log file is required for the correctness of the generated business process model through Alpha miner. Incompleteness or lack of events, as well as describing of some exceptions in the event logs, may lead to incorrect final results. The Alpha algorithm is suitable for use in conjunction with systems that generate complete and well-structured event log files.

2.2. Heuristics miner

This algorithm is considered to be an extended version of the Alpha algorithm, but provides additional options for analyzing system event logs. The application of Heuristics miner is characterized by analyzing log data, sequencing events and determining the occurrence frequency. Based on these data, a network is created that consists of linked sequences and reflects the basic behavior of the workflow. Unlike the Alpha algorithm, Heuristics miner can be used for processing event logs with missing, incomplete and unassigned data, and even there are records of "emergency" situations.

2.3. Genetic miner

The genetic algorithm uses an evolutionary selection approach that mimics the process of natural evolution. This algorithm is significantly slower but offers optimal results based on system event log files [6]. It is characterized by analyzing and identifying initial "process candidates" in the form of related events, most often represented by Petri nets. Genetic miner has a cyclical character, and each of the potential processes is compared, and the one that comes closest to the general scenario in the event log is chosen. The genetic algorithm is suitable for use in system logs with missing or incomplete data, as well as with repeating events. It is not suitable for real-time application, and the final result is a detailed and very accurate business process model.

2.4. Fuzzy miner

Very often, business process models that are generated by event log analysis are large, unstructured and chaotic, without the ability to accentuate important events. The fuzzy algorithm is configurable and allows compact representation of detected patterns from different viewpoints (user or system) [7]. This is done through a set of techniques for identifying and removing isolated nodes, as well as merging several linked nodes into one. As a result of Fuzzy miner performance a logically true and identical copy of the original unstructured process is obtained, which is easy readable [5]. It is suitable for application on complete system log files without missing or incomplete information.

2.5. Selection of algorithm comparison criteria

When comparing algorithms for Process mining in the context of automated creation and assessment of business process models, consideration should be given to criteria that affect the program implementation as well as the number of characteristics of the input and output data. For this purpose, event logs and generated processes are classified according to the following criteria:

- Integrity - this characteristic indicates the ability of the algorithm to work in the case of missing event data.
- Uniqueness – indicates the ability of the algorithm to handle repeating events within a single process.

- Exceptions handling - indicates whether the algorithm can work normally when there are some exceptions.
- Real-time - indicates the ability of the algorithm to be used in real-time on a running system.
- Model integrity - indicates whether the generated model is a complete business process model or only a part of it (a series of related tasks).
- Abstraction level - indicates whether the obtained business process model is generalized and simplified.
- Structured model - indicates whether the obtained business process model is structured and suitable for reading by a human.

2.6. Comparative analysis of process mining algorithms

Automated generation and evaluation of business process models requires the processing event logs that differ by type and structure and the creation of a unified method for working –reading and analyzing the event log files. In order to determine the most appropriate algorithms for different use cases, process mining algorithms are compared and analyzed on criteria affecting their input and output data.

Using the specified criteria and taking into account the characteristics of process mining algorithms discussed in the previous sections, the results of the comparative analysis are summarized and presented in figure 1.

	Alpha Algorithm	Heuristics Miner	Genetic Miner	Fuzzy Miner
Integrity		✓	✓	
Uniqueness	✓	✓	✓	✓
Exceptions		✓	✓	
Real-Time	✓			
Model Integrity	✓	✓	✓	✓
Abstraction level				✓
Structured model				✓

Figure 1. Results from comparative analysis.

In the context of the framework created, when generating a business process model from an event log file, a preliminary analysis of the input data should be performed in order to determine which of the algorithms considered is the most appropriate for the specific case.

3. Framework for automated creation of business process model and evaluation of its performance

The automated creation of business process models combines the optimization and reengineering of existing models. Based on the fact that an activity can be implemented in several different ways, the purpose of the framework is to assemble different business process models that perform a particular business task. Components of each business process are already available business tasks extracted from system event log files. Each business process model will be evaluated for performance in order to identify the most effective for the considered business task.

Figure 2 shows a general architecture of the proposed framework. Each stage is considered and its main features are presented.

3.1. Pre-analysis

At this stage, a preliminary analysis of the system event log file to be processed is performed. The purpose is to determine its integrity and type. Based on the results of the preliminary analysis as well as those from the comparative analysis, an optimal process mining algorithm is selected to be applied in the specific case.

3.2. Extracting business process model

The selected process mining algorithm is applied to the event log file and a business process model is generated. In order to avoid future errors, the model is validated and its integrity is checked. In case

that a correct process is retrieved, it is simulated and evaluated for performance. These data will be used in the future to determine the most efficient process.

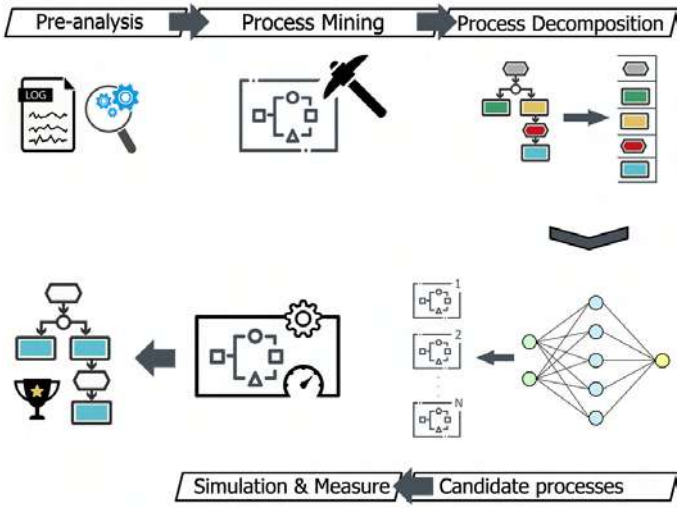


Figure 2. Framework's architecture.

3.3. Business process decomposition

The decomposition consists of segmenting the separate components of the model and storing them as independent elements. Intra-system storage of the elements allows the use of non relational model and, depending on the amount of business components, the application of Big data methods is recommended.

3.4. Generation of "candidate" processes

Creating processes that will be simulated and evaluated is performed using the decomposed business elements from the previous step and with the help of a neural network. Depending on the elements found in the decomposition of the model, the number of layers and neurons in the neural network may vary. The neural network training is based on the dependencies found in the log file, the relationships between the individual components of the process as well as the initial design and the subsequent redesign of the processes.

As a result of this phase, there are several logically identical "candidate" processes that perform the same functionality as the primary process generated by applying process mining algorithm.

3.5. Simulation and assessment

At this stage, it is intended to identify the most appropriate business process model for the particular business task. For this purpose, simulation is performed on each business process in a controlled and preconfigured work environment. During the execution of each process, its behavior is monitored and data is collected that will help to generate baseline assessment of performance.

3.6. Framework operation

After selecting a system event log file, which is the input data of the developed framework, a preliminary analysis is performed on it that aims to determine its structure, integrity, and availability of information about recurring events. After the analysis and on the basis of its results, a process mining algorithm is chosen which is most appropriate for the specific case. The selected algorithm is applied to the input log data and a business process model is derived. It is presented in XML-based format that allows easy processing. Additionally, metadata about the process logic is created used for initial neural network training.

The next step is decomposing the resulting business process model of components and storing them in a data repository, created as Redis database. Next the separate components are used to create "candidate" processes by using a neural network. The number of layers and neurons in the network is

determined dynamically depending on the case and the business process model under consideration. As an environment Java neural network framework Neuroph is used.

The generated "candidate" processes are simulated, taking into account their productivity metrics during the simulation [10]. After comparing the results obtained, the most effective business process model for the particular business task can be determined.

4. Creating an environment for automated generating and evaluating the performance of business process models

The presented framework for evaluating and generating business process models using Process Mining and Artificial Intelligence techniques hypothetically presents the possibilities for automated selection of the most effective business process models for a particular business task. The purpose of this experiment is to confirm the possibility of applying the developed framework, as well as to demonstrate in detail the preparation and implementation of the individual steps of the created algorithm.

4.1. Scope of the experiment

Creating an environment for automated generating and evaluating the performance of business process models covers the stages of decomposing, generating "candidate" processes, simulating and evaluating newly created models. Activities related to preliminary log file analysis and subsequent application of process mining algorithm are replaced by the use of real business processes. This ensures the applicability of the proposed framework in a real business environment.

For the purpose of the experiment, 4,000 business process models from BPMN (versions 1.1 and 2.0) and EPC were selected on a random basis. The processes are provided by the BPM Academic Initiative and are developed by various business and academic organizations.

4.2. Method of performance and expected results

Running the experiment follows the main steps of the proposed algorithm. Designed for automation, the algorithm allows many of the related activities to be automated through program implementation. A part of these processes are the validation, decomposition and storage of business process models. For the use of neural network decision-making, initial training and initialization of threshold values for individual neurons are envisaged. Simulating the generated business process models in a controlled environment allows tracking a number of their features that are taken into account when determining the most effective process for the particular task. Figure 3 shows the stages of the experiment, the basic software components used in each step, the input data used and the results obtained.

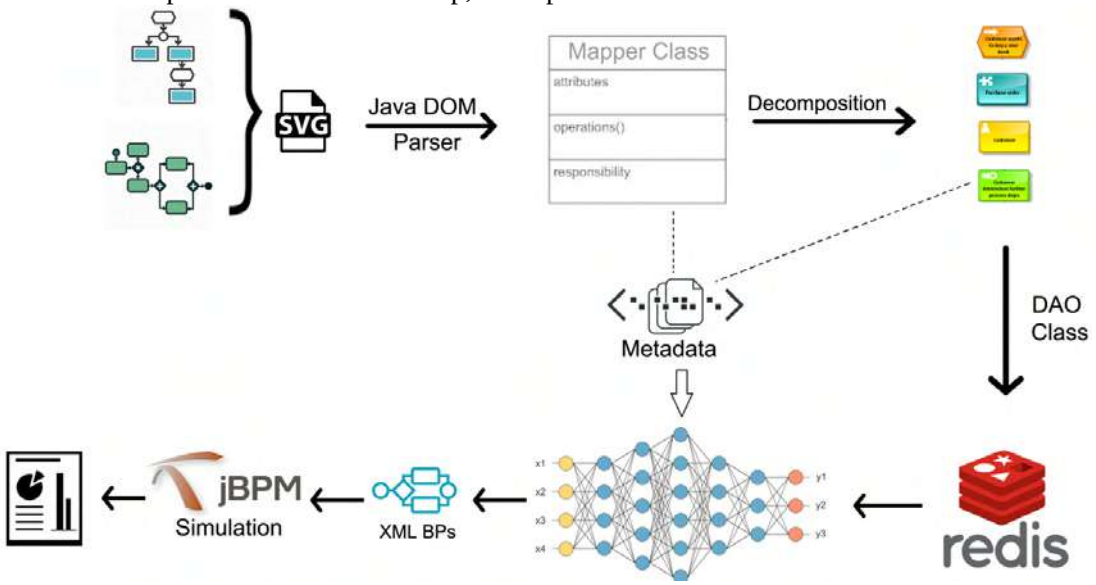


Figure 3. Experiment environment.

4.3. Pre-treatment and preparation of the BP for the experiment

The models used in the experiment need to be pre-treated to make them convenient. The initial state of the models in the database is in the form of vector files (SVG). The XML-based character of SVG files allows information to be extracted through a script using an XML-based software library. Within the experiment, this is done through the Java DOM Parser.

Intra-system presentation of information requires the creation of a Mapper class that allows the extracted data for each process to be represented as a Java object. The class contains a number of parameters, most commonly represented as arrays and collections of classes, designed to store the data specific to each model. Figure 4 shows the structure of the Mapper class.

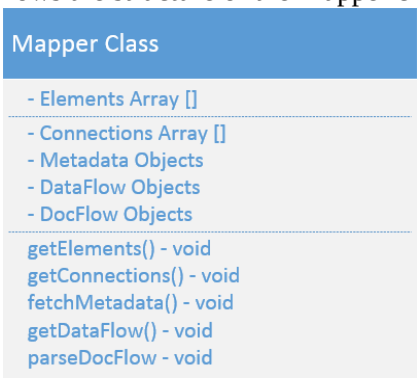


Figure 4. Mapper class structure.

As a result of the pre-processing of the models, an array that is composed of universal class instances is obtained. Each instance stores the transformed XML information for a business process. In this way, business data needed for the experiment is interpreted to an information array independent of file extensions and ready for the next step of the framework.

4.4. Decompiling of business processes

This stage of the developed framework aims to analyze each business process and segment and preserve its main elements. During the analysis, data about the model itself is derived. Later it serves as control results when compared to the generated "candidate" processes as well as is used to generate a data set used for primary training of the neural network. Before starting the decomposition process, a file containing basic metadata about the model is generated for each business process. Figure 5 shows the experimental structure of a JSON file storing metadata about a business process model.

The actual decomposition of the models is done by extracting all the data from the Mapper classes and presenting them as individual elements. Each element is analyzed and a meta model is generated that contains the data about it. Following is the classification and storing the element in a database. The generator automatically creates unique identifiers for each element and puts them in a "key-value" format, so the information is ready to be recorded into the database. The information presented in this way is submitted to the DAO classes that perform the actual storage of the decomposed items in a database. The experiment is conducted using the nonrelational Redis database, which allows fast storage and retrieval of the large volumes of business process models that are provided for the experiment.

4.5. Generation of "candidate" processes

Generating "candidate" processes involves building a logically identical business process to the original through the use of decomposed elements. For this purpose, metadata derived from business process analysis before decomposition, as well as the generated for each element, are used. From the complete meta model, the logic of the initial process is derived and the layers of the neural network are determined. Individual neurons are also determined on the basis of individual metadata for each element [11]. Within the experiment, the Java Neural Network framework - Neuroph is used for realization of this functionality. Figure 6 presents an example of a neural network and possible scenarios that individual neurons create.

```
{
    "Standart": "BPMN",
    "elements": {
        "objects": [
            {
                "ID": 3,
                "type": "Task",
                "title": "Analyze order",
                "metadata": [
                    {
                        "mId": 1,
                        "metadata": "status - pending"
                    }
                ]
            },
            {
                "ID": 4,
                "type": "Parallel Gateway",
                "title": "Concurrency",
                "metadata": [
                    {}
                ]
            }
        ],
        "connections": [
            {
                "from": 3,
                "to": 4,
                "text": "",
                "type": ""
            }
        ]
    }
}
```


Figure 5. Metadata about business process model.

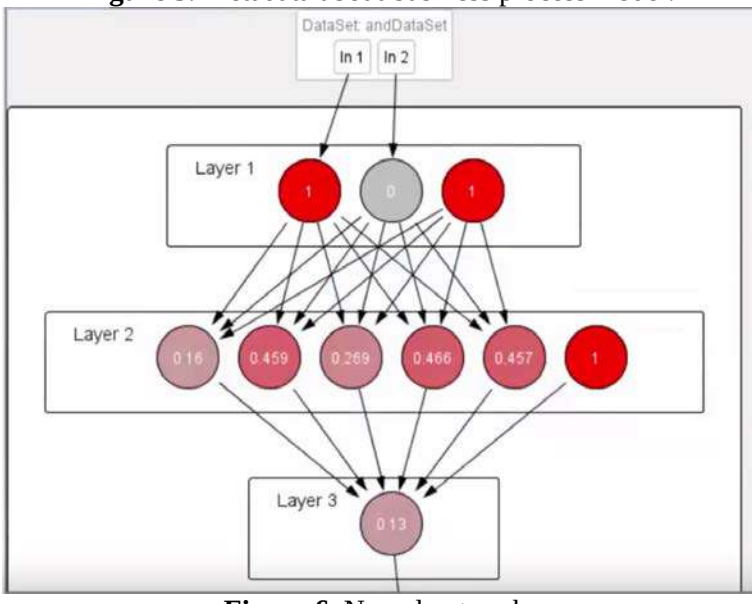


Figure 6. Neural network.

As a result of this stage, the generation of several business process models is expected, logically identical to the original BP model. Candidate processes are stored in an XML-based file that is used in process simulation.

4.6. Simulation and evaluation

This stage of the developed framework aims to determine the availability of more efficient models equivalent to the original one. This is done by simulating the original model and the generated "candidate" processes. The simulation is performed under the same conditions and the basic

parameters accompanying it are traced. For the purpose of the experiment, the Java-based Business Process Management framework - jBPM is used. The generated XML files of the "candidate" processes are loaded into the process server that will perform the simulation. Figure 7 and figure 8 show a view of the simulation and the resulting business process model within the jBPM framework.

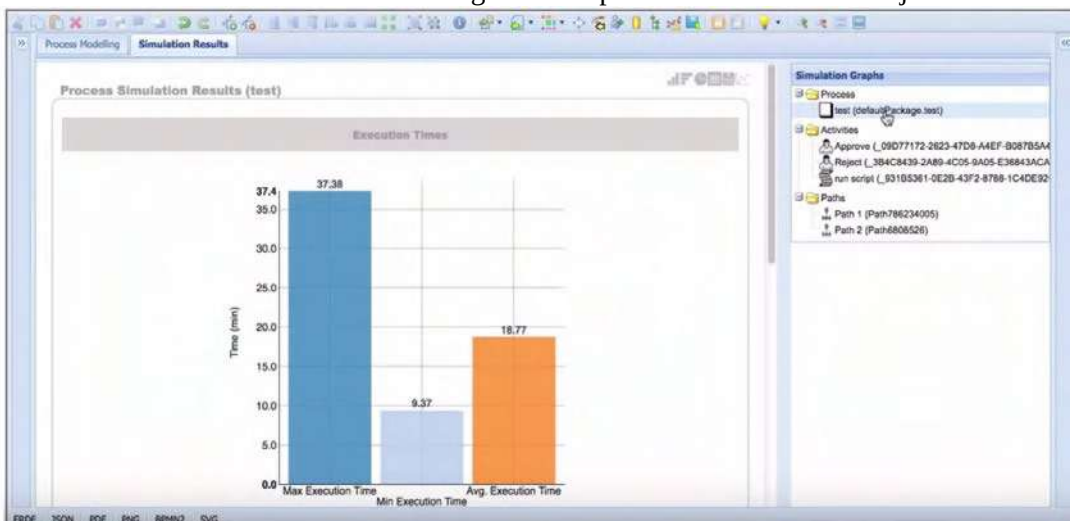


Figure 7. The results of business process simulation.

jBPM offers tracking of over 20 metrics related to process execution and generates a report of the results. By comparing the data about the initial process and those about the "candidate" processes, one can determine whether a more efficient model is found. Depending on the process requirements, performance parameters can be filtered and different performance criteria defined, according to the shortest implementation times, depending on the resources used, risk reduction, and more [8], [9].

The examined experiment shows the possibilities for practical application of the created framework for automated generation and evaluation of the efficiency of business process models. By combining the means of Process Mining and the capabilities of Artificial Intelligence, it is possible to create an environment for optimizing business processes and improving business performance.

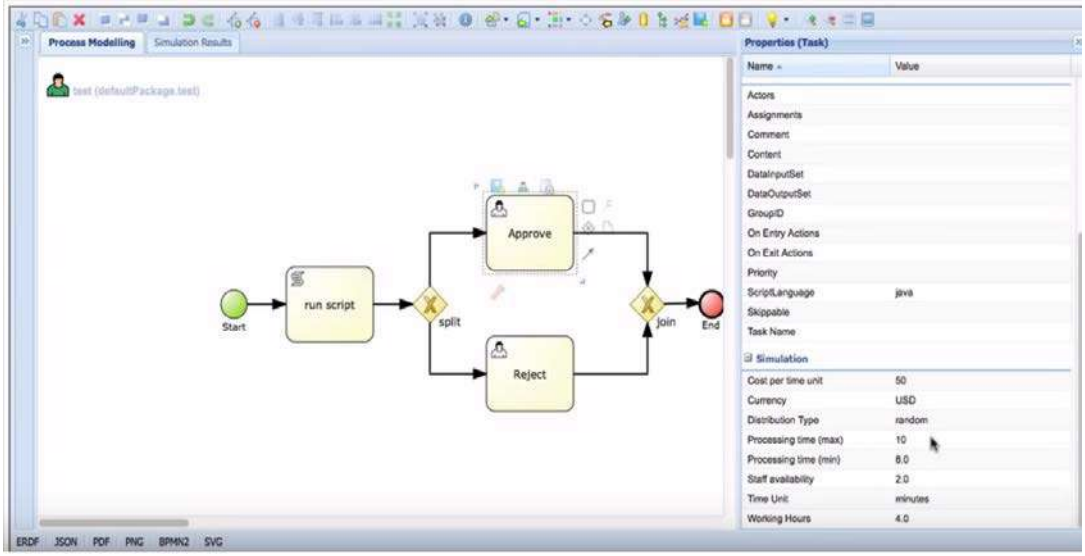


Figure 8. Resulting business process model.

5. Conclusion

The development of technology and increasing competitiveness in business environments contributes to the partial or complete automation of business tasks. The variety of business process modeling systems and the growing tendency for their use by business creates favorable conditions for the

application of process mining techniques. This article offers a solution for automated generating and evaluating the performance of business process models.

The presented framework uses process mining and artificial intelligence techniques and aims to automate part of the monitoring and optimization activities. As a final product of the examined framework, a business process model is generated that is the most effective for a particular business task. This increases the organization's competitiveness, and the automation reduces the resources used.

6. References

[1] Wang J, Wong R K, Ding J, Guo Q and Wen L 2013 *IEEE Transactions on Services Computing* **6(4)** 484-496
[2] van der Aalst W M, Reijers H A, Weijters A J, van Dongen B F, De Medeiros A A, Song M and Verbeek H M W 2007 *Information Systems* **32(5)** 713-732
[3] Tiwari A, Turner C J and Majeed B 2008 *Business Process Management Journal* **14(1)** 5-22
[4] van der Aalst W M, Weijters A J and Maruster L 2004 *IEEE Transactions on Knowledge and Data Engineering* **16(9)** 1128-1142
[5] Weijters A J, van Der Aalst W M and De Medeiros A A 2006 *Technische Universiteit Eindhoven, Tech. Rep. WP* **166** 1-34
[6] De Medeiros A A and Weijters A J 2005 *Lecture Notes in Computer Science* **3536**
[7] van Dongen B F and Adriansyah A 2009 *Business Process Management Workshops* **43** 158-169
[8] Maita A R C, Martins L C, López Paz C R, Peres S M and Fantinato M 2015 *Business Process Management Journal* **21(6)** 1391-1415
[9] Hsiao B, Shu L, Young M and Yang H C 2016 *PACIS* p 264
[10] Hornix P T 2007 *Performance analysis of business processes through process mining. Master Thesis, Eindhoven University of Technology*
[11] Kopenkov V N and Myasnikov V V 2016 *Computer Optics* **40(5)** 713-720 DOI: 10.18287/2412-6179-2016-40-5-713-720

Acknowledgements
This work is supported by the National Scientific Research Fund under the contract DFNI - I02/13.

The ontology-driven approach to support the requirements engineering process in Scrum framework

M Sh Murtazina[1] and T V Avdeenko[1]

[1]Novosibirsk State Technical University, Karla Marks ave 20, Novosibirsk, Russia, 630073

Abstract. The paper presents an approach to the Support of the requirements engineering process in the field of software development process by applying OWL ontology. A brief overview of the capabilities of using ontologies for intellectual support of the process of requirements engineering is given. The main features of requirements engineering for Scrum project management of software development are analyzed.The quality assessment criteria of user stories are explored. The developed ontology accumulates knowledge about the quality assessment criteria of user stories, about requirements artifacts, types of requirements, about elements of Scrum framework. The ontology includes axioms that determine the quality of user story expression and the quality properties of the requirements.Also ontology includes axioms determining the priority and risk of user stories.The ontology is implemented in the Protégé environment.

1. Introduction

The International Standard ISO/IEC/IEEE 29148-2011 *Systems and software engineering – life cycle processes – requirements engineering*determines requirements engineering as "interdisciplinary function that mediates between the domains of the acquirer and supplier to establish and maintain the requirements to be met by the system, software or service of interest" [1, p.8]. A complex system of requirements is developed and maintained in the process of requirements engineering. This system of requirements is built on the basis of needs, expectations, constraints and interactions of the stakeholders. Under the stakeholders the Standard ISO/IEC/IEEE 29148-2011 primarily implies users and customers.The circle of stakeholders also includes software developers and suppliers.

Requirements engineering plays a key role in ensuring the success of the software development. Eliciting and managing requirements is an extremely difficult task for any methodology of software development project management. Project management methodologies are usually subdivided into rigid and agile approaches. A rigid approach to the management of the software development project is characterized by detailed planning. Requirements engineering here is a separate stage that precedes all other stages of the software building. In case of agile methodologies planning is performed only for the current iteration. Requirements engineering here is an iterative process where the requirements constantly evolve. Changeability of requirements is a serious problem for software development projects. In this regard, rigid methodologies are much inferior to agile ones. With an agile approach, precious time is not wasted on trying to anticipate all possible requirements and document them in detail. A typical form of agile high-level requirements in agile methodologies is feature requests and user stories. They are formulated as one or more sentences which illustrate the user's goals that a

software function will satisfy. The details are clarified when the requirements are implemented within the regular iteration. This happens in the process of active interaction with the stakeholders. Incompleteness that is inherent to requirements specification in agile methodologies is compensated by extensive informal communication with stakeholders. Another problem area is to ensure the consistency of the requirements coming from different stakeholders and also to reveal the same requirements presented by different stakeholders with different terminology.

Scrum is one of the most popular and well-developed agile methodologies. According to the Internet survey of Agile Survey, only in 2017 this methodology was used by 56% of the respondents' organizations[2]. Despite the immense popularity of agile methodologies in general, and the Scrum methodology in particular, development of approaches to assessing the quality of specifications for agile requirements is still relevant. Many existing investigations use the INVEST model proposed in 2003 by B. Wake [3]. New approaches to assessing the quality of requirement specifications for agile methodologies began to appear in the mid-2010s. Examples of new approaches are Quality User Story Framework [4] and Agile Requirements Quality Framework [5].

The trend of recent years in the field of supporting the process of requirements engineering and assessment of the software requirements specification quality is knowledge-based systems built on the ontologies. The ontological approach is promising for overcoming some of the deficiencies in requirements engineering [6]. However the existing ontology-base solutions are more focused on assessing the requirements specifications in accordance with international standards and do not take into account the specifics of Scrum and the evaluation of the quality of user stories.

In the present paper we propose an ontology-driven approach to support the requirements engineering process in Scrum framework. The paper is organized as follows. In Section II the literature review is made. Section III analyzes the features of the requirements engine in Scrum. Section IVdiscusses the issues of assessing the quality of user stories.Section V proposes an ontological model for supporting the process of requirements engineering in Scrum. In section VI conclusionsabout prospects for the proposed approach are made.

2. Background

In the early 2000s K. K. Breitman et al. considered the ontology development process as a subprocess of the process of requirements engineering [7]. In [8] it was proposed to use the domain ontology as an infrastructure for refinement of software requirements. G. Dobson et al. pointed out that ontologies are useful for the representation and interconnection of many types of knowledge. Requirements engineering involves the extraction of knowledge from a variety of sources. Ontologies in requirements engineering can be used for representation of the requirements model itself, as well as acquisition structures for domain knowledge, the application domain and the environment [9].

One of the first works, in full demonstrating the possibility of the ontologies for the Requirements engineering, is the work of M. Kossmann et al. in which semi-automated methodology OntoREM (Ontology-driven Requirements Engineering Methodology) was developed [10]. Methodology OntoREM includes processes, methods and tools. The aim of this methodology is to create requirements specifications for systems in less time and at lower costs while improving the quality of such specifications. Methodology OntoREM was applied by its authors in the company Airbus to develop aircraft operability requirements. Using OntoREM resulted in significant savings of money and time [11]. The structure of the OntoREM process developed by M. Kossmann et al. is shown in Figure 1.

To provide the OntoREM process, Kossmann M. et al. used the MindManager Tool Environment, the OntoRAT Tool Environment, the Protégé Tool Environment, and the DOORS Tool Environment. First, the domain ontology intheformofaMindmap is developed using in MindManager. This tool is convenient for quick visualization of the domain ontology and allows saving the result in the OWL format. The Protégé environment is used to manage the ontologies obtained with MindManager. The OntoRAT tool (Ontology-driven Requirements Analysis Tool) is used to analyze requirements by status, purpose, soft targets (i.e. targets without clear criteria) and traceability. The OntoRAT tool was developed during the OntoREM project. The IBM Rational DOORS software package is used for requirements management. The OntoREM project has, to a large extent, led to an active analysis of the

possibilities of using of ontologies in many subject domains including the field of software development.

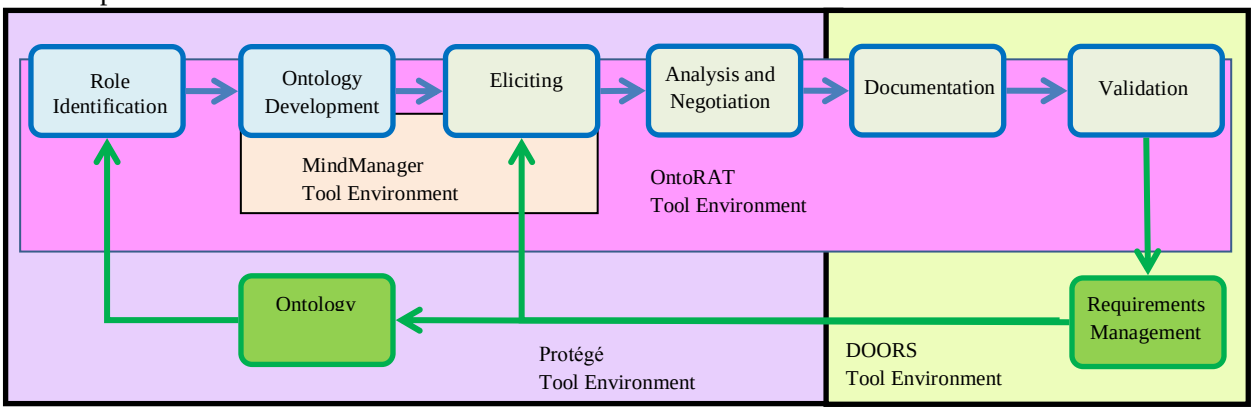


Figure 1. OntoREMprocess [10].

In paper [6] an approach to automating the process for quality evaluation of requirements is presented. B K. Siegemund has distinguished two types of support for the requirements engineer: "(1) support for the specification of the requirements knowledge and (2) validation and error elimination support" [6, p.79]. Support for requirement engineering is provided by two ontologies: "Guidance Ontology" and "Requirements Ontology". The GORE (Goal-Oriented Requirements Engineering) method is used in the approach developed by K. Siegemund. A set of optional and mandatory tasks based on validation rules and their pre- and postconditions are contained in the Guidance Ontology. The basis of the Requirements Ontology is the IEEE 830-1998 standard. The Requirements Ontology accumulates knowledge about the requirements for a particular project

The approach based on frame ontology is proposed in [12-14]. Application of the ontology allows building a harmonized model of requirements for software development process. This is designed to help the analyst to take into account all aspects of the requirements. The theory of the field structure of speech parts and recommendations of the SWEBOK were applied to construct the model. Templates of the specification structure were designed in the Protégé environment.

A metamodel is proposed in [15] which is used to reason about the requirements. This model is implemented as an OWL ontology. The model includes requirements, requirement artifacts and stakeholders. The requirements are linked by four types of relations: Refines, Requires, Conflicts, and Contains. These relations allow building rules to reason about requirements traceability, consistency and completeness.

An ontological approach to improving the requirements engineering in the agile development is presented in [16]. C. Thamrongchote et al. pointed out that although the templates of user stories are easy to use, the application of the domain terminology when writing them is a difficult problem. It is proposed in [16] to accumulate user stories from previous successful projects in the ontology knowledge base, in order to improve the process of subsequent working with them. The ontology schema is designed using class and hierarchy relations. Hierarchy relations are used, for example, to determine which user roles inherit features from other roles. For example, the class "Guest" is a subclass of the "Customer" class. Accordingly, the set of user stories of the "Guest" should be implemented for the "Customer". Also ontology establishes synonymous relations. For example, a synonymous relation exists between the "Guest" and "Visitor" classes. This means that user stories that mention the roles "Guest" and "Visitor" describe the features for users of the same class. The words extracted from the user story are distributed in three classes: "Role", "Action" and "Object". Individuals of the classes "Role" and "Action" are associated with the relation "perform Action" and individuals of the classes "Action" and "Object" – relation "perform Object".

The authors of paper [17] propose to use the ontology to predict the effort estimates in the implementation of user stories. The accuracy of effort estimates depends on the level of analysis of the software development context and the experience of persons making the evaluation. The approach

developed by M. Adnan et al. Proposes to accumulate unique knowledge of key project participants. This knowledge is then used to evaluate efforts to implement user stories.

The analysis of scientific publications shows that the questions of designing the ontology-driven approach to support the requirements engineering process in the Scrum framework are not yet sufficiently developed. The considered complex ontology-oriented approaches do not take into account the peculiarities of working with user stories. At the same time, research in the field of requirements engineering in agile projects focuses on certain aspects of working with user stories and evaluation of their quality.

3. Requirements engineering in Scrum framework

According to the Guide [18], Scrum is a process framework designed for developing and sustaining complex products. The core of the Scrum is *Sprint*, that is a time interval of one month or less, during which the Scrum team creates a potentially releasable product *Increment*. The basic elements of the Scrum framework are depicted in Figure 2. These include the Scrum Team, the Scrum Events and the Scrum Artifacts.

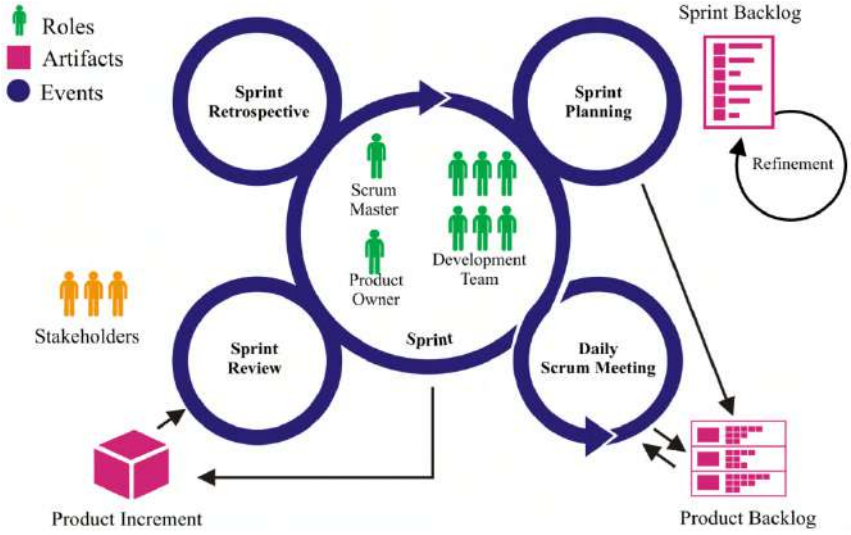


Figure 2. The basic elements of the Scrum framework.

Requirements engineering in Scrum is an iterative process. The requirements evolve in each Sprint [19, p.27]. Eliciting the requirements-needs occurs during the Sprint Review. The next stage is a requirements analysis accepted for execution in the Sprint. From the point of view of the Scrum Team the requirements should be unambiguous, complete and consistent. Conflicts can be resolved by prioritizing user stories or overriding incorrectly formulated requirements.

Documenting the requirements helps to analyze and verify the requirements specification. Although documentation in agile software development methodologies is used in smaller volumes than in rigid methodologies, requirements documentation is an essential part of the work process. Allocate the following base forms of requirements records: feature requests and stories. The following forms of recording the requirements in agile methodologies are allocated: feature request and story. Feature request is a structured query (with header, description and set of attributes) for new or refined feature of the software. Story is a high-level requirement formulated as one or more sentences in the user's everyday or business language so that the developers can give a reasonable estimate of an effort to implement it. The following types of stories are distinguished by the form of the record:

- userstory;
- technical story / technical user story;
- jobstory.

The most common form to record the requirements when applying the Scrum framework is user story. According ISO / IEC / IEEE 26515-2011 user story is “simple narrative illustrating the user goals

that a software function will satisfy" [20, p.3]. In accordance with this standard user story has to include [20, p.18]:

- role of the user;
- goal that the user achieves;
- value for the customer;
- acceptance criteria that allow to determine that the user story is implemented.

The following recording scheme is usually used in Scrum project for the user story:

As a <type of user X >
I want <some goal Y>
So that <some reason Z>

The last part of the user story should show the user's benefit from using the feature (the reason why this story is needed by the user). In Scrum, a detailed analysis of user story occurs at the moment of implementing the feature to which it relates. During the Sprint planning, the user story usually is used as the basis for a conversation between the Development Team and the Product Owner in order to clarify the details of the implementation. The acceptance criteria for the user story are determined by the Development Team during the discussion.

Acceptance criteria are a set of statements that specify both functional and non-functional requirements. It should be noted that attention is usually focused on functional requirements so many non-functional requirements may not be taken into account. Acceptance criteria can be written in a variety of formats. There are two main approaches to the recording of acceptance criteria: rule-oriented or scenario-oriented ones.Gherkin notation can be used when using a scenario-oriented approach.

During the Daily Scrum meetings, it is discussed what was done during the last working day and what difficulties arose. Each member of the Development Team notifies the Team what they are going to implement for the current day. In a Daily Scrum, the problems found in the requirements for the software product can be discussed and a strategy for their solution can be worked out.

Sprint Review is an event that takes place at the end of Sprint to inspect the Increment and adapt the Product Backlog if needed. Validation of the requirements occurs during the Sprint Review. Requirements management is based on the stakeholder feedback during the Sprint Review. The Sprint Retrospective is an event when the Scrum Team analyzes the work done and creates a plan for improvements to be enacted during the next Sprint. These improvements relate to all work processes including engineering requirements.

4. Quality assessment of user stories

There are many approaches to quality assessment of requirements. The standard ISO/IEC/IEEE 29148:2011 defines quality criteria for the individual requirements and sets of requirements. Each individual requirement must be necessary, implementation free, unambiguous, consistent, complete, singular, feasible, traceable, verifiable. The set of requirements must possess the following quality characteristics: completeness, consistency, affordability, boundedness [1, p.11-12].

The QUS (Quality User Story) Framework proposed by Lucassen et al. in [4] can be used to assess the quality of user stories. The structure of the criteria in the QUS framework is based on the understanding of quality in categories O.I. Lindland. There are syntactic quality, semantic quality and pragmatic quality. Syntactic quality criteria are used to evaluate the textual structure of a user story without considering its meaning. These criteria are used to quality assessment of individual user stories. In the structure of user story, three parts are identified: a role, a means and an ends (optionally). The group of syntactic quality criteria includes:

- well-formed: in the user story test there is at least a role and a means;
- atomic: the test user story expresses a requirement for exactly one feature of the software;
- minimal: in the user story test there is nothing except a role, a means and the ends.

Semantic quality criteria are used to evaluate "the relations and meaning of (parts of) the user story text" [4]. The group of semantic quality criteria includes:

- conflict-free: a user story should not conflict with any other user story;

- conceptually sound: the means expresses a feature and the ends expresses a rationale;
- problem-oriented:a user story determines the problem, not the solution to it;
- unambiguous: the text of user story does not contain terms or abstractions that can lead to multiple interpretation.

The criterion "conflict-free" is used to quality assessment of a set of user stories. The other three criteria are used to quality assessment of individual user stories.

Pragmatic quality criteria affect the “choosing the most effective alternatives for communicating a given set of requirements”[4]. The group of pragmatic quality criteria includes:

- fullsentence: a user story is a well-formed full sentence;
- scalable: a user story does not specify the coarse-grained requirements that are difficult to plan and prioritize;
- unique: each user story is unique, duplicates are not allowed;
- uniform: all user stories in the specification use the same template;
- independent: a user story is self-contained and does not have inherent dependencies on other user stories;
- explicit dependencies: if a user story has non-obvious dependencies on others, explicit references should be made to the stories from which there is a dependency
- complete: implementing a set of user stories creates a feature-complete application, no steps are missing (for critical user stories).

The "full sentence" and "scalable" criteria are used to quality assessment of individual user stories. The remaining criteria are used to quality assessment of a set of user stories.

5. Ontology for requirementsengineering process in Scrum framework

OWL ontology to support the engineering requirements process is implemented in the Protégé environment. This ontology accumulates knowledge about the key features of the requirements engineering in the Scrum framework. The taxonomy of the upper level classes, which are direct descendants of the general *Thing class*, is shown in Figure 3.

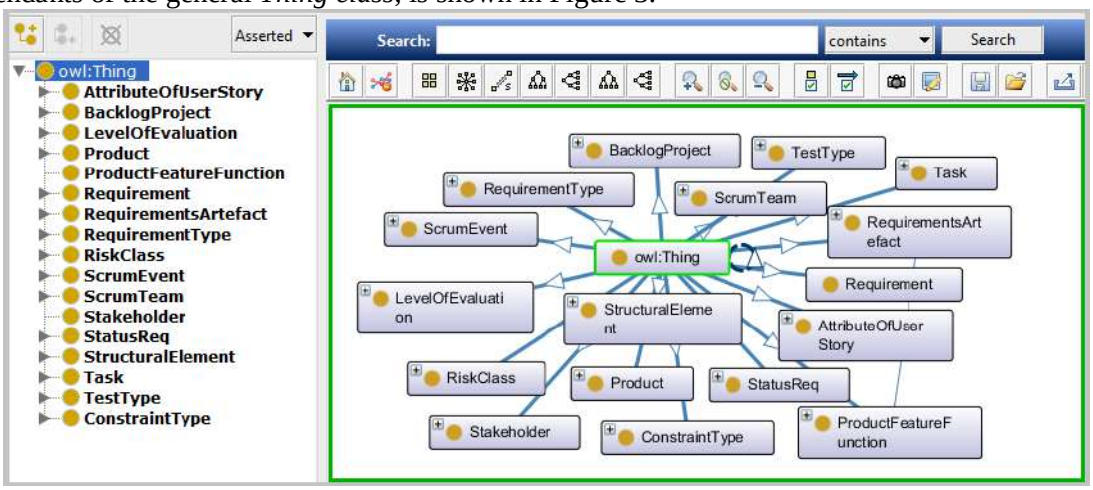


Figure 3. The taxonomy of the upper level classes.

The class *Attribute Of User Story* describes the attributes of user stories such as priority, risk level, etc. The class *Backlog Project* is a class whose subclasses are the types of backlogssuch as Product Backlog, Sprint Backlog and Task Backlog. The class *Level Of Evaluation*is a class whose subclasses are qualitative scales used in assessing priorities, risks, etc. The concept *Product* corresponds to the term "software product". The concept *Product Feature Function* corresponds to the term "function of the software". The class *Requirement* contains subclasses corresponding to the term "requirement". The development of ontology is based on the following understanding of the term "requirement". A requirement is an assertion about some property of the software product. For example, user story, acceptance criteria fixed with use of rule-oriented or scenario-oriented approaches. The class

*Requirements Artefact*contains subclasses corresponding to the concept "requirement artefact". Since such requirements as *Scenarios*, *Acceptance criteria* and *Definition of done* represent information also about the requirements that are created, modified and used in the process of implementing the requirements, they are simultaneously included in the classes *Requirement* and *Requirements Artefact*. Also subclasses of the class *Requirements Artefact* are requirements sources and software features and tests that verify the correctness of the requirements implementation.

The class *Requirement Type* contains requirements classification (for example, functional and non-functional requirements). The class *Risk Class* contains subclasses describing the classification of risks such as a risk of getting a bug, security risk and etc. The class *Scrum Event* contains subclasses corresponding to the Scrum events from the Scrum Guides. The class *Scrum Team* contains subclasses corresponding to the Scrum roles from the Scrum Guides. The class *Stakeholder* corresponds to the term "*stakeholder*". The class *Status Req* contains subclasses describing the status of requirements such as “resolved”, “in progress” and etc. The class *Task* is a class whose instances are tasks performed by the development team. The class *Test Type* contains subclasses describing the classification of test types (for example, performance testing). The class *Constraint Type* contains subclasses describing the types of constraints for non-functional requirements. Constraints express requirements to the internal and external quality of the software product. The class *Structural Element* contains subclasses describing structural elements that must include requirements written using some technique. Figure 4 shows the relations between the class *UserStory* and the classes that describe its structure.

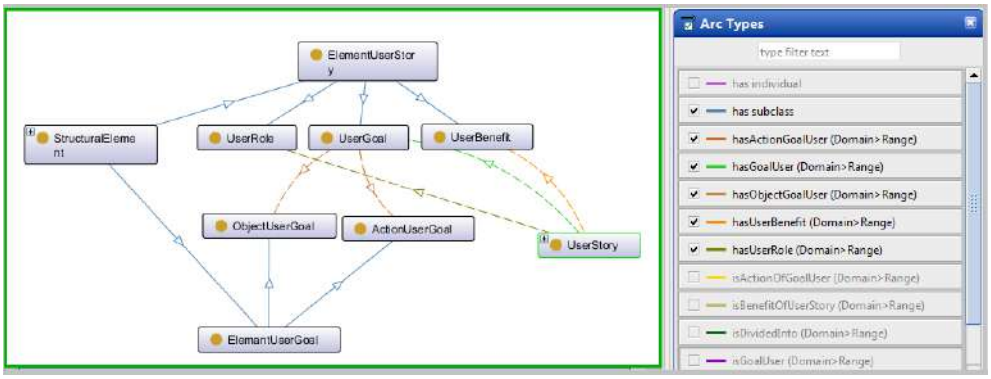


Figure 4. A conceptual structure of user story.

The structure of the user story includes the user role, user goal and the user benefits (or in other words, the reason of the occurrence of the story). The user goal in turn consists of the action that the user makes and the object on which the action is performed. For example, "As a user, I want to sort photos, so that easy to view photos". In this example, the user goal consists of the action "sort" and the object "photos". Theuserbenefitis "easytoviewphotos".

In the ontological model, rules for assessing the quality of individual user stories and sets of user stories are introduced. For example, consider the criterion for evaluating syntactic quality – “Well-formed user story”. A user story is considered well-formed if it specifies the role, goal and benefits. The goal should include the action and the object:

IF User Story has (User Role and Well Formed Goal User and User Benefit) THEN User Story is Well Formed User Story.

IF Goal has (Action Goal User and Object Goal User) THEN Goal is Well Formed Goal User.

The implementation of these rules in the axioms of the ontology in the Protégé environment is shown in Figure 5.

To extract instances of the classes User Role, Action User Goal, Object User Goal and User Benefit from the user story text, morphological, syntactic and semantic analysis of the text can be performed using the appropriate utilities.

A separate user story satisfies the completeness property (internal completeness), if it is included all the information necessary to ensure correct implementation. This means that the user story must be evaluated and acceptance criteria for it must be established, for user story should be indicated the

source and the team member who added it in the baglog. The assessment of user story includes effort estimates, priority and assessing the level of risk. The proposed ontology axiom for assessing the completeness of user story is shown in Figure 6.

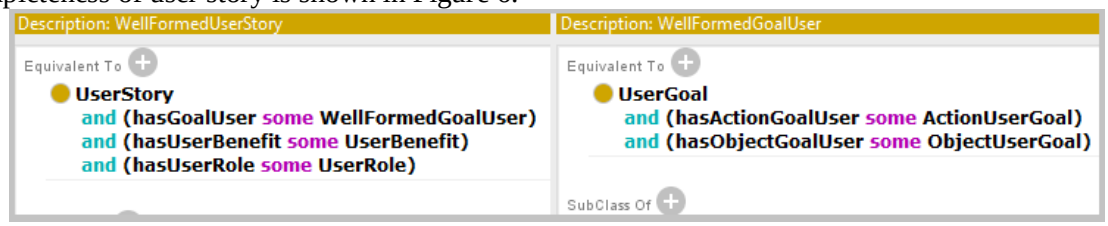


Figure 5. Axioms for evaluating the structure of user story.

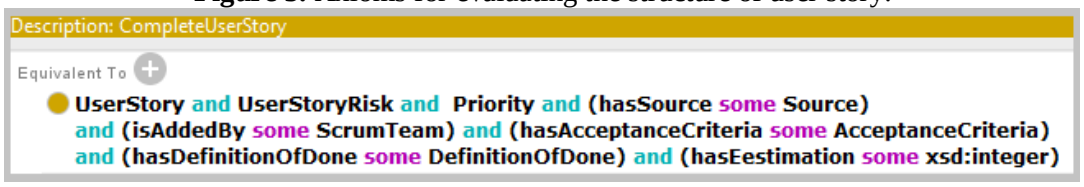


Figure 6. Axioms for evaluating the completeness of user story.

If information is entered to calculate the level of risk then a user story is included in the class *User Story Risk*. If information is entered to calculate the priority of the user story then the user story is included in the class *Priority*. Consider the priority evaluation. In the proposed approach, the user story priority is set depending on the entered parameters – business value and urgency. The simplest from the point of view of the organization of the assessment process is ranking by qualitative scale. In the developed approach, the scale contains four divisions: "critical level" (4), "high" (3), "middle" (2) and "low" (1). The matrix for priority estimation and an example of an axiom for determining a low priority level are shown in Figure 7.

Figure 7. Matrix for priority estimation and an axiom for determining a low priority level.

The developed ontology allows checking the quality criteria of individual user stories. In practice, the definition of completeness and conflictuality of a set of user stories is a context-dependent problem that is difficult to generalize. However, you can talk about the incompleteness of the set of user stories if the user stories meet the requirements for manipulating the elements of the system that were not previously created. You can talk about the conflict of a set of user stories if for one object from one type of user there is permission and lockout the same action. The latter can be partly solved by constructing the domain ontology from the structural elements of user stories and behavior scenarios detailing user stories. Methods of extracting knowledge units from a set of texts can be used for semi-automatic construction of domain ontology (for example, as in [21]).

6. Conclusion

The paper shows that the application of an ontology-oriented approach to supporting the process of requirements engineering in Scrum is an extremely urgent task. Requirements are the raw data for software development and must satisfy certain quality characteristics. Currently, the main characteristics of the requirements quality for software products are defined by the standard ISO/IEC/IEEE 29148-2011 *Systems and software engineering – life cycle processes – requirements engineering*. The analysis of scientific publications allows us to talk about the need to apply assessment models that take into account the specifics of agile requirements.

In present paper the ontology was presented taking into account the features of the Scrum framework and the quality criteria that are characteristic for the user stories. In analyzing the features of the requirements engineering in the Scrum framework found that the main effort is focused on the analysis of the functional component, so non-functional requirements are often not documented. Taking into account high frequency of updating the requirements when working in accordance with the Scrum framework, application of the proposed approach allows to quickly monitoring traceability and completeness of the requirements. Making records in the ontological knowledge base will also serve as a good tool for documenting the progress of the work. The latter can help to increase the productivity of the development team.

7. References

[1] *ISO/IEC/IEEE 29148:2011(E), Systems and software engineering – life cycle processes – requirements engineering* (Version 1.0, 2011-12-01).
[2] VersionOneInc 2018 *The 12th annual State of Agile report* (Access mode: https://explore.versionone.com/state-of-agile/versionone-12th-annual-state-of-agile-report)
[3] Wake B 2003 *INVEST in Good Stories, and SMART Tasks* (Access mode*:* https://xp123.com/articles/invest-in-good-stories-and-smart-tasks/)
[4] Lucassen G, Dalpiaz F, van der Werf J M and Brinkkemper S 2015 *Proc. Int. Conf. on Requirements Engineering* 126-135
[5] Heck P and Zaidman A 2014 *Preprint arXiv*: *1406.4692*
[6] Siegemund K 2014 *Contributions To Ontology-Driven Requirements Engineering : dissertation to obtain the academic degree Doctoral engineer* (Dresden: Technischen Universität Dresden)
[7] Breitman K K and Leite J C S P 2003 *Proc. Int. Conf. on Requirements Engineering* 309-319
[8] Zhu X and Jin Z 2005 *Proc. Int. Conf. onEngineering of Complex Computer Systems* 402-410
[9] Dobson G and Sawyer P 2006 *Int. Seminar on Dependable Requirements Engineering of Computerised Systems at NPPs* (Halden: Institute for Energy Technology)
[10] Kossmann M, Wong R, Odeh M and GilliesA 2008 *Proc. Int. Conf. on Information and Communication Technologies: From Theory to Applications*
[11] Kossmann M, Odeh M, Gillies A and WattsS 2009 *Proc. Int. Conf. onApplications of Digital Information and Web Techn* 95-103
[12] Pustovalova N V and Avdeenko T V 2016 *SPIIRAS Proceedings* **1(44)** 31-49
[13] Avdeenko T V and Pustovalova N V 2016 *Proc. Int. Conf. onActual problems of electronic instrument engineering* **1** 513-518
[14] Avdeenko T V and Pustovalova N V 2015 *Proc. Int. Siberian conference on control and communications*
[15] Goknil A, Kurtev I, van den Berg K 2008 *Proc. Int. Conf. onModel Driven Architecture – Foundations and Applications* 310-325
[16] Thamrongchote C and Vatanawood W 2016 *Proc. Int. Conf. onComputer and Information Science* 633-636
[17] Adnan M and Afzal M 2017 *IEEE Access* **5** 25993-26005
[18] Schwaber K and Sutherland J 2017 *The Scrum Guide* (Access mode: https://www.scrumguides.org /docs/scrumguide/v2017/2017-Scrum-Guide-US.pdf#zoom=100)
[19] Darwish N D and Megahed S 2016 *International Journal of Computer Applications* **149(8)** 24-29
[20] ISO/IEC/IEEE 26515:2011(E), Systems and software engineering *Developing user documentation in an Agile environment (Version 1.0, 2011-12-01)*
[21] Mikhaylov D V, Kozlov A P and Emelyanov G M 2016 *Computer Optics* **40(4)** 572-582 DOI: 10.18287/2412-6179-2016-40-4-572-582

Acknowledgments
The reported study was funded by Russian Ministry of Education and Science, according to the research project No. 2.2327.2017/4.6.

The research of fuzzy decision trees building based on entropy and the theory of fuzzy sets

S B Begenova[1] **and T V Avdeenko**[1]

[1]Novosibirsk State Technical University, Karla Marks ave 20, Novosibirsk, Russia, 630073

Abstract. Decision trees are widely used in the field of machine learning and artificial intelligence. Such popularity is due to the fact that with the help of decision trees graphic models, text rules can be built and they are easily understood by the final user. Because of the inaccuracy of observations, uncertainties, the data, collected in the environment, often take an unclear form. Therefore, fuzzy decision trees are becoming popular in the field of machine learning. This article presents a method that includes the features of the two above-mentioned approaches: a graphical representation of the rules system in the form of a tree and a fuzzy representation of the data. The approach uses such advantages as high comprehensibility of decision trees and the ability to cope with inaccurate and uncertain information in fuzzy representation. The received learning method is suitable for classifying problems with both numerical and symbolic features. In the article, solution illustrations and numerical results are given.Also the comparison of fuzzy logic approaches for building fuzzy rules and classification trees are given.

1. Introduction

Nowadays, in the era of big data, the extraction of knowledge is a bottleneck in the field of knowledge engineering. Computer programs that extract knowledge from data successfully try to solve this problem.Among these programs, systems for building decision trees for decision-making and classification tasks are very popular. The knowledge acquired in the form of decision trees and inference procedures is highly valued for the clarity and visibility of the data. Such assessment, at one time, aroused interest of scientists, which led to a number of methodological and empirical achievements. However, initially decision trees were popularized by Quinlan and his ID3 algorithm [1].

One of the extensions of the classical construction of decision trees is an approach based on fuzzy logic. Fuzzy approach is becoming increasingly popular in solving problems of uncertainty, noise and inaccurate data. It is successfully applied to problems in many industrial spheres. Most studies on the application of this representative framework to existing methodologies are focused mainly on new areas, such as neural networks and genetic algorithms. Nowadays, the fuzzy approach that integrates the concepts of fuzzy sets and entropy is becoming popular.

This article presents a method that includes the features of the two above-mentioned approaches: a graphical representation of the rules system in the form of a tree and the fuzzy representation of the data. Section 2 describes the principle of the decision trees, their advantages and disadvantages, algorithms for their construction. Section 3 shows the principle of constructing fuzzy decision trees, introduces the concepts of fuzzy logic. Section 4 describes the results of the study and the last section gives a conclusion.

2. Decisiontrees

A decision tree (DT) is a common formalization for mapping the transitions of attribute values to classes in the form of a map, which consists of attribute nodes or so-called tests that can have two or more subtrees, leaves, or decision nodes that are labeled with a class indicating the solution. The main advantage of this approach is the visualization of the solution. One of the most commonly used algorithms for constructing decision trees is the ID3 method, formalized by Quinlan in 1986 [1].

Decision trees create efficient models for machine learning [11, 12]. Let us give the following characteristics of decision trees:

- they are easily interpretable and visible;
- the model can be expressed both graphically and with text rules;
- they are competitive in comparison with more expensive approaches;
- decisiontrees are scalable;
- they can process discrete and continuous data;
- decision trees can be applied to different sizes of data sets, including large sample sets.

In the process of tree constructing, the pattern is represented by a set of features that are expressed in some descriptive language. Samples whose characteristics are known are called examples. The purpose of constructing a tree is to solve the problem of classification or regression.

ID3 and CART are the two most important discriminating learning algorithms that work by recursive partitioning. Their basic ideas are approximately the same: splitting the incoming sample into subsets and representing the partitions as a tree. An important property of these algorithms is that they simultaneously try to minimize the size of the tree with the optimization of some quality measure. Subsequently, they use the same logical inference.

3. Fuzzy decisiontrees

To construct a fuzzy decision tree, the following procedure is proposed [4]:

1. Define the fuzzy data base, i.e., the fuzzy granulation for the domains of the continuous features.
2. Replace the continuous attributes of the training set using the linguistic labels of the fuzzy sets with highest compatibility with the input values [5, 6].
3. Calculate the entropy and information gain of each feature to split the training set and define the test nodes of the tree until all features are used or all training examples are classified.

Figure 1 shows an example of the fuzzification of continuous data.

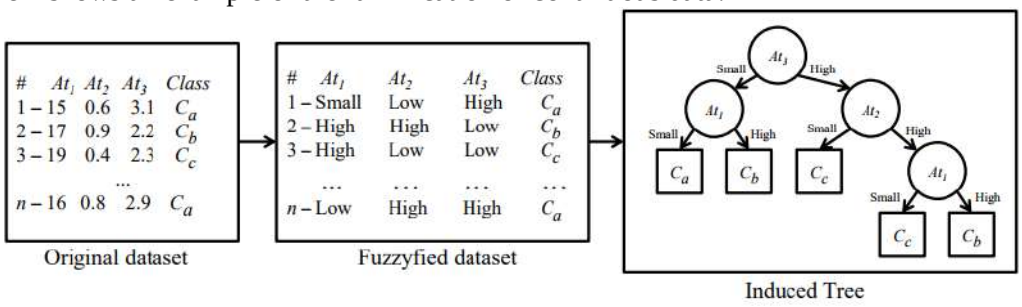


Figure 1. Algorithm for constructing a fuzzy decision tree.

The first block of Figure 1 illustrates a dataset with n examples, three attributes (At_1, At_2,At_3) and a class attribute. The fuzzified version of this dataset is presented in the second block. This fuzzified set of examples is used to induce the final DT, illustrated in the last block of Figure 1.

The entropy and information gain formulas remain the same for the classical version of the ID3 algorithm [10]. Let us introduce the following notation:

$U = \{u_1, u_2, \dots, u_s\}$ – set of data samples; $C = \{c_1, c_2, \dots, c_n\}$ – set of attributes; $D = \{d\}$ – a singleton set with a solution attribute or class attribute. Let this attribute have m different values, thens_i is the number of samples of set U in class d_i.

Information gain I relative to subset S_j is equal to:

$$I(s_{1j}, s_{2j}, \dots, s_{mj}) = -\sum_{i=1}^{m} p_{ij} log_2 p_{ij},$$
$$p_{ij} = \frac{s_{ij}}{|s_j|},$$

where $|S_j|$– the number of samples in a subset of S.

Entropy $E(c_i)$ is:

$$E(c_i) = \sum_{j=1}^{v} \frac{s_{1j} + s_{2j} + \dots + s_{mj}}{s} I(s_{1j}, s_{2j}, \dots, s_{mj});$$

accordingly, the criterion for selecting an attribute is the increase in information:

$$Gain(c_i) = I(s_{1j}, s_{2j}, \dots, s_{mj}) - E(c_i).$$

The difference between common algorithm ID3 and the fuzzy version of algorithm ID3 is that the attributes of objects have degrees of belonging to a particular node, and it is quite possible that an attribute with certain probabilities belongs to several nodes.

Figure 2 shows two decision trees that were built using the above-mentioned algorithms.

As an example, a classic data set was taken, Fisher's iris [9], which has 4 attributes: the length and width of the cup, the length and width of the petal and the three resultant classes - setosa, versicolor and virginica.

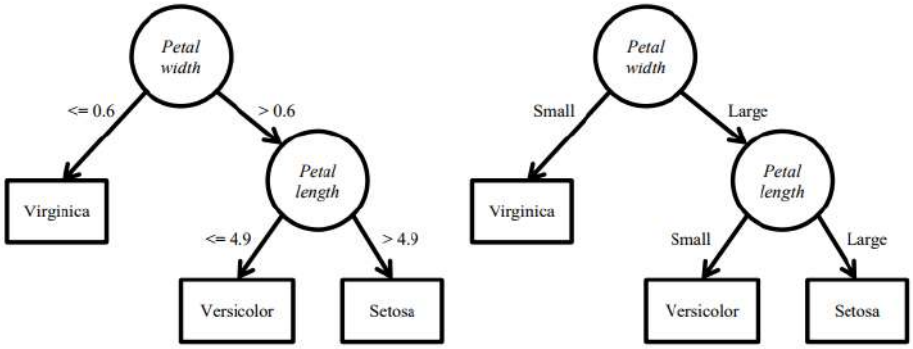


Figure 2. Classical (left) and fuzzy (left) decision trees.

4. Theresearch results

As a research object, just as in the previous example, a Fisher's iris data set was used.To construct a fuzzy decision tree at the first stage, it is necessary to perform a fuzzification procedure.

While performing the fuzzification procedure, the definition set of fuzzy attributes is divided into fuzzy subsets. The value of the fuzzy attribute is put in correspondence with the term, and this correspondence is found using the membership function. The division of the definition set into fuzzy subsets can be made evenly, that is, the definition set is divided into equal intervals. However, in most real data sets obtained from the environment, it is preferable to perform the partitioning taking into account the features of the original sample. For example, it may happen that most of the sampling objects lie in the first third of the definition set and, in this case, uniform partitioning will not give the desired effect.The results of fuzzification for attributes SepalLength, SepalWidth, PetalLength and PetalWidthare shown in Figures 3, 4, 5 and 6 respectively.

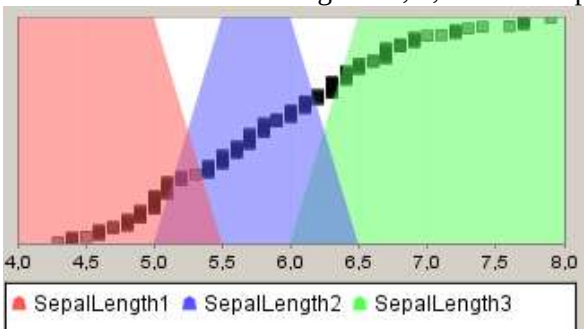


Figure 3. Sepal Length Attribute fuzzification.

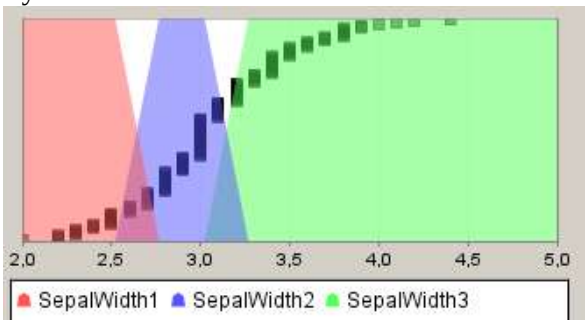


Figure 4. Sepal Width Attribute fuzzification.

Figure 5. Petal Length Attribute fuzzification.

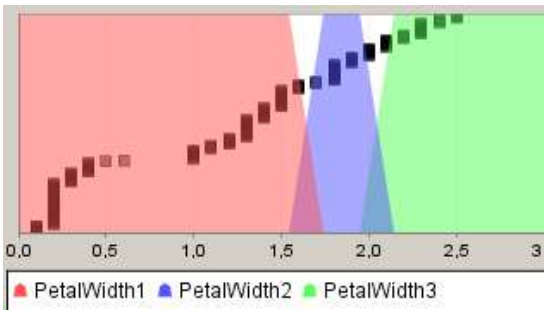


Figure 6. Petal Width Attribute fuzzification.

Numeric attributes have been assigned 3 terms (low, medium, high). Let us present the results of the classification obtained with the help of fuzzy decision trees. In the context of this studies, let us introduce the following notations:

Correct - the number of correctly classed sample objects.

Incorrect - the number of incorrectly classified sample objects.

WithoutClass - the number of objects without a class.

PercentCorrect is the percentage of correctly-categorized sample objects that is calculated as follows:

$$PercentCorrect = 100 - \frac{(Incorrect + WithoutDecision) * 100}{Correct}.$$

To study the hypothesis that with a decrease of the sample size, the accuracy of the classification of fuzzy decision trees is better than classifying with classical ones, the dependence of the classification accuracy on the number of instances in the data set was constructed.

In this study, the trees were constructed for 3 randomly selected N samples and the table shows the averaged values obtained (the sum of the values of the attributes / 3).

According to the data presented in Table 1, it can be seen that when the sample is reduced from 150 to 90, the accuracy of classification using fuzzy decision trees is three percent higher than results of classifying with classical decision trees, and when the sample is reduced to 60, the accuracy is higher by 0.82 percent.

Table 1. Comparison of classification results obtained using fuzzy decision trees and classical decision trees.

The number of instances in the data set (N)	Fuzzy decision trees	Classical decision trees
120	Correct = 115 Incorrect = 5 WithoutClass = 0 PercentCorrect= 95.65	Correct = 117 Incorrect = 3 WithoutClass = 0 PercentCorrect= 97.43
90	Correct= 88 Incorrect = 2 WithoutClass = 0 PercentCorrect= 97.72	Correct= 85 Incorrect = 5 WithoutClass = 0 PercentCorrect= 94.11
60	Correct= 58 Incorrect = 2 WithoutClass = 0 PercentCorrect= 95.55	Correct= 57 Incorrect = 3 WithoutClass = 0 PercentCorrect= 94.73

Table 2 shows the dependence of the accuracy of data classification on the number of terms.

According to the data in the table, it is clear that the optimal number of terms for the test set of data is 5. Such quantity gave a higher percentage of correctly classified data compared to 3 terms.

Table 3 shows the dependence of the accuracy of the data classification on the value of the information gain. In this method, the increment of information will be the breakpoint of the algorithm, that is, when the specified value is reached, further building of the tree stops. According to the data it is clear that the lower the information gain, the more accurate and "deeper" the tree will be built.

Table 2. Dependence of the accuracy of classification of data on the number of terms.

The number of terms	Classification results
3	Correct = 142
	Incorrect = 8
	WithoutClass = 0
	Percent Correct= 94.36
5	Correct = 143
	Incorrect = 7
	WithoutClass = 0
	Percent Correct = 95.33
7	Correct = 143
	Incorrect = 7
	WithoutClass = 0
	Percent Correct= 95.33

Table 3. Dependence of the accuracy of classification of data on the information gain.

Information gain	Classification results
0.02	14 leaves
	Correct = 142
	Incorrect = 8
	WithoutClass = 0
	PercentCorrect= 94.67
0.2	5 leaves
	Correct = 139
	Incorrect = 11
	WithoutClass = 0
	PercentCorrect= 92.67
0.4	3 leaves
	Correct = 119
	Incorrect = 31
	WithoutClass = 0
	PercentCorrect= 79.33

As a part of fuzzy decision trees research, we compare two methods of classification based on fuzzy logic. The first one is algorithm of direct generating fuzzy linguistic rules, proposed in [3]. The second method of fuzzy decision trees which is proposed in this article was used. Fig.7 gives visual illustration of comparison between the two methods for sequentially growing number of terms.

Here we can observe that the classification accuracy for sequentially growing number of terms (from 3 cases to 7) remains quite high. Method of fuzzy decision trees is better for medium and high sizes of terms while the method of direct generating fuzzy rules is better for small size of training sample. Further research will be in development of algorithm based on combination of both approaches.

On fig. 8 we can observe that the classification accuracy for sequentially reducing size of training sample (from 105 cases to 45) remains quite high. Method of fuzzy decision trees is better for medium size of training sample while the method of direct generating fuzzy rules is better for small size of training sample. Further research will be in development of algorithm based on combination of both approaches.

Fig. 9 and Fig. 10 illustrate the comparison between the two methods with sequentially reducing size of training sample using T-class and S- class membership functions respectively.

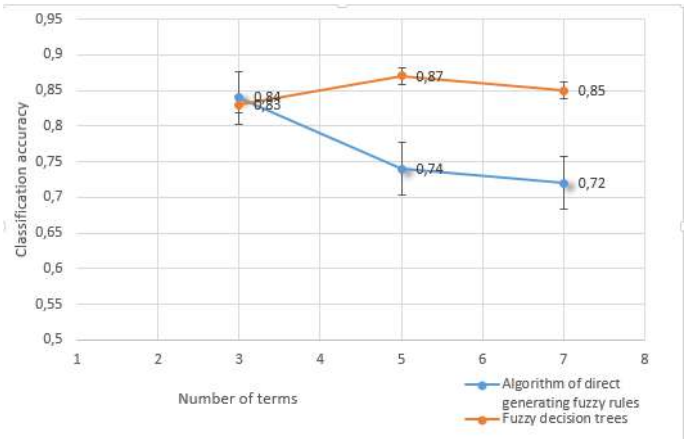


Figure 7. Classification accuracy for sequentially growing number of terms.

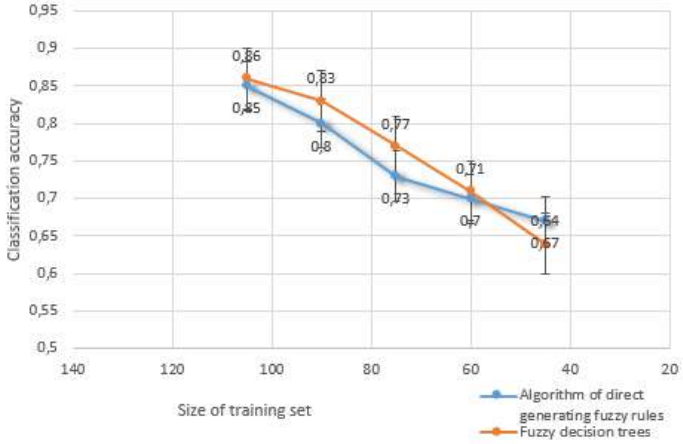


Figure 8. Classification accuracy for sequentially reducing size of training sample.

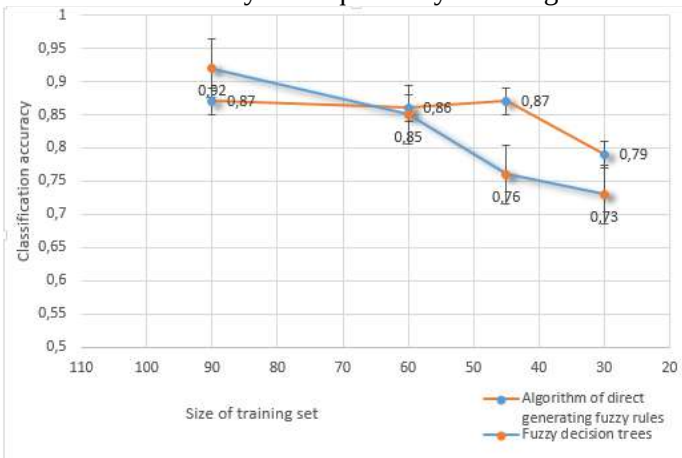


Figure 9. Classification accuracy for sequentially reducing size of Iris dataset training sample (T – class).

T – class membership function also known as triangular is specified by three parameters {a, b, c} as follows:

$$t(x,a,b,c) = \begin{cases} 0, x \le a \\ \dfrac{x-a}{b-a}, a \le x \le b \\ \dfrac{c-x}{c-b}, b \le x \le c \\ 0, c \le x \end{cases}$$

S – class membership function is specified by two parameters {a, b} as follows:

$$s(x,a,b) = \begin{cases} 0, x \le a \\ \dfrac{1}{1+e^{-a(x-b)}} \\ 0, x \ge b \end{cases}$$

For T-class membership functions, we observe that fuzzy decision trees method is better for large size of training sample while the method of direct generating fuzzy rules is better for small and medium size of training sample. In case of S-class membership functions, we observe the same situation.

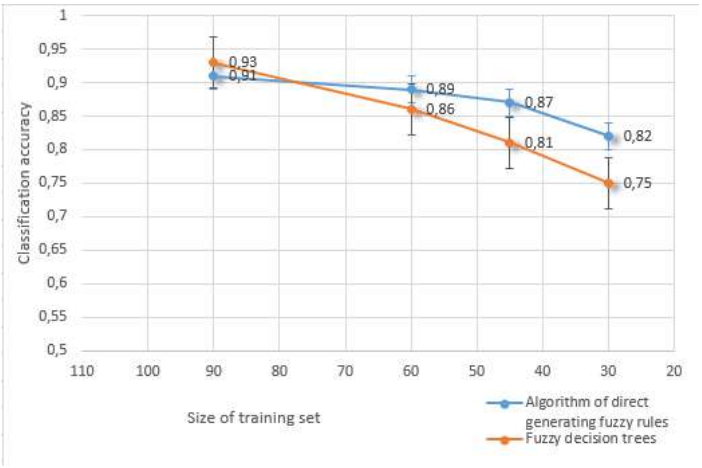


Figure 10. Classification accuracy for sequentially reducing size of Iris dataset training sample (S – class).

5. Conclusion

Decision trees are successfully used to solve regression and classification problems. They are popular in the field of machine learning, because decision trees build graphic models, along with text rules that are easily interpreted by the users. On the other hand, fuzzy systems can solve classification problems with input inaccurate and noisy data.

The combination of fuzzy trees and fuzzy logic makes it possible to construct intuitive graphic models for qualitative and quantitative data [2, 7, 8]. Usage of this type of decision tree gives us several solutions with different probabilities of belonging to a particular class.

In addition, in the course of the conducted studies, the advantage of classification using fuzzy decision trees with respect to classical ones was revealed, by comparing the percentage of correctly classed objects. Also, a direct correlation between the accuracy of the classification and the value of the information gain was revealed (the increment is a criterion for stopping the further construction of the tree).

The comparison between algorithm of direct generating fuzzy linguistic rules and method of fuzzy decision trees didn't reveal the one and only right one. Both methods show high classification accuracy under certain conditions. The proposed approach can be applied to built fuzzy neural networks [13].

6. References

[1] Quinlan J R 1986 Induction of decision trees *Machine learning* **1** 81-106
[2] Cintra M E, Meira C A A, Monard M C, Camargo H A and Rodrigues L H 2011 The use of fuzzy decision trees for coee rust warning in Brazilian *Int. Conf. Int. Sys. Design & Applications* **1** 1347- 1352
[3] Avdeenko T V and Makarova E S 2017 Acquisition of knowledge in the form of fuzzy rules for cases classification *Lecture Notes in Computer Science* **10387** 536-544
[4] Cintra M E, Monard M C and Camargo H A 2012 Fuzzy DT- a fuzzy decision tree algorithm based on C4.5 *CBSF -Brazilian Congress on Fuzzy Systems* 199-211

[5] Janikow C Z 1998 Fuzzy Decision Trees: Issues and Methods *IEEE Transactions of Man, Systems, Cybernetics* **28(1)**
[6] Faifer M and Janikow C Z 2000 Bottom-up Partitioning in Fuzzy Decision Trees *Proceedings of the 19th International Conference of the North American Fuzzy Information Society* 326-330
[7] Tokumaru M and Muranaka N 2010 Impression analysis using fuzzy c4.5 decision tree *Int. con. on Kansei engineering and emotion research*
[8] Janikow C Z 2004 Fid 4.1: an overview *Proc. of the North American Fuzzy Information Processing Society* 877-881
[9] *Machine Learning Repository* (Access mode: http://archive.ics.uci.edu/ml/datasets/Iris) (30.05.2018)
[10] Begenova S B and Avdeenko T V 2018 Building of fuzzy decision trees using ID3 algorithm *Journal of Physics: Conference Series* **1015**
[11] Cintra M E, Meira C A A, Monard M C, Camargo H A and Rodrigues L H 2011 The use of fuzzy decision trees for coffee rust warning in Brazilian crop *Int. Conf. Int. Sys. Design & Applications* **1** 1347-1352
[12] Olaru C and Wehenkel L 2003 Fuzzy Sets and Systems **138** 221-254
[13] Soldatova O P, Lezin I A, Lezina I V, Kupriyanov A V and Kirsh D V 2015 Application of fuzzy neural networks to determine the type of crystal lattices observed on nanoscale images *Computer Optics* **39(5)** 787-795 DOI: 10.18287/0134-2452-2015-39-5-787-794

Acknowledgments
The work is supported by a grant from the Ministry of Education and Science of the Russian Federation within the framework of the project part of the state task, project No. 2.2327.2017 / 4.6 “Integration of knowledge representation models based on intellectual analysis of large data to support decision making in the field of software engineering.”

Dual ordered structures of binary relations

V P Tsvetov[1]

[1]Samara National Research University, Moskovskoe shosse 34, Samara, Russia, 443086

Abstract. The theory of ordered structures like a (lattice) ordered semigroups is applied to graphs and automatons as well as to coding, programming and artificial intelligence. In this paper an algebraic structure on an underlying set of binary relations is considered. The structure includes the operations of Boolean algebra, inverse and composition. It is defined a dual semigroup to the binary relations ordered semigroup, and then the general properties of dual operations are studied.

1. Introduction

Abstract theory of algebraic structures (sometimes called universal algebra) forms the basis for various applications [1-8]. Semigroups and lattices are the simplest structures but not the least ones.

Let's recall some definitions:

The semigroup is a set with single binary operation $*$ satisfying associative low. A semigroup with neutral (identity) element is called a monoid;

The semiring is a set with couple of binary operations – addition and multiplication - satisfying associative lows. There are neutral elements for both of them and addition is commutative. Also multiplication distributes over addition and multiplication by zero annihilates semiring;

The lattice (as an algebraic structure) is a set with pair of binary operations – join and meet - satisfying associative lows, commutative lows, and absorption lows. A distributive lattice is a lattice in which the operations of join and meet distribute over each other. A bounded lattice is a lattice with neutral elements. The lattice's bottom is a neutral element for the join operation and the lattice's top is a neutral element for the meet operation;

The lattice (as a poset) is a partial ordered set such that each finite-elements subset has *supremum* (join) and *infimum* (meet). A bounded lattice is a lattice with bottom and top elements;

The ordered semigroup is a semigroup together with a partial order $\prec$ that is compatible with the semigroup operation i.e. $\forall u,v,w\, u \prec v \to w*u \prec w \circ v \wedge u*w \prec v*w$. The bounded semigroup is an ordered semigroup with bottom and top elements.

It's well known that any ordered semigroup is isomorphic to a subsemigroup of binary relations ordered by subset relation. In this paper we deal with a left composition of binary relation as a semigroup operation, i.e. we set

$$R_1 \circ R_2 = \left\{ (u_1,u_2) \mid \exists u_3 \, (u_1,u_3) \in R_1 \wedge (u_3,u_2) \in R_2 \right\} \quad (1)$$

At first, we denote a universe as U and consider a power set of Cartesian square $2^{U\times U}$ as a collection of binary relations on U. The traditional approach to studying binary relations leads to ordered semigroup $S_{\mathrm{R}} = \left\langle 2^{U\times U}, (\circ, \subseteq) \right\rangle$ and bounded distributive lattice $L_{\mathrm{R}} = \left\langle 2^{U\times U}, (\cup, \cap) \right\rangle$. In this way we don't take into account a complement operation

$$\overline{R} = \{(u_1,u_2) \mid (u_1,u_2) \notin R\} \quad (2)$$

However, it's very convenient to use a complement element. For example, we can write the trichotomy low for relation R in several forms. First, we can write it as in equation (3)

$$I_R \cup R \cup R^{-1} = 1_R = U \times U = \overline{\varnothing} = \overline{0_R} \quad (3)$$

Then, we can rewrite it in alternative form as antisymmetric low for the complement $\bar{R}$ as in equation (4)

$$\bar{R} \cap \bar{R}^{-1} \subseteq I_R = \{(u,u) \mid u \in U\} \quad (4)$$

In this case and below we use the notations 1_R, 0_R and I_R for complete relation, empty relation and identity relation respectively. Note that 1_R and 0_R are top and bottom elements for lattice L_R. Also we denote the inverse relation of R as

$$R^{-1} = \{(u_2,u_1) \mid (u_1,u_2) \in R\} \quad (5)$$

2. Algebraic structure h_R

Let's consider an algebraic structure $h_R = \left\langle 2^{U \times U}, \left(\cup, \cap, \circ, \bar{}, {}^{-1}, \subseteq, 0_R, 1_R, I_R\right)\right\rangle$. It's easy to prove properties (6)-(41):

$$R_1 \cup (R_2 \cup R_3) = (R_1 \cup R_2) \cup R_3 \quad (6)$$
$$R_1 \cap (R_2 \cap R_3) = (R_1 \cap R_2) \cap R_3 \quad (7)$$
$$R_1 \circ (R_2 \circ R_3) = (R_1 \circ R_2) \circ R_3 \quad (8)$$
$$R_1 \cup R_2 = R_2 \cup R_1 \quad (9)$$
$$R_1 \cap R_2 = R_2 \cap R_1 \quad (10)$$
$$0_R \cup R = R \quad (11)$$
$$1_R \cap R = R \quad (12)$$
$$I_R \circ R = R \circ I_R = R \quad (13)$$
$$1_R \cup R = 1_R \quad (14)$$
$$0_R \cap R = 0_R \quad (15)$$
$$0_R \circ R = R \circ 0_R = 0_R \quad (16)$$
$$1_R \circ 1_R = 1_R \quad (17)$$
$$R_1 \cup (R_2 \cap R_3) = (R_1 \cup R_2) \cap (R_1 \cup R_3) \quad (18)$$
$$R_1 \cap (R_2 \cup R_3) = (R_1 \cap R_2) \cup (R_1 \cap R_3) \quad (19)$$
$$R_1 \circ (R_2 \cup R_3) = (R_1 \circ R_2) \cup (R_1 \circ R_3) \quad (20)$$
$$(R_2 \cup R_3) \circ R_1 = (R_2 \circ R_1) \cup (R_3 \circ R_1) \quad (21)$$
$$R_1 \circ (R_2 \cap R_3) \subseteq (R_1 \circ R_2) \cap (R_1 \circ R_3) \quad (22)$$
$$(R_2 \cap R_3) \circ R_1 \subseteq (R_2 \circ R_1) \cap (R_3 \circ R_1) \quad (23)$$
$$R_1 \cup (R_1 \cap R_2) = R_1 \quad (24)$$
$$R_1 \cap (R_1 \cup R_2) = R_1 \quad (25)$$
$$R \cup R = R \quad (26)$$
$$R \cap R = R \quad (27)$$
$$\overline{\overline{R}} = R \quad (28)$$
$$\overline{1_R} = 0_R \quad (29)$$
$$\overline{0_R} = 1_R \quad (30)$$

$$\overline{R_1 \cup R_2} = \overline{R}_1 \cap \overline{R}_2 \tag{31}$$

$$\overline{R_1 \cap R_2} = \overline{R}_1 \cup \overline{R}_2 \tag{32}$$

$$\left(R_1 \cup R_2\right)^{-1} = R_1^{-1} \cup R_2^{-1} \tag{33}$$

$$\left(R_1 \cap R_2\right)^{-1} = R_1^{-1} \cap R_2^{-1} \tag{34}$$

$$\left(R_1 \circ R_2\right)^{-1} = R_2^{-1} \circ R_1^{-1} \tag{35}$$

$$\overline{R^{-1}} = \overline{R}^{-1} \tag{36}$$

$$R_1 \subseteq R_2 \Leftrightarrow R_1 \cup R_2 = R_2 \Leftrightarrow R_1 \cap R_2 = R_1 \tag{37}$$

$$R_1 \subseteq R_2 \Rightarrow R_1 \cup R_3 \subseteq R_2 \cup R_3 \tag{38}$$

$$R_1 \subseteq R_2 \Rightarrow R_1 \cap R_3 \subseteq R_2 \cap R_3 \tag{39}$$

$$R_1 \subseteq R_2 \Rightarrow R_1 \circ R_3 \subseteq R_2 \circ R_3 \wedge R_3 \circ R_1 \subseteq R_3 \circ R_2 \tag{40}$$

$$R_1 \subseteq R_2 \Leftrightarrow \overline{R}_1 \supseteq \overline{R}_2 \tag{41}$$

The typical algebraic structures we can obtain by restriction of structure $h_{R(U\times U)}$ are as follows:

The bounded lattices of binary relations $LO_{\mathrm{R}}^1 = \left\langle 2^{U\times U}, \left(\subseteq, 0_{\mathrm{R}}, 1_{\mathrm{R}}\right)\right\rangle$ and $LO_{\mathrm{R}}^2 = \left\langle 2^{U\times U}, \left(\supseteq, 1_{\mathrm{R}}, 0_{\mathrm{R}}\right)\right\rangle$;

The bounded monoids of binary relations $M_{\mathrm{R}}^1 = \left\langle 2^{U\times U}, \left(\cup, \subseteq, 0_{\mathrm{R}}, 1_{\mathrm{R}}\right)\right\rangle$, $M_{\mathrm{R}}^2 = \left\langle 2^{U\times U}, \left(\cap, \subseteq, 0_{\mathrm{R}}, 1_{\mathrm{R}}\right)\right\rangle$, $M_{\mathrm{R}}^3 = \left\langle 2^{U\times U}, \left(\cup, \supseteq, 1_{\mathrm{R}}, 0_{\mathrm{R}}\right)\right\rangle$, $M_{\mathrm{R}}^4 = \left\langle 2^{U\times U}, \left(\cap, \supseteq, 1_{\mathrm{R}}, 0_{\mathrm{R}}\right)\right\rangle$ and $M_{\mathrm{R}} = \left\langle 2^{U\times U}, \left(\circ, \subseteq, I_{\mathrm{R}}, 0_{\mathrm{R}}, 1_{\mathrm{R}}\right)\right\rangle$;

The bounded semirings (with multiplicative identity) of binary relations $SR_{\mathrm{R}}^1 = \left\langle 2^{U\times U}, \left(\cup, \cap, \subseteq, 0_{\mathrm{R}}, 1_{\mathrm{R}}\right)\right\rangle$, $SR_{\mathrm{R}}^2 = \left\langle 2^{U\times U}, \left(\cap, \cup, \supseteq, 1_{\mathrm{R}}, 0_{\mathrm{R}}\right)\right\rangle$, and $SR_{\mathrm{R}} = \left\langle 2^{U\times U}, \left(\cup, \circ, \subseteq, 0_{\mathrm{R}}, I_{\mathrm{R}}\right)\right\rangle$;

The Boolean algebras of binary relations $B_{\mathrm{R}}^1 = \left\langle 2^{U\times U}, \left(\cup, \cap, \bar{}, 0_{\mathrm{R}}, 1_{\mathrm{R}}\right)\right\rangle$ and $B_{\mathrm{R}}^2 = \left\langle 2^{U\times U}, \left(\cap, \cup, \bar{}, 1_{\mathrm{R}}, 0_{\mathrm{R}}\right)\right\rangle$.

3. Dual semigroup to S_{R}

Let's consider a Boolean isomorphism $F\left(R\right) = \overline{R}$ from B_{R}^1 onto B_{R}^2. We define a binary operation • in accordance with duality principle $F\left(R_1 \bullet R_2\right) = F\left(R_1\right) \circ F\left(R_2\right)$, i.e. we set

$$R_1 \bullet R_2 = \overline{\overline{R_1} \circ \overline{R_2}} = \left\{\left(u_1, u_2\right) \mid \forall u_3 \left(u_1, u_3\right) \in R_1 \vee \left(u_3, u_2\right) \in R_2\right\}. \tag{42}$$

Note that $F\left(0_{\mathrm{R}}\right) = 1_{\mathrm{R}}$, $F\left(1_{\mathrm{R}}\right) = 0_{\mathrm{R}}$, $F\left(I_{\mathrm{R}}\right) = \overline{I_{\mathrm{R}}}$ and moreover

$$R_1 \circ R_2 = \overline{\overline{R_1} \bullet \overline{R_2}} = \left\{\left(u_1, u_2\right) \mid \exists u_3 \left(u_1, u_3\right) \in R_1 \wedge \left(u_3, u_2\right) \in R_2\right\} \tag{43}$$

$$R_1 \bullet \left(R_2 \bullet R_3\right) = \left(R_1 \bullet R_2\right) \bullet R_3 \tag{44}$$

$$\overline{I_{\mathrm{R}}} \bullet R = R \bullet \overline{I_{\mathrm{R}}} = R \tag{45}$$

$$1_{\mathrm{R}} \bullet R = R \bullet 1_{\mathrm{R}} = 1_{\mathrm{R}} \tag{46}$$

$$0_{\mathrm{R}} \bullet 0_{\mathrm{R}} = 0_{\mathrm{R}} \tag{47}$$

$$R_1 \bullet \left(R_2 \cap R_3\right) = \left(R_1 \bullet R_2\right) \cap \left(R_1 \bullet R_3\right) \tag{48}$$

$$\left(R_2 \cap R_3\right) \bullet R_1 = \left(R_2 \bullet R_1\right) \cap \left(R_3 \bullet R_1\right) \tag{49}$$

$$R_1 \bullet \left(R_2 \cup R_3\right) \supseteq \left(R_1 \bullet R_2\right) \cup \left(R_1 \bullet R_3\right) \tag{50}$$

$$\left(R_2 \cup R_3\right) \bullet R_1 \supseteq \left(R_2 \bullet R_1\right) \cup \left(R_3 \bullet R_1\right) \tag{51}$$

$$\left(R_1 \bullet R_2\right)^{-1} = R_2^{-1} \bullet R_1^{-1} \tag{52}$$

$$R_1 \supseteq R_2 \Rightarrow R_1 \bullet R_3 \supseteq R_2 \bullet R_3 \wedge R_3 \bullet R_1 \supseteq R_3 \bullet R_2 \tag{53}$$

By our construction semigroup S_{R} is isomorphic to semigroup $\overline{S}_{\mathrm{R}}=\left\langle 2^{U\times U},(\bullet,\supseteq)\right\rangle$ as well as monoid M_{R} and semiring SR_{R} are isomorphic to $\overline{M}_{\mathrm{R}}=\left\langle 2^{U\times U},\left(\bullet,\supseteq,\overline{I_{\mathrm{R}}},1_{\mathrm{R}},0_{\mathrm{R}}\right)\right\rangle$ and $\overline{SR}_{\mathrm{R}}=\left\langle 2^{U\times U},\left(\cap,\bullet,\supseteq,1_{\mathrm{R}},\overline{I_{\mathrm{R}}}\right)\right\rangle$ respectively. In such cases, we'll say that the algebraic structures are dual.

Now we use the previous definitions to argue the following logical consequences:

$$u_1R_1\circ(R_2\bullet R_3)u_2\Leftrightarrow\exists u_3\ u_1R_1u_3\wedge u_3(R_2\bullet R_3)u_2\Leftrightarrow\exists u_3\ u_1R_1u_3\wedge(\forall u_4\ u_3R_2u_4\vee u_4R_3u_2)$$
$$\Leftrightarrow\exists u_3\forall u_4\ u_1R_1u_3\wedge(u_3R_2u_4\vee u_4R_3u_2)\Leftrightarrow\exists u_3\forall u_4\ (u_1R_1u_3\wedge u_3R_2u_4)\vee(u_1R_1u_3\wedge u_4R_3u_2)$$
$$\Rightarrow\forall u_4\exists u_3\ (u_1R_1u_3\wedge u_3R_2u_4)\vee(u_1R_1u_3\wedge u_4R_3u_2)\Leftrightarrow\forall u_4\ (\exists u_3\ u_1R_1u_3\wedge u_3R_2u_4)\vee(\exists u_3\ u_1R_1u_3\wedge u_4R_3u_2)$$
$$\Leftrightarrow\forall u_4\ u_1R_1\circ R_2u_4\vee(\exists u_3\ u_1R_1u_3\wedge u_4R_3u_2)\Leftrightarrow\forall u_4\ (u_1R_1\circ R_2u_4\vee u_4R_3u_2)\wedge(u_1R_1\circ R_2u_4\vee\exists u_3\ u_1R_1u_3)$$
$$\Leftrightarrow(\forall u_4\ u_1R_1\circ R_2u_4\vee u_4R_3u_2)\wedge(\forall u_4\ u_1R_1\circ R_2u_4\vee\exists u_3\ u_1R_1u_3)$$
$$\Leftrightarrow u_1R_1\circ(R_2\bullet R_3)u_2\wedge(\forall u_4\ u_1R_1\circ R_2u_4\vee\exists u_3\ u_1R_1u_3)$$
$$\Leftrightarrow u_1R_1\circ(R_2\bullet R_3)u_2\wedge((\forall u_4\exists u_3\ u_1R_1\circ R_2u_4)\vee(\exists u_3\ u_1R_1u_3))\Rightarrow u_1R_1\circ(R_2\bullet R_3)u_2\wedge(\exists u_3\ u_1R_1u_3)$$
$$\Leftrightarrow u_1R_1\circ(R_2\bullet R_3)u_2\wedge(u_1\in D_{R_1})$$

Let's denote a domain of R_1 as $D_{R_1}=\{u_1\mid\exists u_3\ u_1R_1u_3\}\subseteq U$ and then consider a binary relation $E(D_{R_1},1_{\mathrm{R}})=\{(u_1,u_2)\mid u_1\in D_{R_1}\wedge u_2\in U\}=D_{R_1}\times U\subseteq 1_{\mathrm{R}}=U\times U$. Now we can write

$$R_1\circ(R_2\bullet R_3)\subseteq(R_1\circ R_2)\bullet R_3\cap E(D_{R_1},1_{\mathrm{R}}) \tag{54}$$

Note that

$u_1R_1\circ 1_{\mathrm{R}}u_2\Leftrightarrow\exists u_3\ u_1R_1u_3\wedge u_3 1_{\mathrm{R}}u_2\Leftrightarrow\exists u_3\ u_1R_1u_3\wedge u_2\in U\Leftrightarrow u_1\in D_{R_1}\wedge u_2\in U\Leftrightarrow u_1E(D_{R_1},1_{\mathrm{R}})u_2$

and so we obtain

$$R_1\circ(R_2\bullet R_3)\subseteq(R_1\circ R_2)\bullet R_3\cap R_1\circ 1_{\mathrm{R}} \tag{55}$$

Similarly, we get

$$(R_2\bullet R_3)\circ R_1\subseteq R_2\bullet(R_3\circ R_1)\cap 1_{\mathrm{R}}\circ R_1 \tag{56}$$
$$R_1\bullet(R_2\circ R_3)\supseteq(R_1\bullet R_2)\circ R_3\cup R_1\bullet 0_{\mathrm{R}} \tag{57}$$
$$(R_2\circ R_3)\bullet R_1\supseteq R_2\circ(R_3\bullet R_1)\cup 0_{\mathrm{R}}\bullet R_1 \tag{58}$$

In the latter, we have taken into account the following equalities:

$$E(1_{\mathrm{R}},D_{R_1^{-1}})=U\times D_{R_1^{-1}}=1_{\mathrm{R}}\circ R_1$$
$$E(\overline{D_{R_1}},1_{\mathrm{R}})=\overline{D_{R_1}}\times U=R_1\bullet 0_{\mathrm{R}}$$
$$E(1_{\mathrm{R}},\overline{D_{R_1^{-1}}})=U\times\overline{D_{R_1^{-1}}}=0_{\mathrm{R}}\bullet R_1$$

4. Extension of algebraic structure h_{R}

At first, we denote $O_{\mathrm{R}}=\overline{I_{\mathrm{R}}}$ and then we consider $H_{\mathrm{R}}=\left\langle 2^{U\times U},\left(\cup,\cap,\bullet,\circ,^{-},^{-1},\subseteq,0_{\mathrm{R}},1_{\mathrm{R}},O_{\mathrm{R}},I_{\mathrm{R}}\right)\right\rangle$ as an extension of algebraic structure h_{R}. It is clear that all of the properties (6)-(58) are true for structure H_{R}.

Let's rewrite (57)-(58) in the form

$$(R_2\bullet R_3)\circ R_1\cup R_2\bullet 0_{\mathrm{R}}\subseteq R_2\bullet(R_3\circ R_1)$$
$$R_1\circ(R_2\bullet R_3)\cup 0_{\mathrm{R}}\bullet R_3\subseteq(R_1\circ R_2)\bullet R_3$$

So we can rewrite (55)-(58) as:

$$R_1 \circ (R_2 \bullet R_3) \subseteq (R_1 \circ R_2) \bullet R_3 \cap R_1 \circ 1_{\mathrm{R}} \tag{59}$$

$$R_1 \circ (R_2 \bullet R_3) \cup 0_{\mathrm{R}} \bullet R_3 \subseteq (R_1 \circ R_2) \bullet R_3 \tag{60}$$

$$(R_2 \bullet R_3) \circ R_1 \subseteq R_2 \bullet (R_3 \circ R_1) \cap 1_{\mathrm{R}} \circ R_1 \tag{61}$$

$$(R_2 \bullet R_3) \circ R_1 \cup R_2 \bullet 0_{\mathrm{R}} \subseteq R_2 \bullet (R_3 \circ R_1) \tag{62}$$

Obviously, for all binary relations $R_1, R_2, R_3 \in 2^{U \times U}$ we have

$$R_1 \circ (R_2 \bullet R_3) \subseteq (R_1 \circ R_2) \bullet R_3 \tag{63}$$

$$(R_2 \bullet R_3) \circ R_1 \subseteq R_2 \bullet (R_3 \circ R_1) \tag{64}$$

This is immediate from the inclusions (59)-(62).

Properties like the (55)-(58), (63)-(64) we'll call the laws of semi-compatibility. Now we are interested in cases of compatibility (low) of dual operation with each other

$$R_1 \circ (R_2 \bullet R_3) = (R_1 \circ R_2) \bullet R_3 = R_1 \circ R_2 \bullet R_3 \tag{65}$$

$$(R_2 \bullet R_3) \circ R_1 = R_2 \bullet (R_3 \circ R_1) = R_2 \bullet R_3 \circ R_1 \tag{66}$$

Note that we won't find algebraic substructures of H_{R} satisfying (65)-(66). Indeed, from (16), (17), (46), (47) we obtain

$$0_{\mathrm{R}} \circ (0_{\mathrm{R}} \bullet 1_{\mathrm{R}}) = 0_{\mathrm{R}} \neq 1_{\mathrm{R}} = (0_{\mathrm{R}} \circ 0_{\mathrm{R}}) \bullet 1_{\mathrm{R}} \tag{67}$$

$$0_{\mathrm{R}} \circ (1_{\mathrm{R}} \bullet 1_{\mathrm{R}}) = 0_{\mathrm{R}} \neq 1_{\mathrm{R}} = (0_{\mathrm{R}} \circ 1_{\mathrm{R}}) \bullet 1_{\mathrm{R}} \tag{68}$$

$$(1_{\mathrm{R}} \bullet 0_{\mathrm{R}}) \circ 0_{\mathrm{R}} = 0_{\mathrm{R}} \neq 1_{\mathrm{R}} = 1_{\mathrm{R}} \bullet (0_{\mathrm{R}} \circ 0_{\mathrm{R}}) \tag{69}$$

$$(1_{\mathrm{R}} \bullet 1_{\mathrm{R}}) \circ 0_{\mathrm{R}} = 0_{\mathrm{R}} \neq 1_{\mathrm{R}} = 1_{\mathrm{R}} \bullet (1_{\mathrm{R}} \circ 0_{\mathrm{R}}) \tag{70}$$

Hence, we have to restrict structure H_{R} to find algebraic substructures satisfying (65)-(66) and we'll call them compatible (sub)structures. Let's consider H_{R} without 0_{R} or 1_{R}.

We are studying the simplest cases of subsets of $2^{U \times U}$ as an underlying set for the operations from H_{R} below.

Let's denote the collection of (partial) functions from U to U as $U^U_{+1} \subseteq 2^{U \times U}$. It's easy that $\langle U^U_{+1}, (\circ, \subseteq, I_{U \times U}, 0_{U \times U}) \rangle$ is a bounded below submonoid of M_{R}.

We want to prove that U^U_{+1} is closed under the dual operation $\bullet$.

At first, we suppose that U contains only one element. In this case $U^U_{+1} = 2^{U \times U} = \{I_{\mathrm{R}}, 0_{\mathrm{R}}\}$, where I_{R} is identity function and 0_{R} is empty function. So U^U_{+1} is closed under $\bullet$ because $2^{U \times U}$ is closed.

Let now U contains more than one element. Suppose $R_1, R_2 \in U^U_{+1}$, but $R_1 \bullet R_2 \notin U^U_{+1}$. Hence, there are $u_1, u_2, u_3 \in U$ such that $u_2 \neq u_3$ and $((u_1, u) \in R_1 \vee (u, u_2) \in R_2) \wedge ((u_1, u) \in R_1 \vee (u, u_3) \in R_2)$ for all $u \in U$. However, $R_1 \in U^U_{+1}$ and so there is no more than one $u_0 \in U$ satisfying relation $(u_1, u_0) \in R_1$. Whence $u_2 \neq u_3 \wedge (u, u_2) \in R_2 \wedge (u, u_3) \in R_2$ for all $u \in U \setminus \{u_0\} \neq \varnothing$. The latter is in contradiction with $R_2 \in U^U_{+1}$.

The proof is complete.

Let's consider the scale of sets $U^U_0 \subseteq U^U \subseteq U^U_{+1}$, where U^U is a collection of total functions and U^U_0 is a collection of total bijections from U to U.

We assume U to be a two-element set and denote the cardinality of $U = \{u_1, u_2\}$ as $|U|$. Obviously, $|2^{U \times U}| = 16$, $|U^U_{+1}| = 9$, $|U^U| = 4$, $|U^U_0| = |U| = 2$. Note $1_{\mathrm{R}} \notin U^U_{+1}$, $0_{\mathrm{R}} \in U^U_{+1}$, $0_{\mathrm{R}} \notin U^U$.

We have simulated some interesting cases of algebraic substructures to check irregularities in (65)-(66). Table 1 contains statistics on the incompatibility of dual operations.

Table 1. A total amount of incompatibility of $\circ$ and $\bullet$.

$R_1, R_2, R_3 \in W$	$R_1 \circ (R_2 \bullet R_3) \neq (R_1 \circ R_2) \bullet R_3$	$(R_2 \bullet R_3) \circ R_1 \neq R_2 \bullet (R_3 \circ R_1)$
$W = 2^{U \times U} \setminus \{0_R\}$	706 from 3375	706 from 3375
$W = 2^{U \times U} \setminus \{1_R\}$	706 from 3375	706 from 3375
$W = U_{+1}^U$	90 from 729	20 from 729
$W = U^U$	0 from 64	0 from 64
$W = U^U \cup \{0_R\}$	10 from 75	4 from 75
$W = U^U \cup \{1_R\}$	4 from 75	10 from 75
$W = U_0^U$	0 from 8	0 from 8
$W = U_0^U \cup \{0_R\}$	0 from 27	0 from 27
$W = U_0^U \cup \{1_R\}$	0 from 27	0 from 27

Cayley tables 2 and 3 describes the dual operations on the set $U_0^U \cup \{0_R, 1_R\}$.

Table 2. A Cayley table for $\circ$ on the set $U_0^U \cup \{0_R, 1_R\}$.

$\circ$	O_R	I_R	0_R	1_R
O_R	I_R	O_R	0_R	1_R
I_R	O_R	I_R	0_R	1_R
0_R	0_R	0_R	0_R	0_R
1_R	1_R	1_R	0_R	1_R

Table 3. A Cayley table for $\bullet$ on the set $U_0^U \cup \{0_R, 1_R\}$.

$\bullet$	O_R	I_R	0_R	1_R
O_R	O_R	I_R	0_R	1_R
I_R	I_R	O_R	0_R	1_R
0_R	0_R	0_R	0_R	1_R
1_R	1_R	1_R	1_R	1_R

The cases of incompatibility on the set $U_0^U \cup \{0_R, 1_R\}$ are listed below apart from (67)-(70).

$R_1 \circ (R_2 \bullet R_3) \neq (R_1 \circ R_2) \bullet R_3$:

$$1_R \circ (I_R \bullet 0_R) = 0_R \neq 1_R = (1_R \circ I_R) \bullet 0_R$$
$$1_R \circ (O_R \bullet 0_R) = 0_R \neq 1_R = (1_R \circ O_R) \bullet 0_R$$
$$0_R \circ (I_R \bullet 1_R) = 0_R \neq 1_R = (0_R \circ I_R) \bullet 1_R$$
$$0_R \circ (O_R \bullet 1_R) = 0_R \neq 1_R = (0_R \circ O_R) \bullet 1_R$$

$(R_2 \bullet R_3) \circ R_1 \neq R_2 \bullet (R_3 \circ R_1)$:

$$(0_R \bullet I_R) \circ 1_R = 0_R \neq 1_R = 0_R \bullet (I_R \circ 1_R)$$
$$(0_R \bullet O_R) \circ 1_R = 0_R \neq 1_R = 0_R \bullet (O_R \circ 1_R)$$
$$(1_R \bullet I_R) \circ 0_R = 0_R \neq 1_R = 1_R \bullet (I_R \circ 0_R)$$
$$(1_R \bullet O_R) \circ 0_R = 0_R \neq 1_R = 1_R \bullet (O_R \circ 0_R)$$

It's easy to see that $U_0^U,(\bullet,O_{\mathrm{R}})$ and $\langle U_0^U,(\circ,I_{\mathrm{R}})\rangle$ are abelian groups. Moreover, $\langle U_0^U\cup\{0_{\mathrm{R}}\},(\cap,\bullet,\circ,^{-1},\subseteq,0_{\mathrm{R}},O_{\mathrm{R}},I_{\mathrm{R}})\rangle$ is a bounded below compatible algebraic structure and $\langle U_0^U\cup\{1_{\mathrm{R}}\},(\cup,\bullet,\circ,^{-1},\subseteq,1_{\mathrm{R}},O_{\mathrm{R}},I_{\mathrm{R}})\rangle$ is a bounded above compatible algebraic structure.

Let's give other examples.

Let $F_1\subseteq U_{+1}^U$ be a set of partial and total functions are listed as $O_{\mathrm{R}}=\{(u_1,u_2),(u_2,u_1)\}$, $I_{\mathrm{R}}=\{(u_1,u_1),(u_2,u_2)\}$, $0_{\mathrm{R}}=\varnothing$, $f_1=\{(u_1,u_1)\}$, $f_2=\{(u_1,u_2)\}$, $f_3=\{(u_2,u_1)\}$, $f_4=\{(u_2,u_3)\}$.

Cayley tables 4 and 5 describes the dual operations on the F_1.

Table 4. A Cayley table for $\circ$ on the set F_1.

$\circ$	O_{R}	I_{R}	0_{R}	f_1	f_2	f_3	f_4
O_{R}	I_{R}	O_{R}	0_{R}	f_3	f_4	f_1	f_2
I_{R}	O_{R}	I_{R}	0_{R}	f_1	f_2	f_3	f_4
0_{R}	0_{R}	0_{R}	0_{R}	0_{R}	0_{R}	0_{R}	0_{R}
f_1	f_2	f_1	0_{R}	f_1	f_2	0_{R}	0_{R}
f_2	f_1	f_2	0_{R}	0_{R}	0_{R}	f_1	f_2
f_3	f_4	f_3	0_{R}	f_3	f_4	0_{R}	0_{R}
f_4	f_3	f_4	0_{R}	0_{R}	0_{R}	f_3	f_4

Table 5. A Cayley table for $\bullet$ on the set F_1.

$\bullet$	O_{R}	I_{R}	0_{R}	f_1	f_2	f_3	f_4
O_{R}	O_{R}	I_{R}	0_{R}	f_1	f_2	f_3	f_4
I_{R}	I_{R}	O_{R}	0_{R}	f_3	f_4	f_1	f_2
0_{R}	0_{R}	0_{R}	0_{R}	0_{R}	0_{R}	0_{R}	0_{R}
f_1	f_1	f_2	0_{R}	0_{R}	0_{R}	f_1	f_2
f_2	f_2	f_1	0_{R}	f_1	f_2	0_{R}	0_{R}
f_3	f_3	f_4	0_{R}	0_{R}	0_{R}	f_3	f_4
f_4	f_4	f_3	0_{R}	f_3	f_4	0_{R}	0_{R}

It's easy to see that $\langle F_1,(\cap,\bullet,\circ,^{-1},\subseteq,0_{\mathrm{R}},O_{\mathrm{R}},I_{\mathrm{R}})\rangle$ is a bounded below compatible algebraic structure.

Now let $F_2\subseteq U^U$ be a set of total functions are listed as O_{R}, I_{R}, $g_1=\{(u_1,u_1),(u_2,u_1)\}$, $g_2=\{(u_1,u_2),(u_2,u_2)\}$.

Cayley tables 6 and 7 describes the dual operations on the F_2.

Table 6. A Cayley table for $\circ$ on the set F_2.

$\circ$	O_{R}	I_{R}	g_1	g_2
O_{R}	I_{R}	O_{R}	g_1	g_2
I_{R}	O_{R}	I_{R}	g_1	g_2
g_1		g_1	g_1	g_2
	g_1		g_1	g_2

Table 7. A Cayley table for $\bullet$ on the set F_2.

$\bullet$	O_{R}	I_{R}	g_1	g_2
O_{R}	O_{R}	I_{R}	g_1	g_2
I_{R}	I_{R}	O_{R}	g_1	g_2
g_1	g_1	g_2	g_1	g_2
g_2	g_2	g_1	g_1	g_2

In this case $\left\langle F_2,\left(\cap,\bullet,\circ,^{-1},\subseteq,O_{\mathrm{R}},I_{\mathrm{R}}\right)\right\rangle$ is unbounded compatible algebraic structure. Taking into account (42)-(43) we can write

$$R_1\circ\left(R_2\bullet R_3\right)=\left(R_1\circ R_2\right)\bullet R_3\Leftrightarrow\overline{R}_1\bullet\left(\overline{R}_2\circ\overline{R}_3\right)=\left(\overline{R}_1\bullet\overline{R}_2\right)\circ\overline{R}_3 \tag{71}$$

$$\left(R_2\bullet R_3\right)\circ R_1=R_2\bullet\left(R_3\circ R_1\right)\Leftrightarrow\left(\overline{R}_2\circ\overline{R}_3\right)\bullet\overline{R}_1=\overline{R}_2\circ\left(\overline{R}_3\bullet\overline{R}_1\right) \tag{72}$$

Let's denote the sets of relations are complement of functions from F_1 and F_2 as $\overline{F_1}=\left\{O_{\mathrm{R}},I_{\mathrm{R}},1_{\mathrm{R}},\overline{f_1},\overline{f_2},\overline{f_3},\overline{f_4}\right\}$ and $\overline{F_2}=\left\{O_{\mathrm{R}},I_{\mathrm{R}},\overline{g_1},\overline{g_2}\right\}$. In accordance with (71)-(72) we get ordered (not bounded and not lattice) compatible algebraic structures $\left\langle\overline{F_1},\left(\cup,\bullet,\circ,^{-1},\subseteq,1_{\mathrm{R}},O_{\mathrm{R}},I_{\mathrm{R}}\right)\right\rangle$ and $\left\langle\overline{F_2},\left(\bullet,\circ,\subseteq,O_{\mathrm{R}},I_{\mathrm{R}}\right)\right\rangle$.

Of cause, the list of examples can be continued.

5. Conclusion

We have studied non-traditional algebraic structures on the underlying set of binary relations. Starting from left composition, inclusion and Boolean isomorphism we defined dual ordered semigroups. Then we extended them to the more general ordered algebraic structure with a couple of dual operations. We have proved that these operations satisfy the semi-compatibility laws. This is notable and important fact. We paid special attention to the algebraic substructures satisfying the compatibility laws. So we have considered interesting examples of compatible algebraic structures.

The results will be useful for graphs and automatons as well as for coding, programming and artificial intelligence.

References

[1] Clifford A H and Preston G B 1961 The Algebraic Theory of Semigroups *Mathematical Surveys and Monographs* **7(1)**
[2] Clifford A H and Preston G B 1967 The Algebraic Theory of Semigroups *Mathematical Surveys and Monographs* **7(2)**
[3] Birkhoff G 1967 *Lattice Theory* (Providence RI: American Mathematical Society)
[4] Fuchs L 1963 *Partially Ordered Algebraic Systems* (Oxford: Pergamon Press)
[5] Ershov A P 1982 Abstract computability in algebraic systems *Proc. Int. Symp. Algorithms in Modern Mathematics and its Applications* (Novosibirsk: Computing Center of the Siberian Branch of the USSR Academy of Sciences) **2** 194-299
[6] Chernov V M 2015 Quasiparallel algorithm for error-free convolution computation using reduced Mersenne-Lucas codes *Computer Optics* **39(2)** 241-248 DOI: 10.18287/ 0134-2452-2015-39-2-241-248
[7] Chernov V M 2018 Calculation of Fourier-Galois transforms in reduced binary number systems *Computer Optics* **42(3)** 495-500 DOI: 10.18287/ 0134-2452-2018-42-3-495-500
[8] Makhortov S D and Shurlin M D 2013 LP-Structures analysis: substantiation of refactoring in object-oriented programming *Automation and Remote Control* **74(7)** 1211-1217

Geometric and game approaches for some discrete optimization problems

B F Melnikov[1], E A Melnikova[1], S V Pivneva[1], V A Dudnikov[2] and E V Davydova[3]

[1]Russian State Social University, Wilhelm Pieck str. 4, Moscow, Russia, 129226
[2]Togliatti State University, Belorusskaya str. 16, Togliatti, Russia, 445020
[3]Moscow Aviation Institute (State Technical University), Volokolamskoe shosse 4, Moscow, Russia, 125993

Abstract. We consider in this paper the adaptation of heuristics used for programming non-deterministic games to the problems of discrete optimization. In particular, we use some "game" heuristic methods of decision-making in various discrete optimization problems. The object of each of these problems is programming anytime algorithms. Among the problems described in this paper, there are the classical traveling salesman problem and some connected problems of minimization for nondeterministic finite automata. The first of the considered methods is the geometrical approach to some discrete optimization problems. For this approach, we define some special characteristics relating to some initial particular case of considered discrete optimization problem. For instance, one of such statistical characteristics for the traveling salesman problem is a significant development of the so-called "distance functions" up to the geometric variant such problem. And using this distance, we choose the corresponding specific algorithms for solving the problem.

Besides, other considered methods for solving these problems are constructed on the basis of special combination of some heuristics, which belong to some different areas of the theory of artificial intelligence. More precisely, we shall use some modifications of unfinished branch-and-bound method; for the selecting immediate step using some heuristics, we apply dynamic risk functions; simultaneously for the selection of coefficients of the averaging-out, we also use genetic algorithms; and the reductive self-learning by the same genetic methods is also used for the start of unfinished branch-and-bound method again. This combination of heuristics represents a special approach to construction of anytime-algorithms for the discrete optimization problems. This approach can be considered as an alternative to application of methods of linear programming, and to methods of multi-agent optimization, and also to neural networks.

1. Introduction. A brief survey of discrete optimization problems

We consider in this paper the adaptation of heuristics used for programming non-deterministic games to the problems of discrete optimization, in particular, some heuristic methods of decision-making in various discrete optimization problems (DOP). The object of each of these problems is programming anytime algorithms, i.e., the algorithms, which can provide a near-to-optimal solution in real time. The basic purposes of the paper are practical questions of construction of algorithms, as well as the creation of the corresponding theory. Let us briefly list the considered problems, more precisely, the classes of considered problems.

First, it is the classical traveling salesman problem (TSP; [1] etc.). Certainly, an universal methods for solving TSP simply cannot exist. Some last years, the authors of papers for heuristic

methods of TSP-solution consider most often so-called *metric TSP*. For their solving, some methods (of linear programming, of multi-agent optimization, etc.) are used; [1, 2, 3, 4, 5] etc. However, the classical branch-and-bound method (BBM) can also be used not only for the exact (optimal) solution of considered TSP, but also for quasi-optimal heuristic solutions. We shall write below about these things more detailed.

Second, these are some related problems of minimization for nondeterministic finite automata (Rabin-Scott automata, NFA). Probably, the main for them is state-minimization, i.e., the problem of constructing NFA, which defines the given regular language and has minimum possible number of states. Since 1970 (i.e., since [6]), there are a few changes in description of the exact algorithms for this problem: all the algorithms are exponential relative to the number of states of considered NFA. The last argument is true because all the algorithms need to construct equivalent automaton of canonical form (or, maybe, some similar graphs or other objects). Let us remark, that from the point of view of the theory of complexity of algorithms, all the algorithms of [6, 7, 8, 9, 10, 11] are equivalent.

Besides, there are other problems for NFA-minimization, the following ones:

- edge-minimization [12];
- and also the star-height-minimization. There exists two solutions of the last problem ([13], and also [14] with the simplification of the proof [15]). However, the authors think that there is impossible *to make a computing algorithm* on basis of these papers.

Third, this is the problem of minimization of disjunctive normal forms (DNF). The exact algorithms for this minimization are obtained for ages (and are considered in the classical textbooks for first-year students), however the computer programs making on basis of such solutions cannot work in real time even for the number of variables, which is equal to 20, except, certainly, for a lot of trivial cases. The unified approach of this paper is used in some versions of computer programs.

Let us remark, that, certainly, these groups of problems do not formulate the whole set of problems, which can be heuristically solved by the methods considered in this paper. Let us also mention, for instance, [16]. Some other groups of problems are given in the conclusion, and, probably, each DOP can be solved in such a way.

The methods of solution DOP, considered in this paper, are constructed on the basis of special combination of some heuristics, which belongs to some different areas of the theory of artificial intelligence. Firstly, we shall use some modifications of unfinished branch-and-bound methods. Secondly, for the selecting immediate step using some heuristics, we use dynamic risk functions. Thirdly, simultaneously for the selection of coefficients of the averaging-out, we also use genetic algorithms. Fourthly, the reductive self-learning by the same genetic methods is also used for the start of unfinished branch-and-bound method. Let us consider now the detailed description of these heuristic methods of DOP-solution.

2. Traveling salesman problem and its quasi-metric variant

Often, a mathematical model, as well as algorithms based on this model, created for one area, find application in many other subject areas. As we said before, an example of such a model is the traveling salesman problem. The peculiarity of this problem is that, with the relative simplicity of its formulation, finding the optimal solution (the optimal route) is a very complex problem and relates, both in its generalized formulation and for most of its variations, to the NP-complete class. Moreover, according to the classification given in [1] etc., the traveling salesman problem is an example of the optimization problem included in the most complex NPO(V) class: it contains all optimization problems for which (with some additional "natural" assumption, for example, $P \neq NP$), the time complexity of all possible polynomial algorithms cannot be limited

by any polylogarithmic function. (Another example of such a problem is the maximum clique problem.)

Among many various versions of TSP, one of the most studied is the *geometric* version (also referred as *Euclidean*, see also [1]): the cost of the route is equal to the distance between points ("cities") on the plane, calculated as the Euclidean norm. For further discussion, that a characteristic feature of this problem is the fulfillment of the triangle inequality for any three cities. Namely, for any $u, v, w \in E$ the following inequality holds:

$$c(\{u, v\}) \leq c(\{u, w\}) + c(\{w, v\}).$$

We note that this formula assumes only so-called *symmetric* variants of TSP (this term is also defined in [1]); however, in *not* symmetric variants, everything is the same. In general, according to [1], the variants of the problem of TSP with the triangle inequality satisfied for any three cities are called *metric*, that is, all geometric TSP represent a proper subset of metric ones.

However, the main research of the authors of this paper is aimed at studying not geometric, but so-called *pseudo-geometric* version of TSP, [17] etc. In it, the input data is formed as follows. To the data of some *predefined* TSP, which is *geometric*, a vector

$$R = (r_1, \ldots, r_m), \quad (\text{where} \quad m = |E|)$$

is added. Here, all the values r_i are independent and identically distributed random variables (IID); we consider normal distribution only, and allow $\mu = 1$ and some predefined "acceptable" value σ. In this case, each of the elements of the value matrix (i.e., $c(\{u, v\})$; let it be c_i for some $i \in \{1, \ldots, m\}$) is changed for the following value:

$$\max(c_i \cdot r_i, 0).$$

An important difference between the pseudo-geometric version of TSP and the geometric one is the possibility of violating the triangle inequality for some triples of cities. It is also important to note, that the geometric version of TSP can be considered as a special case of pseudo-geometric one (with the value $\sigma = 0$).

Here is our view on the "ecological niche" of the version of the traveling salesman problem we are considering. As we said before, the geometric version of TSP defined above is one of the most studied; as a result of theoretical and practical research, many approaches to solving its particular cases have been developed. Among them, one can single out the so-called geometric approaches using in the search for the solution information that the coordinates of the cities *were previously obtained* as a particular case of a geometric TSP. As examples of algorithms belonging to a group of such geometric approaches and described 15 and more years ago, we can cite the so-called "onion-peeling" algorithms, see [18], and also "elastic network" ones, see [19]. With a small computational complexity, these approaches make it possible to obtain a fairly good solution: for particular cases containing millions of points, it practically coincides with the optimal one. However, outside the geometric version of TSP, these algorithms usually turn out to be meaningless; and this paper can be considered as an attempt to apply similar algorithms to other versions of TSP.

Thus, for TSP, we use the following names:

- "accidental TSP", when all the elements of the TSP-matrix are generated by the variate having the given equipartition law;
- "metric TSP", when we consider towns as the accidental points of the unit square (both the coordinate have the given equipartition law), and the elements of the TSP-matrix are their distances. And here is also evident the following symmetric condition: $a_{ij} = a_{ji}$ for each possible i and j.

- "quasi-metric TSP", when all the elements of the metric TSP-matrix are multiplied a posteriori by a random number, which is obtained by a given normal low.

Some last years, the papers for metric TSP are mostly published. The consideration of metric TSP as the problems of linear programming or the applications of so-called methods of multi-agent optimization was started long before 2000, [20, 21]. And the quasi-metric TSP, which is almost not considered, is more interesting, because of the following:

- first, it is more closely to various practical problems;
- second, various heuristics can be checked up here, which are not connected to use of an arrangement of cities on a plane; moreover, the reduction of this problem to a problem of linear programming is here ineffectively;
- and third, the simplification of the TSP-matrix by one step of BBM is here "less significant", than in other TSP-variants.

Therefore, the metric TSP is the most important scope for algorithms considered in this paper.

However, in [1], the only exact BBM was considered; it finishes by constructing the optimal solution. And in practice, we can rarely obtain the exact TSP-solution using only algorithms of [1]. Using some special programming techniques (e.g., special data structures for quick making the next step of BBM, organization of swapping by the programmer), we can only a little improve the situation. In fact, each of such programming techniques is a new heuristics, which is used in addition to considered applying BBM. However, we are considering the exact solution of TSP (and other DOP) only, and, for now, are not considering the things connected with anytime algorithms. Before the formulation such algorithms, let us consider the following definitions.

Considering a BBM-step, we have to designate the problem for the next solution, obtained by reduction of dimension, exactly *the right problem* (like [1] etc.). For example, the right problem of TSP is obtained when we include the edge between two considered towns; and the right problem of NFA-minimization is obtained when we include the selected grid (see [9] etc.). The other alternative, i.e., when we make a decision about the absence of some element in the optimal solution (e.g., if we consider TSP not containing the edge between two considered towns) is designated by *the left problem*. It is evident, that the object of each modification of BBM (for each DOP) is to obtain the case, when the probability of belonging the optimal solution in the right problem is more then the same probability for the left problem; *we make this thing using some special heuristics*. The explanation of this fact is trivial: dimension of the right problem is less than the left one.

The simple heuristics, which reforms the usual BBM into the unfinished one (and on basis of unfinished BBM we construct any anytime algorithms in this paper), is the following. Each time, when we obtain the next right problem (let us call it problem T) we make at the same time also the sequence of the right (sub)problems (SRP), i.e., T, then the right problem of T, then the right problem of the right problem of T, etc. Certainly, we make each time also the corresponding left problems, i.e., the left problem of T, then the left problem of the right problem of T, etc. This process finishes:

- when we obtain a trivial problem (e.g., of dimension 1), then we use its solution (i.e., its bound, and also the obtained path, and similar behavior) by the current quasi-optimal solution of the considered anytime algorithm;
- or when we obtain the big value of the bound, for example, if this value is more than the current (existing) quasi-optimal solution.

Let us remark, that in practice such process of SRP-constructing does not require a lot of time, and the increasing the dimension of the list of problems for the solution in the future is very reasonable.

Thus, we have described the simple process of constructing the anytime algorithm on the basis of the given version of BBM. And it is unlikely, that such process is described here at the first time (it is really very simple), however, the authors have not the references for this thing. Let us also remark, that this algorithm of SRP-constructing is used as the sub-algorithm not only for the unfinished BBM (like this section), but also for so called algorithm of tournament self-learning.

3. Nondeterministic games and dynamic estimation of a position. Dynamic risk functions in games and in discrete optimization

Because of space limitations, the rules of backgammon are not presented in the paper. Among scientific works devoted to programming of this game, we mention the papers [22, 23, 24]. However, the authors of this paper hope, that the use of dynamic risk functions (DRF) considered here simplifies conventional methods of neural network programming and learning; they are an alternative to these methods.

What is the difference between backgammon (and other nondeterministic games) and, to say, chess (or other deterministic games) from the programming standpoint? The difference is that the game tree constructed for backgammon includes not only the vertexes where the players choose a next move but also those where they wait for a particular realization of a random event. Therefore, the standard minimax method is to be generalized for programming of search in nondeterministic games. In this paper, we will only briefly describe this generalization.

We assume that the reader knows the canonical minimax method. And in nondeterministic games, we have the following alternate actions:

- a particular realization of a certain random event;
- a move of one player;
- another realization of the random event;
- and a move of the other player.

The number of possible outcomes of the random event is to be finite (otherwise, we need different models). As a result, the game tree contains additional levels between those corresponding to the players' moves. These new levels correspond to the moments when the random event is realized. It is such a tree that the generalization of the minimax method.

Assume that we can construct a static estimator of a position. By temporally eliminating indeterminacy, we preliminary estimators of the game tree positions. For this purpose, we assume first that a particular outcome of the random event has been already realized and calculate the dynamic estimator of the position in the same way as in the conventional minimax method. Then, we calculate the dynamic estimator for the next outcome of the random event, and so on for all possible outcomes.

The final dynamic estimator of the position is based on the deterministic estimators of all possible outcomes of the random event. The values of the deterministic (usually, static) estimators are averaged in a special way resulting in the dynamic estimator. From the physical standpoint, such averaging gives us the coordinate of the gravity center of a one-dimensional system of masses whose values are determined by a specially chosen function (risk function). Coordinates of the masses are equal to the values of the corresponding deterministic estimator, which is determined by only deterministic factors of the game, like in the conventional minimax. Let ai be values of deterministic estimators and f be a risk function. Then, according to [25, 26], the dynamic estimator is calculated by the formula

$$\frac{\sum_{i=1}^{k} a_i \cdot f(a_i)}{\sum_{i=1}^{k} f(a_i)}.$$

Let us note, that purpose of the paper [25] was to generalize the minimax method. The important thing to note, however, is that the program based on this generalization only, with the simplest risk function for static estimation of a backgammon position, showed good results and won most part of the programs that the authors of [25] could find that time in Internet. This program has been gradually improving since then. Note also, that the ways of improving the program were very different from those discussed in "classical" works on programming of backgammon [22, 23, 24]. In the latter works, one or another way of calculation of static estimators of a position is optimized. Some ways to improve programs, which were used by the authors after the simplest dynamic estimator had been already introduced, are described in this paper. Note that they concern not only improvements of the static estimator of a position.

The question is to what extent the weights of the opponent's casts that are favorable for us are to be reduced? Even if we simply take the risk function $y = 1 - 0.4x$ and use this function independent of any other circumstances with the simplest static position estimator, we will get rather good results. The short description of practical results can be found in [25]. Note also that in that work, different decreasing risk functions were considered.

But to get a stronger program, it is required *to change* strategies $y = 1 - 0.4x$ during the game. One way to improve the dynamic estimator of a position (described in [25]) is as follows. We qualitatively estimate the position (whether we are about to win or to loose). Then:

- if we are about to win or loose a little, we should be pessimists and adopt a risk function similar to above-mentioned function $y = 1 - 0.4x$;
- if we loose more, the risk function should be close to constant one;
- and if the loss is great, the risk function should be increasing; in this case we need to be super-optimists and hope against hope (what else can we do?).

Of course, there are many other, intermediate, variants of risk functions. And a possible approach to dynamic selecting these intermediate variants was described in detail in [26]. Thus, the authors of this paper believe that, in the given case, the methods of modifications of plots of risk functions simplify conventional methods of neural network programming and learning. This follows, for example, from the fact that one of the authors created a good program for playing backgammon using less than 3 self-learning coefficients, whereas programs described in [23, 24] (see also above-mentioned web sites) use several hundreds such parameters for neural networks.

4. Geometric approach based on the classification of input data

Usually, when developing algorithms, an idealized model of input data is used, most often it is a model with uniform or normal distributions of values of specially allocated *characteristics* of the problem under consideration. However, in practice, the input data, as a rule, come in accordance with some other probability distribution, i.e. distribution, other than uniform or normal. As a result, the performance of the algorithm on real data is calculated inadequately to the input data; moreover, this situation occurs in both the average and the worst case.

Another situation is possible: when the algorithm is designed to take into account the characteristics of the input data that are characteristic of one subject area; and the application of such algorithms to another domain can be extremely inefficient due to the fact that the characteristic features of real data turn out to be different from those considered in the development of the algorithm.

Many scientists are working on a *theory* describing algorithms for a supposedly correct distribution in each particular case, which would sufficiently adequately describe many practical situations. In the opinion of the authors of this paper, the greatest progress in this direction was achieved in the works of Yu. Gurevich ([27] etc.). In some cases, the evaluation of representativeness may not be an end in itself, but should serve as a basis for approaches to evaluating the effectiveness of algorithms.

In the opinion of the authors of this paper, some specific models are needed for each problem under consideration, i.e., the *concrete interpretation* of similar approaches, the so-called. methods "ad hoc". To implement this approach, we have defined special characteristics relating to some initial particular case of TSP; it should be noted that they were chosen similarly to the characteristics used in [28] for a completely different discrete optimization problem. The statistical characteristics used by us for TSP, in fact, are a significant development of the so-called "distance functions" *up to the geometric variant* of TSP $dist(G, c)$, described in [1, Ex. 4.2.3.2]; we can say that these characteristics reflect the cases of non-fulfillment of this inequality in the given particular case of TSP. [1]

For their calculation, similarly to the above example of [1], we considered all possible pairs of points $u, v \in V$ $(u \neq v)$. For each of these pairs and each point $p \in V$, the value

$$\max\Big(0, \frac{c(\{u,v\})}{c(\{u,p\}) + c(\{p,v\})}\Big), \quad p \neq u,\ p \neq v$$

was calculated. For a collection [2] D of such values, and also for the collection $\mathcal{D}$, composed of non-zero elements of collection D, we considered the following characteristics:

- $|D|/N$, where N is the maximum possible number of non-zero elements [3] of collection D, i.e.,
$$\frac{n \cdot (n-1) \cdot (n-2)}{2};$$
- $|\mathcal{D}|/N$;
- the expected value of D, where D is considered as a realization of a random variable;
- the expected value of $\mathcal{D}$;
- the variance of $\mathcal{D}$; [4]
- the average harmonic of the collection $\mathcal{D}$;
- the median of the collection $\mathcal{D}$;
- value of the Durbin – Watson statistic ([29, 30]) for several variants of collections $\mathcal{D}$ (i.e., if there are different inputs).

Based on these characteristics with a perceptron ([31] etc.), a decision was taken on the class for which the particular case of TSP was generated.

The specific description of computational experiments is as follows. *Previously*, for the chosen dimension of the problem, a collection of these characteristics was created (in this case, we have always considered 70 cities); for them, we used various values σ, namely, from 0.00 to 2.20 with the step 0.01. Each of the characteristic values was taken equal to the average of the corresponding characteristics, obtained for 10 various random generations of special cases. At the same time, a self-study of the perceptron was carried out, which classified the special case under consideration on the basis of the obtained characteristic vector, i.e., giving an exit on the assumption of the original value σ.

Further, for each of the following values σ:

$$\sigma = 0.00\,;\quad \sigma = 0.20\,;\quad \sigma = 0.50\,;\quad \sigma = 0.90\,;\quad \sigma = 1.40\,;\quad \sigma = 2.00$$

[1] Note that *in practice* the most interesting application of the pseudo-geometric version of TSP in the case when the dimension of the problem is approximately 100. (The complexity of the algorithm used to calculate all these characteristics is $O(n^3)$, where, as before, n is the dimension of the problem.)

[2] The elements of a collection can be repeated.

[3] As we said before, we considered a symmetric TSP.

[4] For this and the following characteristics, it hardly makes sense to consider their analogues for the collection D.

we made 100 experiments. In each of them, we generated a special case of TSP, and after that *based on the characteristic vector obtained for this special case*, we selected the assumption of the initial value σ in two ways:

- based on the maximum probability value given by a previously trained perceptron;
- using the method of least squares.

The results of the experiments are summarized in the following Table 1. In it, the first column (P) corresponds to the results obtained with the perceptron. Each cell contains a segment of the values σ, into which the results were obtained, as well as the number of cases (from 100 observed), in which the value σ differs from the initial value by not more than 10%.

Table 1.

σ	P	MLS
0.00	$[0.00, 0.00]$ 100	$[0.00, 0.00]$ 100
0.20	$[0.16, 0.23]$ 83	$[0.16, 0.24]$ 81
0.50	$[0.43, 0.59]$ 81	$[0.42, 0.58]$ 80
0.90	$[0.76, 1.03]$ 79	$[0.76, 1.02]$ 79
1.40	$[1.20, 1.57]$ 78	$[1.22, 1.55]$ 78
2.00	$[1.74, 2.20]$ 79	$[1.76, 2.20]$ 79

Thus, apparently, we can say that the *both* versions of the classification give acceptable results, which practically do not depend on the specific method of classification (i.e., using a neural network or using the method of least squares). It is also important that both algorithms can be considered a *combination* of:

- the *rule-based approach*;
- the *alternative-based approach*, for which there is as yet no final name. [5]

A rule-based approach is to select the characteristics of the *expert*, and an alternative approach is to a posteriori use of the neural network or the method of least squares.

However, both the more detailed results of the above computations and the possible improvement of both classification algorithms that we have applied are hardly of great interest: as already noted, these calculations (that is, the classification) are regarded as an auxiliary task for another, more important task (i.e., the continuation of the *geometrical solving* of pseudo-geometric TSP), which we propose to consider in subsequent publications.

5. Conclusion

Let us mention another difficult problem for the following solution. For different sub-classes of problems, we try to use different genomes (using different genetic algorithms for the self-learning is also possible); they are analogues of classes of positions in nondeterministic games. But this is a simple problem, rather, a problem, which main complication belongs not to the programmer, but to the experts (who understand the whole specificity of the considered DOP). There would be much more important a possibility of automatic generating conditions for belonging some DOP to a class of problems, which should be considered separately from other classes (let us

[5] The final name is not in the literature, not only in Russian, but also in English: the often used words “not rule-based approach” are unlikely to be “claimed” for it. Possible alternatives to a rule-based approach, *in different subject areas*, are the so-called. a connectionism approach, a structural approach, and, for example, in the case of machine translation, the so-called statistical translation..

In connection with the last, it should be noted its long-standing use in the service “Yandex.Translate” [32], as well as the emergence of recently *hybrid* translation systems, actually being a *combination* of the two above-mentioned approaches.

also remark, that in the case of programming intelligent games, this problem is connected with a problem of automatic generating a new parameter for the function of static estimating position). After such automatic generating we use the usual self-learning by some genetic methods, and then we make a testing, whether we really have constructed a new class of sub-problems for considered DOP. Let us remark, that the possible algorithms of such testing are evident (they are similar to usual algorithms of clustering), and the first part of this problem, i.e., automatic generating conditions of a class of problems, is much more important.

Let us remark also the following fact. Using heuristics of this paper in various DOP, the authors have none example, when the optimal solution needs more than 5% of the time required for the common solution of the considered problem. This fact marginally explains all the heuristics considered in this paper, even the optimum solution is unknown, e.g., if it cannot be obtained in the reasonable time.

Here are the links to recent papers related to the subject matter discussed in this paper:[33, 34]. We also cite some recent works by the authors of this paper [35, 36, 37, 38, 39, 40].

6. References

[1] Hromkovič J 2004 *Algorithmics for Hard Problems. Introduction to Combinatorial Optimization, Randomization, Approximation, and Heuristics* (Berlin: Springer) p 548
[2] Gutin G and Punnen A 2002 The Traveling Salesman Problem and its Variations (Berlin: Springer) p 829
[3] Applegate D, Bixby R, Chv´atal V and Cook W 2007 *The Traveling Salesman Problem. A Computational Study* (Princeton: Princeton University Press) p 608
[4] Gutin G, Karapetyan D 2010 A memetic algorithm for the generalized traveling salesman problem *Natural Computing* **9(1)** 47-60
[5] Raman G and Gill N 2017 Review of different heuristic algorithms for solving Travelling Salesman Problem *Int. J. of Advanced Research in Computer Science* **8(5)** 423-449
[6] Kameda T and Weiner P 1970 On the state minimization of nondeterministic finite automata *IEEE Transactions on Computers* **C-19** 617-627
[7] Hashiguchi K 1991 Algorithms for determining the smallest number of nonterminals (states) sufficient for generating (accepting) a regular language *Automata, Languages and Programming Lecture Notes in Computer Science* **510** 641-648
[8] Jiang T and Ravikumar B 1993 Minimal NFA problems are hard *SIAM J. Comput.* **22(6)** 1117-1141
[9] Melnikov B 2000 Once more about the state-minimization of the nondeterministic finite automata *Journal of Applied Mathematics and Computing* **7(3)** 655-662
[10] Pol´ak L 2005 Minimizations of NFA using the universal automaton I*nternational Journal of Foundations of Computer Science* **16(5)** 999-1010
[11] Han Y-S 2013 State elimination heuristics for short regular expressions *Fundamenta Informaticae* **128** 445-462
[12] Melnikov B 2010 Once more on the edge-minimization of nondeterministic finite automata and the connected problems *Fundamenta Informaticae* **104(3)** 267-283
[13] Hashiguchi K 1988 Algorithms for determining relative star height and star height *Information and Computation* **78** 124-169
[14] Kirsten D 2005 Distance desert automata and the star height problem *Theoretical Informatics and Applications* **39** 455-509
[15] Bojanczyk M 2015 Star Height via Games *30th Annual ACM/IEEE Symposium on Logic in Computer Science* **104(3)** 214-219
[16] Melnikov B and Panin A 2012 On a parallel implementation of the multi-heuristic approach in the problem of comparison of genetic sequences *Vector Science of Togliatti State University* **4(22)** 83-86 (in Russian)
[17] Makarkin S and Melnikov B 2013 Geometrical methods of solving pseudo-geometrical version of traveling salesman problem *Stochastic optimization in informatics* **9(2)** 54-72 (in Russian)

[18] Liew S 2012 Introducing convex layers to the Traveling Salesman Problem *Preprint arXiv:1204.2348*
[19] Somhom S, Modares A and Enkawa T 1999 Competition-based neural network for the multiple travelling salesmen problem with minimax objective *Computers & Operations Research* **26(4)** 395-407
[20] Dorigo M and Gambardella L 1997 Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem *IEEE Transactions on Evolutionary Computation* **1(1)** 53-66
[21] Johnson D and McGeoch L 1997 The Traveling Salesman Problem: A Case Study in Local Optimization *Local Search in Combinatorial Optimization* 215-310
[22] Berliner H 1980 Computer Backgammon *Scientific American* **243** 64–72
[23] Tesauro G and Sejnowski T 1989 A Parallel Network that Learns to Play Backgammon *Artificial Intelligence* **39** 357-390
[24] Tesauro G 1985 Temporal Difference Learning and TD-Gammon *Temporal Difference Learning and TD-Gammon* **38(3)** 58-68
[25] Melnikov B and Radionov A 1998 A choice of strategy in nondeterministic antagonistic games *Programming and Computer Software* **24(5)** 247-252
[26] Melnikov B 2001 Heuristics in programming of nondeterministic games *Programming and Computer Software* **27(5)** 277-288
[27] Gurevich Y and Veanes M 2007 Can abstract state machines be useful in language theory? *Theoretical Computer Science (Developments in Language Theory)* **376(1-2)** 17-29
[28] Melnikov B, Pivneva S and Rogova O 2010 The representativeness of randomly generated nondeterministic finite automata from the point of view of the corresponding basis automata *Stochastic optimization in informatics* **6(1)** 74-82 (in Russian)
[29] Watson G 1962 Goodness-of-fit tests on a circle *Biometrika* **48(1-2)** 109-114, **49(1-2)** 57-63
[30] Lagutin M 2012 *Visual mathematical statistics* (Moscow: Binom) p 472 (in Russian)
[31] Haikin S 2008 *Neural networks: the full course* (Moscow: Williams) p 1104 (in Russian)
[32] Yandex.Translate (Access mode: https://en.wikipedia.org/wiki/Yandex.Translate)
[33] Evdokimova N I and Kuznetsov A V 2017 Local patterns in the copy-move detection problem solution *Computer Optics* **41(1)** 79-87 DOI: 10.18287/2412-6179-2017-41-1-79-87
[34] Evsutin O O, Shelupanov A A, Meshcheryakov R V and Bondarenko D O 2017 An algorithm for information embedding into compressed digital images based on replacement procedures with use of optimization *Computer Optics* **41(3)** 412-421 DOI: 10.18287/2412-6179-2017-41-3-412-421
[35] Melnikov B and Dudnikov V NFA project (Access mode: https://github.com/va-dudnikov/nfa)
[36] Melnikov B, Korabelshhikova S and Churikova N 2017 On verification algorithms for some binary relations on the general supermonoid of a free monoid *Izvestiya of Higher Educational Institutions. Volga Region. Physics and Mathematical Sciences* **3(43)** 87-99 (in Russian)
[37] Melnikov B and Dudnikov V 2018 Problem of pseudo-optimal placement on the graph and one heuristic approach of its solution *Informatization and Communication* **1** 63-70 (in Russian)
[38] Melnikov B and Zubova T 2018 Mathematical modeling of organization management by value guidelines: algorithms for complex estimation and selection of pseudo-optimal actions *International Journal of Open Information Technologies* **3** 1-8 (in Russian)
[39] Melnikov B and Davydova E 2018 Mathematical modeling of increasing the level of safety in case of failures of space technology *International Journal of Open Information Technologies* **5** 1-6 (in Russian)
[40] Melnikov B and Trenina E 2018 On a problem of the reconstruction of distance matrices between DNA sequences *International Journal of Open Information Technologies* **6** 1-13 (in Russian)

Acknowledgements
The reported study was partially supported by the research project of Russian State Social University.

The using of fractal measures for network state monitoring and probabilistic network attack type determination

O Yu Gubareva[1], O V Osipov[1], A O Pocheptsov[1] and V V Pugin[1]

[1]Povolzhskiy State University of Telecommunications and Informatics, L. Tolstoy Street 23, Samara, Russia, 443010

Abstract. In the academic paper for network traffic analysis concerning risk assessment of network nodes infosecurity fractal analysis is used which takes into consideration system past history and makes it possible to randomly determine probable network attack types on the system of interest. There has been developed a network traffic analysis technique based on fractal measures set with a focus on network state analysis and probabilistic attack type determination. Following on from the thesis results there is possible the creation of network traffic analyzer (sniffer) for time estimate of infosecurity state as well as further computation of previously attacked devices and network nodes.

1. Introduction

Any organization when operating is liable to various infosecurity risks which one way or another affect particular business processes characteristics and can negatively influence on financial data as well as the opportunity for the organization to go on its activity. Current business requirements necessitate using well-grounded technical-and-economic methods and means in operation providing quantitative and qualitative infosecurity (IS) level determination both with assessing infosecurity cost efficiency. For the purpose of efficient organization infosecurity a serious, systematic and integrated approach is required.

Almost any infosecurity system building must start form risks analysis. Before infosecurity system designing one should specify what threatening (in other words conditions and factors which can become the reason for breaking system integrity, its security and privacy and also facilitating unauthorized access to it) exists for the given infosystem and to what extent it is potentially critical.

Telecommunications networks have numerous vulnerabilities arising both in system software development and in misconfiguration and equipment operation. The presence of security threat makes it possible for intruders to put into operation various types of network attack. Nowadays software tools development for infosecurity risks analysis by means of network traffic online analysis is of great interest. Clearing up possible threat aims makes the basis for providing safety-related system design. The threat aims show what should be protected. As a rule network state is analyzed with a focus on network administration problem solving, routing device monitoring, etc. To become aware of abnormal system behavior there is often used various statistic information collection and analysis via IP-traffic. In this paper for network traffic data accessing a free given software Zabbix under GNU GPL license was used. The

monitoring system in this case builds software set for current traffic measuring and software system of its analysis constructed with computing entries so called fractal measures which will be specified in the given paper.

2. The research objective and solution method

The research objective is network traffic analysis technique development based on fractal measures set aimed at network state analysis and probabilistic attack type determination. Resting upon the research described in this paper the authors are planning to create a network traffic analyzer (sniffer) currently left on a company server for its time estimate, consecutive defining of previously attacked devices and network nodes (network vulnerabilities) and as a result further IS risks assessment.

The paper [1] gives an overview of scientific research in the field of analysis real-time network traffic, and specific hardware and software solutions are considered.

In the work [2] the use of the Hurst index for the analysis of the traffic subject to anomalous intrusions in the form of DoS-attacks is considered. The studies conducted in [2] showed that traffic has the property of self-similarity during abnormal intrusions, which proves the possibility of determining traffic anomalies in real time.

To take the set goal in the academic paper the following challenges are met: running the process analysis of the infosystem in question (the infosystem of an academic institution was taken as a basis) as an object to protect; Hurst exponent assessment, power-density spectrum and network traffic fractal measures in normal state and in the time of attack on the infosystem resources; executing the attack on the system resources.

Fractal analysis is statistic in its nature and in addition it gives the possibility to find self-similarity markers in the traffic of interest. The fact permits first to become aware of minimal required time for making the experiment. Second, it makes it possible to rely on the opportunity to forecast the system behavior dynamics in the nearest future. Fractal model is a set of fractal parameters (measures) put in accordance with the current network traffic state. The fractal measures changes dynamics when involving a series of measurements of one and the same telecommunication node lets us estimate traffic condition dynamics that is about the presence or absence of attacks on infosystem resources. Jumping ahead we can mention that as a result of the performed experiment it was brought to light that in case of DoS-attacks the self-similarity network traffic level reduces as well as there takes place power-density spectrum transformation.

The experiment idea is the following: there is some telecom traffic which is network load to timing dependency diagram (figure 1). From mathematical analysis perspective the traffic in question represents univariate time series the observations of which are channel occupancy levels at different moments. The current series can be analyzed with various fractal measures calculating (Hurst exponent, etc.) as well as power-density spectrum.

At the first stage Hurst exponent and power-density spectrum were calculated for the normal network condition.

First they determine Hurst exponent for calculating network traffic self-similarity level. For its determination they find average channel occupancy value $\langle U \rangle_{\mathrm{N}}$ for N tick marks [3, 4]:

$$\langle U \rangle_{\mathrm{N}} = \frac{1}{N} \sum_{n=1}^{N} U(n). \tag{1}$$

Then they define $X(n, N)$ which is accumulated divergency $U(n)$ from average value $\langle U \rangle_{\mathrm{N}}$, which is determined with the help of the following total (union):

$$X(n, N) = \sum_{p=1}^{n} \{U(p) - \langle U \rangle_{N}\}, \tag{2}$$

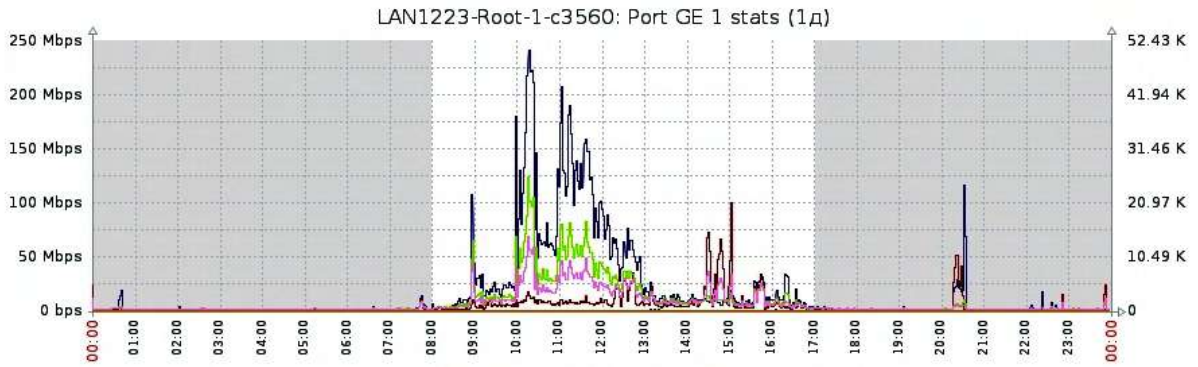


Figure 1. Network load to timing dependency diagram.

where the average value $\langle U \rangle_{\mathrm{N}}$ is defined by the formula (1).

According to standardized Hurst range [3, 4], the divergence range is determined via minimal and maximal values of the accumulated divergence $X(n, N)$ (2):

$$R(N) = \max_{1 \le n \le N} X(n, N) - \min_{1 \le n \le N} X(n, N). \tag{3}$$

The standard divergence $S\,(N)$ can be computed with the following known formula via dispersivity [3, 4]:

$$S\,(N) = \left\{ \frac{1}{N} \sum_{n=1}^{N} [U(n) - \langle U \rangle_N]^2 \right\}^{1/2}. \tag{4}$$

For most timing series the observed standardized range R/S is described by empiric relation and with the help of (3) and (4) appears as [3, 4]:

$$R/S = (\alpha N)^H, \tag{5}$$

where H is the Hurst exponent; α is an arbitrary parameter (constant).

The described procedure in scientific literature got the name of R/S-analysis.

In figure 2 there is shown R/S telecom traffic dependency in normal state upon N in log-log scale. The axis of ordinate shows the value of $\lg(R/S)$, on the x-axis — $\lg N$.

Hurst exponent value for the traffic in question in normal condition turned out to be equal 0.68. In accordance with the theory of fractals if the got Hurst exponent value $H < 0.5$ then the under study series has "short" memory. In other words it is antipersistent. It means that recent events in the begetter system produce much more influence on the following system behavior than less recent events. If $H > 0.5$ the timing series is persistent and has fractal nature. With the value $H = 0.5$ the signal represents stochastic noise and doesnt have any useful information. As can be seen from the above, it was proved that the traffic in question in normal state is self-similar and has fractal nature.

Further in the paper there was made power-density spectrum estimation which represents rapid inverse Fourier transform of autocorrelation function.

The network traffic autocorrelation function is determined by the following formula:

$$R\,(j) = \frac{1}{N} \sum_{i=1}^{N-j} U\,(i)\,U\,(i+j), \tag{6}$$

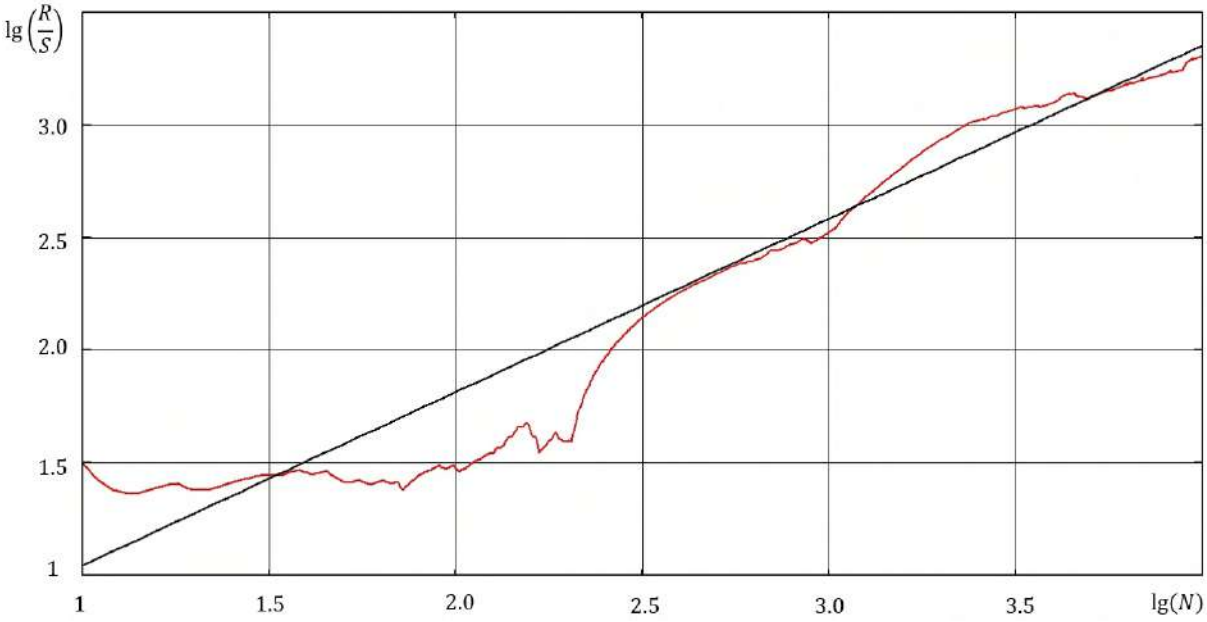


Figure 2. The R/S dependency of telecom traffic on the number of timing counts.

where N is the total number of network traffic tic marks. The signal power-density spectrum is defined by direct inverse Fourier transform of autocorrelation function (6):

$$S_k = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{p=1}^{N-i} U(p)\, U(p+i) \exp\left[-\mathrm{j}\frac{2\pi k i}{N}\right], \ \left(k = \overline{0, N-1}\right). \tag{7}$$

In figure 3 there is introduced network traffic power-density spectrum $S(f) = S_k(U)$ in normal state (with no network attack).

On the second stage there were studied fractal measures and network traffic power-density spectrum with DoS-attack.

During the DoS attack, the channel was fully loaded at 70 MB per second. It is worth noting here that the use of fractal measures (in particular, the parameter R/S) allows to guarantee the scalability of the obtained results in the case of higher channel utilization.

For this a before vulnerable web-system which before-known IP-address was developed. To perform DoS-attack there was used the software which is similar to LOIC program that allows to execute an attack of the given in advance IP-address with variable transactions amount. In addition to that simultaneously with this there was executed an attack on MySQL-server using SQL-injection implementation through get-parameter of the vulnerable system.

To do that they used an enquiry with SQL-function benchmark (n, q) that gives the possibility to do n times function q [5].

For attacking SQL-server there was written a script which given number of times issued such requests in cycle. After executing DoS-attack network traffic was taken during its time which was again analyzed about fractal measures and power-density spectrum. Hurst exponent for the traffic in question in the time of attack equaled 0.54 that speaks of sharp decrease self-similarity level of the traffic of interest.

Currently, experiments are being conducted on the backbone network with a load of 1.2 GB

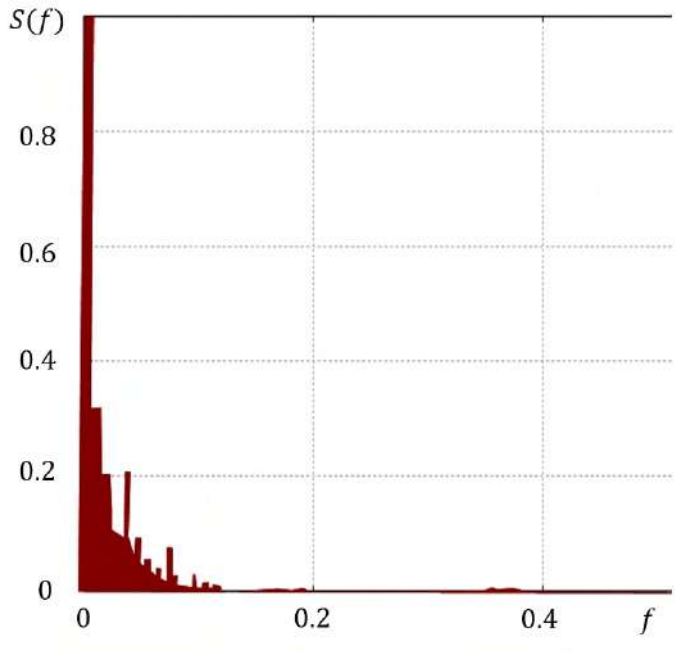


Figure 3. Network traffic power-density spectrum in normal condition.

per second with a time sample duration of 24 hours (86,000 calculated values of the channel load).

In figure 4 there is shown power-density spectrum for the case in question which allows to visually classify the signal in question as "brown" noise.

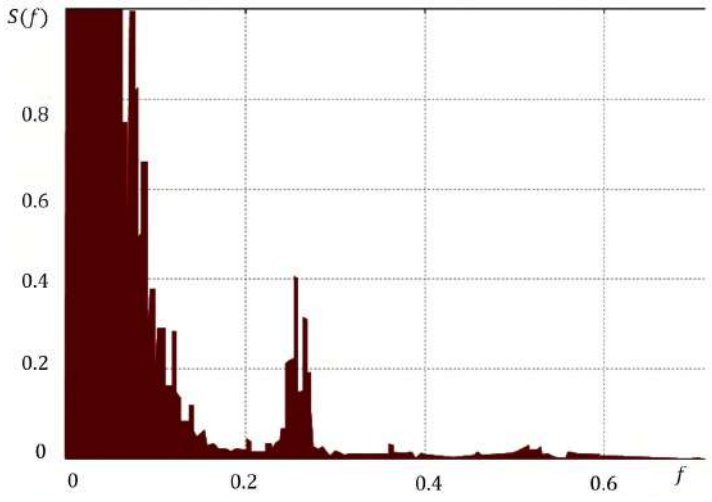


Figure 4. Network traffic power-density spectrum at the time of attack.

Consequently as a result of the experiment in real network fractal measures changing and power-density spectrum were proved with DoS-attack.

Not without interest are studying of fractal measures and network traffic power-density

spectrum while various network attacks are that can lead to creation of some on-line "patterns" database (library) of power-density spectra and fractal measures values. In other words it is referred to the opportunity to make some fractal network-status indicator for the time of high probability to determine the threat type. Worth making a point in this regard is that subtle fractal analysis allows to reveal minimal traffic changing despite full channel occupancy in case of network attack. However, here we need further experimental research aimed at revealing specific attacks and building "patterns" of fractal characteristics.

Note that Hurst exponent computing of network traffic with 10000 time samples takes around 1,5 seconds when using Intel Core i5 and power-density spectrum calculation takes about 4 seconds.

3. Conclusion

In conclusion, let us dwell on the main conclusions of the work done. Fractal network indicator led has to perform the following functions:

- saving channel occupancy entries sampling for certain time windows that are enough for network condition diagnostics;
- fractal measures and power-density spectrum calculation for every set timing series for the definite interval times for the purpose of further comparison with ¡¡patterns¿¿ from the information base (using neural networks);
- finding of network state totally in the current and precedent time points;
- probability forecast about the network attack nature in the future.

Thus, in the paper for network traffic state analysis (including DoS-attacks) there is offered to use fractal measures and power-density spectrum which allow by indirect hints for agreeable time limit to determine threat level.

The algorithms developed in this work may be useful for the analysis of "smartlink connections" [6]. Another object of the fractal technique is the stochastic network [7].

In conclusion, we note that the proposed method is the basis for creating a fractal indicator for analyzing the state of the network, while specialized software (iptables, ipwf, etc.) should be used to determine the sources of the DoS attack.

4. References

[1] Get'man A I, Markin Yu V, Evstropov E F and Obydenkov D O 2017 Analysis of network traffic in the mode real-time: overview of applied tasks, approaches and solutions *Trudy ISP RAN* **29(3)** 117-150 (in Russian)

[2] Shelukhin O I and Antonyan A A 2014 Analysis of changes in the fractal properties of telecommunications traffic caused by abnormal intrusions *T-COMM: Telecommunications and transport* **8(6)** 61-64 (in Russian)

[3] Feder J 1991 *Fractals* (Springer Science + Business Media, LLC) 305 p

[4] Golovko V A 2005 Neural network methods for processing chaotic processes *VII All-Russian scientific-technical conference "Neuroinformatics"* 43-91 (in Russian)

[5] Nizamutdinov M F 2005 *The tactics of protecting and attacking WEB applications* (SPb.: BHV-Peterburg Publisher) p 432 (in Russian)

[6] Nikitin V S, Semyonov E I, Solostin A V, Sharov V G and Chayka S V 2016 Modeling the "smartlink connection" performance *Computer Optics* **40(1)** 64-72 DOI: 10.18287/2412-6179-2016-40-1-64-72

[7] Agafonov A A, Myasnikov V V 2016 Method for the reliable shortest path search in timedependent stochastic networks and its application to GIS-based traffic control *Computer Optics* **40(2)** 275-283 DOI: 10.18287/2412-6179-2016-40-2-275-28

Development of a knowledge base based on context analysis of external information resources

N Yarushkina[1], A Filippov[1] and V Moshkin[1]

[1]Ulyanovsk State Technical University, Severny Venetz str. 32, Ulyanovsk, Russia, 432027

Abstract. The article describes the process of developing a knowledge base (KB). The content of KB is formed as a result of the analysis of the contexts of external information resources. In this case, the context is a certain "point of view" on the problem area (PrA) and its features. A graph database (DB) Neo4j is used as the basis for storing the contents of the KB in the form of an ontology. An attempt is made to implement the mechanism of inference by the contents of a graph database. The mechanism is used to dynamically generate the screen forms of the user interface to simplify the work with the KB. This article also describes the method of extension of KB based on the content of the wiki-resources and relational databases.

1. Introduction

Post-industrial society operates with huge volumes of information both in everyday and professional activities. A large amount of information causes difficulties in making decisions within the framework of rigid time constraints.

A variety of software automation of human activities are used to solve this problem. However, it is necessary to adapt them to the specifics of a particular problem area (PrA) and its contexts for the effective operation of these tools [1, 2, 7, 10, 18, 19, 20].

Thus, "trained" automation software solves the tasks more efficiently, but they require considerable resources (human and temporary) for training.

In this paper, an attempt is made to construct a KB. The content of the KB is an applied ontology. The basic requirements for KB are [26]:

- adaptation to the specifics of PrA based on contexts;
- reliability and speed of ontology storage;
- the presence of a mechanism of logical inference;
- availability of tools to simplify work with the KB for unprepared users;
- availability of mechanisms for importing data from external information resources.

As you can see from Figure 1, the KB consists of the following subsystems:

(i) Ontology store:
  - Neo4j [12];
  - content management module;
  - ontology import/export module.

(ii) Inference subsystem:

- inference module.

(iii) A subsystem for interaction with users:

- screen forms generation module.

(iv) A subsystem for importing data from wiki-resources:

- a module for importing data from wiki-resources.

(v) A subsystem for importing data from relational databases:

- a module for importing data from relational databases.

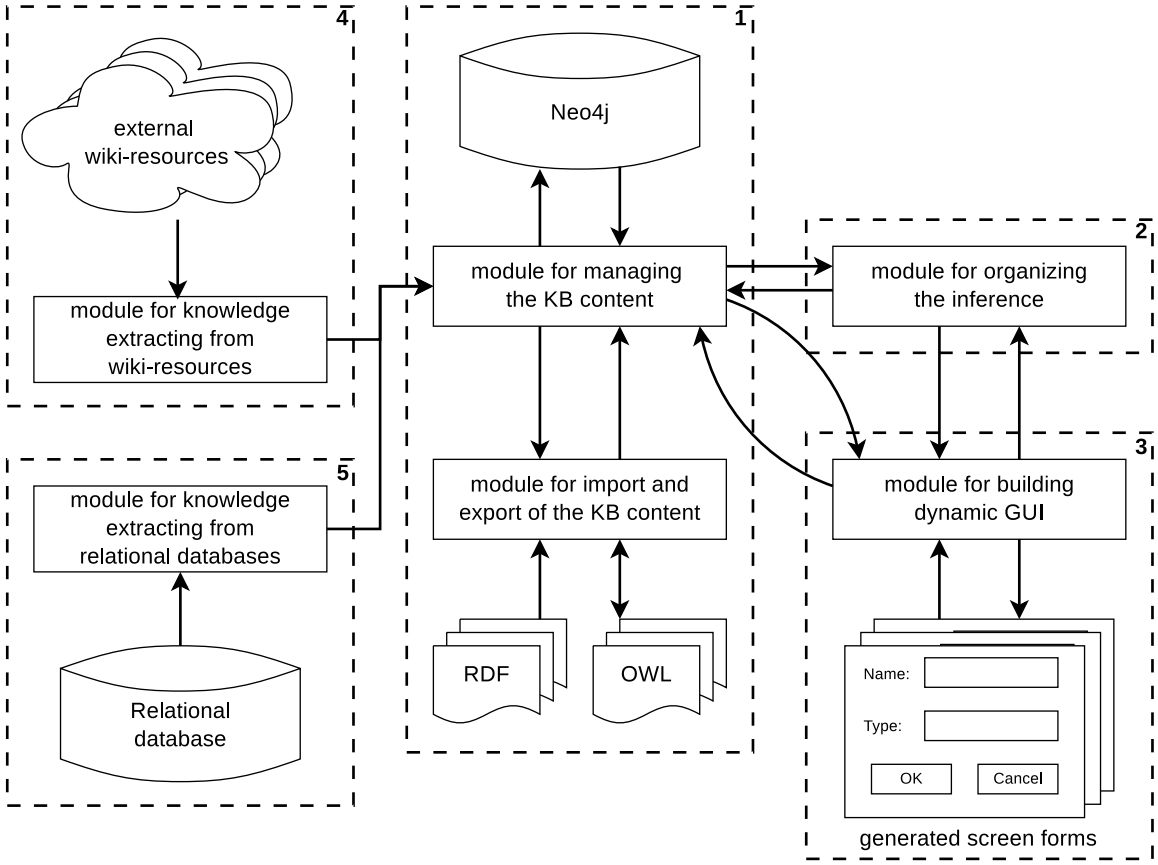


Figure 1. Knowledge base architecture.

2. The organization of the ontology store of KB

Ontology is a model of the representation of the PrA in the form of a semantic graph [9].

Graph-oriented database management system (Graph DBMS) Neo4j is the basis of the ontology store for KB. Neo4j is currently one of the most popular graph databases and has the following advantages:

(i) Having a free community version.

(ii) Native format for data storage.

(iii) One copy of Neo4j can work with graphs containing billions of nodes and relationships.

(iv) The presence of a graph-oriented query language Cypher.

(v) Availability of transaction support.

Neo4j was chosen to store the description of the PrA in the form of an applied ontology, since the ontology is actually a graph. In this case, it is only necessary to limit the set of nodes and graph relations into which ontologies on RDF and OWL will be translated.

The context of an KB is some state of content of KB, obtained during versioning or building a content of KB using different "points of view" [6, 8].

Figure 2 shows an example of the translation of the owl representation of ontology of family relations into the entities of the KB.

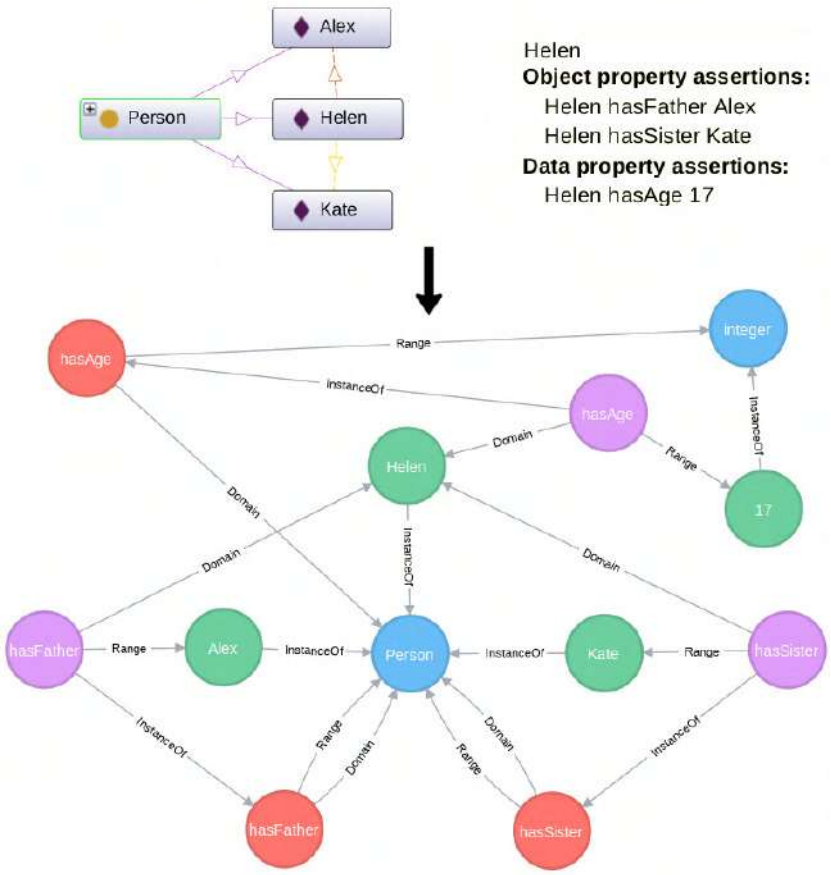


Figure 2. Example of the translation of the owl representation of ontology of family relations into the content of the KB.

Formally, the content of the KB can be represented by the following equation:

$$O = \langle T, C^{T_i}, I^{T_i}, P^{T_i}, S^{T_i}, F^{T_i}, R^{T_i} \rangle, i = \overline{1,t}, \tag{1}$$

where t is a number of the KB contexts,
$T = \{T_1, T_2, \ldots, T_t\}$ is a set of KB contexts,
$C^{T_i} = \{C_1^{T_i}, C_2^{T_i}, \ldots, C_n^{T_i}\}$ is a set of KB classes within the i-th context,
$I^{T_i} = \{I_1^{T_i}, I_2^{T_i}, \ldots, I_n^{T_i}\}$ is a set of KB objects within the i-th context,
$P^{T_i} = \{P_1^{T_i}, P_2^{T_i}, \ldots, P_n^{T_i}\}$ is a set of KB classes properties within the i-th context,
$S^{T_i} = \{S_1^{T_i}, S_2^{T_i}, \ldots, S_n^{T_i}\}$ is a set of KB objects states within the i-th context,
$F^{T_i} = \{F_1^{T_i}, F_2^{T_i}, \ldots, F_n^{T_i}\}$ is a set of the logical rules fixed in the KB within the i-th context,

R^{T_i} is a set of KB relations within the i-th context defined as:

$$R^{T_i} = \{R_C^{T_i}, R_I^{T_i}, R_P^{T_i}, R_S^{T_i}, R_F^{T_i}\},$$

where $R_C^{T_i}$ is a set of relations defining hierarchy of KB classes within the i-th context,
$R_I^{T_i}$ is a set of relations defining the "class-object" KB tie within the i-th context,
$R_P^{T_i}$ is a set of relations defining the "class-class property" KB tie within the i-th context,
$R_S^{T_i}$ is a set of relations defining the "object-object state" KB tie within the i-th context,
$R_F^{T_i}$ is a set of relations generated on the basis of logical KB rules in the context of i-th context.

Principles similar to the paradigm of object-oriented programming are at the basis of the content of the KB:

- KB classes are concepts of the PrA;
- classes can have properties, the child-class inherits properties of the parent class;
- objects of KB describe instances of the concepts of the PrO;
- specific values for the properties of objects inherited from the parent class are determined by the states;
- logical rules are used to implement the functions of inference by the content of KB.

3. The inference on the contents of KB

The inference is the process of reasoning from the premises to the conclusion. Reasoners are used to implement the function of inference. Reasoners form logical consequences on the basis of many statements, facts and axioms. The most popular at the moment reasoners are [5, 17]:

- Pellet;
- FaCT++;
- Hermit;
- Racer, etc.

These reasoners are actively used in the development of intelligent software. However, Neo4j does not assume the possibility of using similar default reasoners. Thus, there is a need to develop a mechanism for inference based on the content of a KB [3, 4].

Currently the Semantic Web Rule Language (SWRL) is used to record logical rules [24].

These SWRL rules describe the conditions under which object a has "nephew-uncle" relation with object c. Formally the logical rule of the KB is:

$$F^{T_i} = \langle A^{Tree}, A^{SWRL}, A^{Cypher} \rangle,$$

where T_i is the i-th context of the the KB, A^{Tree} is the tree-like representation of a logical rule F^{T_i}, A^{SWRL} is the SWRL representation of the logical rule F^{T_i}, A^{Cypher} is the Cypher representation of the logical rule F^{T_i}.

The tree-view A^{Tree} of a logical rule F^{T_i} is:

$$A^{Tree} = \langle Ant, Cons \rangle,$$

where $Ant = Ant_1 \Theta Ant_2 \Theta \ldots \Theta Ant_n$ is the antecedent (condition) of the logical rule F^{T_i};
$\Theta \in \{AND, OR\}$ is a set of permissible logical operations between antecedent atoms;
$Cons$ is the consequent (consequence) of a logical rule F^{T_i}.

Figure 3 shows an example of a tree-like representation of two logical rules for the ontology of family relations. That rules describes the father-child relationships.

The tree-like logical rule is translated into the following SWRL:

```
hasFather(?a,?b) => hasChild(?b,?a)
hasSister(?c,?a) & hasFather(?c,?b) => hasChild(?b,?a)
```

and the following Cypher view:

```
MATCH (s1:Statement{name: "hasChild", lr: true})
MATCH (r1a)<-[:Domain]-(:Statement{name:"hasFather"})-[:Range]->(r1b)
MERGE (r1b)-[:Domain]->(s1)
MERGE (r1a)-[:Range]->(s1)

MATCH (s1:Statement{name: "hasChild", lr: true})
MATCH (r2c)<-[:Domain]-(:Statement{name:"hasSister"})-[:Range]->(r2a)
MATCH (r2c)<-[:Domain]-(:Statement{name:"hasFather"})-[:Range]->(r2b)
MERGE (r2b)-[:Domain]->(s1)
MERGE (r2a)-[:Range]->(s1)
```

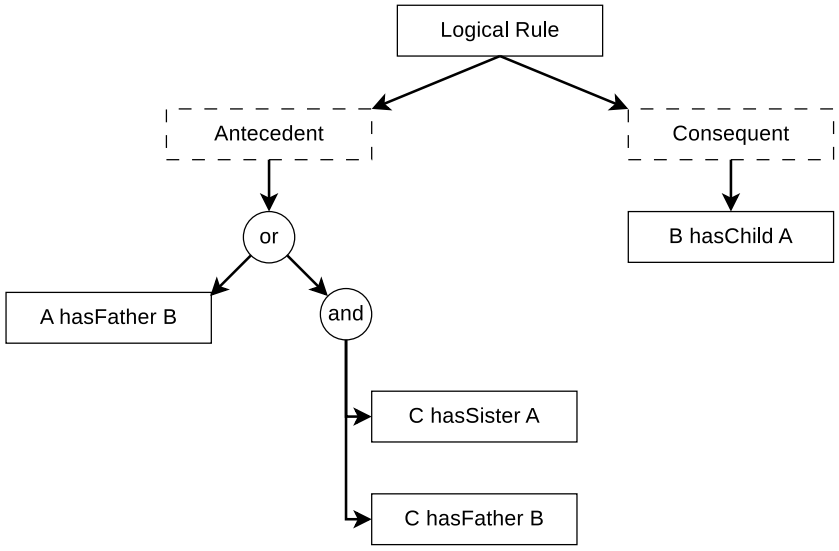


Figure 3. Example of a tree-like representation of a logical rule.

Thus, the rules are translated into their tree-view when imported into the KB of logical rules in the SWRL language.

The presence of a tree-like representation of a logical rule allows to form both a SWRL-representation of a logical rule and a Cypher-representation based on it.

Relations of a special type are formed by using Cypher to represent the logical rule between entities of the KB. Figure 4 shows the content of KB after executing the Cypher queries that were built for the logical rule shown in Figure 3. These relations correspond to the antecedent atoms of the logical rule. Formed relationships provide the inference from the contents of the KB.

4. Building a Graphical User Interface based on the contents of a KB

The dynamic graphical user interface (GUI) mechanism is used to simplify the work with KB of untrained users and control of user input [11, 13, 21].

You need to map the KB entities to the GUI elements to build a GUI based on the contents of the KB. Formally, the GUI model can be represented as follows:

$$UI = \langle L, C, I, P, S \rangle, \tag{2}$$

where $L = \{L_1, L_2, \ldots, L_n\}$ is a set of graphical GUI components (for example, ListBox, TextBox, ComboBox, etc.), $C = \{C_1, C_2, \ldots, C_n\}$ is a set of KB classes, $I = \{I_1, I_2, \ldots, I_n\}$ is a set of KB objects, $P = \{P_1, P_2, \ldots, P_n\}$ is a set of properties of KB classes, $S = \{S_1, S_2, \ldots, S_n\}$ is a set of states of KB objects.

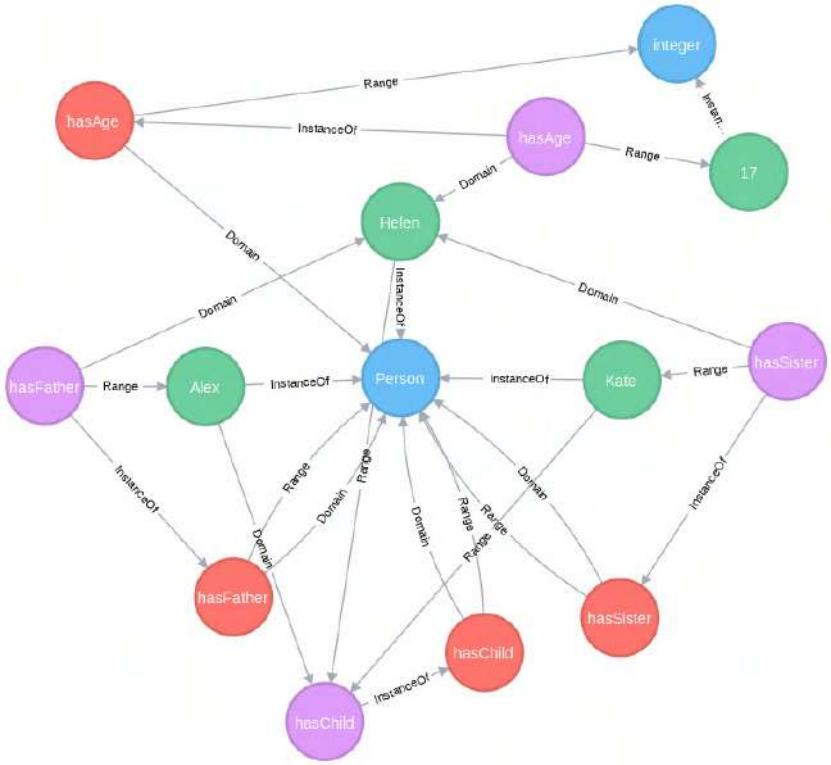


Figure 4. The result of executing Cypher queries for logical rule.

The following function is used to build a GUI based on content of KB:

$$\phi(O) : \{C^O, I^O, P^O, S^O, F^O, R^O\}^{T_i} \rightarrow \{L^{UI}, C^{UI}, I^{UI}, P^{UI}, S^{UI}\},$$

where $\{C^O, I^O, P^O, S^O, F^O, R^O\}^{T_i}$ is a set of entities of KB represented by expression 1 within the i-th context;
$\{L^{UI}, C^{UI}, I^{UI}, P^{UI}, S^{UI}\}$ is a set of GUI entities of KB represented by the expression 2.

Thus, the contents of the KB are mapped to set of GUI components. This makes it easier to work with KB for a user who does not have skills in ontological analysis and knowledge engineering. It also allows you to monitor the logical integrity of the user input, which leads to a reduction in the number of potential input errors.

5. Extracting knowledge from wiki-resources

At present, wiki-technologies are used to organize corporate KB. It is necessary to solve the task of knowledge extracting from wiki-resources [14, 15, 16, 23, 27]. Table 1 contains the result of mapping the KB entities to the wiki-resource entities [22]. Thus, it becomes possible to import the structure of external wiki resources for initial filling of the KB contents.

Table 1. The correspondence between the wiki-resource entities and the entities of KB.

The entities of knowledge base	The entities of wiki-resource
Class	Category
Subclass	Subcategory
Object	Page
Class properties	The infobox elements (properties)
Object states	The infobox elements (values)
Relations	Hyperlinks

Also, a content of KB can be built on the basis of an analysis of the content of wiki-resources pages. In this work the Syntaxnet [22] framework to construct a syntactic tree $Synt$ of content of wiki-resources pages is used. Further, using a set of rules $Rule^{Synt}$, a syntax tree $Synt$ is translated into entities of KB.

Formally the functions of translating a syntactic tree into entities of KB:

$$\phi^{Struct}(Synt) : \{N_{Synt}, Rule^{Struct}_{Synt}\} \rightarrow \{C^O, P^O, R^O_P\}^{T_i},$$
$$\phi^{Content}(Synt) : \{N_{Synt}, Rule^{Content}_{Synt}\} \rightarrow \{I^O, S^O, R^O_I, R^O_S\}^{T_i},$$

where N_{Synt} is a set of nodes of the syntactic tree $Synt$,
$Rule^{Struct}_{Synt}$ is a set of rules to translating nodes of syntactic into structure entities of the KB,
$Rule^{Content}_{Synt}$ is a set of rules to translating nodes of syntactic into content entities of the KB,
$\{C^O, P^O, R^O_P\}^{T_i}$ is a set of structure entities of the KB within the context T_i (eq. 1),
$\{I^O, S^O, R^O_I, R^O_S\}^{T_i}$ is a set of content entities of the KB within the context T_i (eq. 1).

Formally the rules to translating nodes of syntactic into entities of the KB:

$$Rule^{Struct}_{Synt} = \left(N^{Synt}_1, N^{Synt}_2, \ldots, N^{Synt}_i, \ldots, N^{Synt}_n\right) \rightarrow \{C^O, P^O, R^O_P\},$$
$$Rule^{Content}_{Synt} = \left(N^{Synt}_1, N^{Synt}_2, \ldots, N^{Synt}_i, \ldots, N^{Synt}_m\right) \rightarrow \{I^O, S^O, R^O_I, R^O_S\},$$

where N^{Synt}_i is the i-th node of syntactic tree.

Thus, it becomes possible to extract knowledge from the structure of wiki-resource and contents of wiki-resource pages and present the extracted knowledge as a content of KB.

6. Extracting knowledge from relational databases

Relational databases are widely used for data storing and contains subject area description in the form of interconnected tables. Nowadays, researchers of various scientific groups are involved in solving the problem of extracting knowledge from relational databases.

The relational data model can be represented as the following expression:

$$RDM = (E, R),$$

where $E = \{E_1, E_2, \ldots, E_n$ is a set of database tables (entities),
$R = \{R_1, R_2, \ldots, R_i, \ldots R_n$ is a set of relationships between database tables:

$$R_i = E_j \frac{F(x)}{G(x)} E_k,$$

where E_j, E_k are database entities;
$F(x)$ is the relationship between entity E_j and entity E_k,
$G(x)$ is the relationship between entity E_k and entity E_j.

Scope of functions $F(x)$ and $G(x)$ are U – single relationship and N – multiply relationship.

For mapping of relational database structure with KB structure special functions are used:

$$\phi^{Struct}(RDM) : \{E^{RDM}, R^{RDM}\} \rightarrow \{C^O, P^O, R_P^O\}^{T_i},$$
$$\phi^{Content}(RDM) : \{E^{RDM}, R^{RDM}\} \rightarrow \{I^O, S^O, R_I^O, R_S^O\}^{T_i},$$

where $\{E^{RDM}, R^{RDM}\}$ is a set of entities of relational database and relationships between them, $\{C^O, P^O, R_P^O\}^{T_i}$ is a set of structure entities of the KB within the context T_i (eq. 1), $\{I^O, S^O, R_I^O, R_S^O\}^{T_i}$ is a set of content entities of the KB within the context T_i (eq. 1).

Importing data from a relational database to the KB were finish after mapping the structure of the relational database to the set of structure entities $\{C^O, P^O, R_P^O\}^{T_i}$ of the KB ends. Set of content entities of KB $\{I^O, S^O, R_I^O, R_S^O\}^{T_i}$ are created during the import of data basis from the relational database (*row set*) to the T_i context. Table 2 contains a comparison of KB entities with relational database entities.

Table 2. The correspondence between the relational database entities and the entities of KB

The entity of knowledge base	The entity of relational database
Class	Table
Object	Table row
Class properties and Relations	Relations between tables, table columns
Object states	Content of cells

Thus, it becomes possible to extract knowledge from the contents of relational databases and present the extracted knowledge as a content of KB.

7. Conclusion

Thus, the use of KB stored in the Graph DBMS in the decision support process presupposes the existence of a certain set of mechanisms:

- organization of inference on the content of KB by translating SWRL-rules into Cypher-structures;
- building a graphical user interface based on the contents of KB;
- automated import of knowledge from structure and content of wiki-resources;
- automated import of knowledge from relational databases.

These mechanisms allow to automate the learning process of KB and simplify the work of specialists with KB. The application of a contextual approach to the storage of knowledge raises the effectiveness of the use of subject ontologies, allowing to adapt the KB to the characteristics of the PrA and to the requirements of specialists. This approach provides them with a tool that is convenient in a software dynamically changeable depending on the contents of the KB.

8. References

[1] Berant J, Chou A, Frostig R and Liang P 2013 Semantic parsing on freebase from question-answer pairs *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)* 1533-1544

[2] Bianchini D, De Antonellis V, Pernici B and Plebani P 2005 Ontology-based methodology for e-service discovery *Information Systems* **31** 361380
[3] Bobillo F and Straccia U 2008 FuzzyDL: an expressive fuzzy description logic reasoner *Proceedings of the 17th IEEE International Conference on Fuzzy Systems* 923-930
[4] Bobillo F and Straccia U 2010 Representing fuzzy ontologies in OWL 2 *Proceedings of the 19th IEEE International Conference on Fuzzy Systems* 2695-2700
[5] Dentler K, Cornet R, Aten Teije and N de Keizer 2011 Comparison of reasoners for large ontologies in the OWL 2 EL profile *Semant. web* **2** 7187
[6] Falbo R A, Quirino G K, Nardi J C, Barcellos M P, Guizzardi G and Guarino N 2016 An ontology pattern language for service modeling *Proceedings of the 31st Annual ACM Symposium on Applied Computing* 321-326
[7] Farid DM, Al-Mamun MA, Manderick B and Nowe A2016 An adaptive rule-based classifier for mining big biological data *Expert Systems with Applications* **64** 305316
[8] Gao M and Liu C2005 Extending OWL by fuzzy description logic *Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence* 562-567
[9] Guarino N and Musen M A 2015 Ten years of Applied Ontology *Applied Ontology* **10** 169170
[10] Guizzardi G, Guarino N, Almeida J P A 2016 Ontological Considerations About the Representation of Events and Endurants in Business Models *International Conference on Business Process Management* 20-36
[11] Hattori S and Takama Y 2014 Recommender System Employing Personal-Vallue-Based User Model *J Adv. Comput. Intell. Intell. Inform.* **18** 157165
[12] *Neo4j* (Access mode: https://neo4j.com/product) (14.05.2018)
[13] Ltifi H, Kolski C, Ayed M B and Alimi A M 2013 A human-centred design approach for developing dynamic decision support system based on knowledge discovery in databases *Journal of Decision Systems* **22** 6996
[14] Mikhaylov D V, Kozlov A P and Emelyanov G M 2015 An approach based on TF-IDF metrics to extract the knowledge and relevant linguistic means on subject-oriented text sets *Computer Optics* **39(3)** 429-438 DOI: 10.18287/0134-2452-2015-39-3-429-438
[15] Mikhaylov D V, Kozlov A P and Emelyanov G M 2016 Extraction of knowledge and relevant linguistic means with efficiency estimation for the formation of subject-oriented text sets *Computer Optics* **40(4)** 572-582 DOI: 10.18287/2412-6179-2016-40-4-572-582
[16] Mikhaylov D V, Kozlov A P and Emelyanov G M 2017 An approach based on analysis of n-grams on links of words to extract the knowledge and relevant linguistic means on subject-oriented text sets *Computer Optics* **41(3)** 461-471 DOI: 10.18287/2412-6179-2017-41-3-461-471
[17] *Pellet Framework* (Access mode: http://github.com/stardog-union/pellet) (14.05.2018)
[18] Rajpathak D, Chougule R and Bandyopadhyay P 2012 A domain-specific decision support system for knowledge discovery using association and text mining *Knowledge and Information Systems* **31** 405432
[19] Renu R S, Mocko G and Koneru A 2013 Use of Big Data and Knowledge Discovery to Create Data Backbones for Decision Support Systems *Procedia Computer Science* **20** 446453
[20] Rubiolo M, Caliusco ML, Stegmayer G, Coronel M, Fabrizi MG 2012 Knowledge discovery through ontology matching: An approach based on an Artificial Neural Network model *Information Sciences* **194** 107-119
[21] Ruy F B, Reginato C C, Santos V A, Falbo R A and Guizzardi G 2015 Ontology Engineering by Combining Ontology Patterns *34th International Conference on Conceptual Modeling* 173-186
[22] Shestakov V K 2011 Development and maintenance of information systems based on ontology and Wiki-technology *Advanced Methods and Technologies, Digital Collections* 299-306 (in Russian)

[23] Suchanek F M, Kasneci G and Weikum G 2007 YAGO: A Core of Semantic Knowledge Unifying WordNet and Wikipedia *Proceedings of the 16th International Conference on World Wide Web* 697706
[24] *SWRL: A Semantic Web Rule Language Combining OWL and RuleML* (Access mode: https://www.w3.org/Submission/SWRL) (14.05.2018)
[25] *SyntaxNet: Neural Models of Syntax* (Access mode: https://github.com/tensorflow/models/tree/master/research/syntaxnet) (14.05.2018)
[26] Yarushkina N, Filippov A and Moshkin V 2017 Development of the Unified Technological Platform for Constructing the Domain Knowledge Base Through the Context Analysis *Creativity in Intelligent Technologies and Data Science* 6272
[27] Zarubin A, Koval A, Filippov A and Moshkin V 2017 Application of Syntagmatic Patterns to Evaluate Answersto Open-Ended Questions *Creativity in Intelligent Technologies and Data Science* 150162

Acknowledgments
This work was financially supported by the Russian Foundation for Basic Research (Grant No. 16-47-732054).

Linear codes invariant with respect to generalized shift operators

V G Labunets[1] and E Ostheimer[2]

[1]Ural State Forest Engineering University, Sibirsky trakt 37, Ekaterinburg, Russia, 620100
[2]Capricat LLC, Pompano Beach, Florida, US

Abstract. The purpose of this paper is to introduce new linear codes with generalized symmetry. We extend cyclic and group codes in the following way. We introduce codes, invariant with respect to a family of generalized shift operators (GSO). In particle case when this family is a group (cyclic or Abelian), these codes are ordinary cyclic and group codes. They are invariant with respect to this group. We deal with GSO-invariant codes with fast code and encode procedures based on fast generalized Fourier transforms. The hope is that thesemore general structures will lead to larger classes of useful codes "good" properties.

1. Introduction

Let $\mathbf{F}$ be a finite field. A *block code* of length N is a subset C of $\mathbf{F}^N$, i.e., a collection of N length vectors with components from $\mathbf{F}$. Most of the literature on block codes pertains to block codes over finite fields $\mathbf{F}=\mathbf{GF}(q)$ or finite rings $\mathbf{F}=\mathbf{GR}(q)$, where $q=p^s$ and p is a prime. Although any subset forms a code, there are codes with more structure that are very useful and compose the majority of block codes in practice. A *linear block code* is a block code that is an $\mathbf{F}$-subspace of the $\mathbf{F}$-vector space $\mathbf{F}^N$. In addition to linearity, there are many structural properties that make for good codes. One of the most prevalent such structural properties is symmetry of code, that is described as invariance with respect to a group. Invariance (code symmetry), in many circumstances, leads to some nice encoding and decoding algorithms yet it is a very simple structure to describe. For these reasons, it is one of the most studied structural properties in coding theory.

Definition 1 [1,2]. *A cyclic block code* $\mathsf{C}\subset\mathbf{F}^N$ *of length* N *over a finite field* $\mathbf{F}$ *is a linear block code with the property that if* $(c_0,c_1,...,c_{N-2},c_{N-1})\in\mathsf{C}$ *then* $(c_{N-1},c_0,...,c_{N-3},c_{N-2})\in\mathsf{C}$.

It means that group of code symmetry of a cyclic code $\mathsf{C}\subset\mathbf{F}^N$ is $\mathbf{Symm}\{\mathcal{C}\}\approx\mathbf{Z}_N$. Cyclic codes are studied from many points of view. One way is to view them as ideals of an algebra. Define $\rho:\mathbf{F}^N\to\mathbf{F}[x]/\langle x^N-1\rangle$ via $\rho:(c_0,c_1,...,c_{N-2},c_{N-1})\mapsto c_0+c_1x+...+c_{N-2}x^{N-2}+c_{N-1}x^{N-1}$. It can be shown that ρ is an isomorphism. Let $\mathsf{C}\subset\mathbf{F}^N$ be cyclic block code. Then $\rho(\mathsf{C})$ is a subspace of the $\mathbf{F}$-vector space $\mathbf{F}[x]/\langle x^N-1\rangle$. Now the added condition of being cyclic translates to the following: if $\rho(c)\in\rho(\mathsf{C})$ then $x\cdot\rho(c)=x\cdot\left(c_0+c_1x+...+c_{N-2}x^{N-2}+c_{N-1}x^{N-1}\right)=\left(c_{N-1}+c_0x+c_1x^2+...+c_{N-2}x^{N-1}\right)\in\rho(\mathsf{C})$.
With this extra condition, $\rho(\mathsf{C})\lhd\mathbf{F}[x]/\langle x^N-1\rangle$.

There are many generalizations of cyclic codes, some of which may be viewed as ideals of particular rings [3]:

- negacyclic (skew-cyclic) codes[4-11]- ideal of the ring $\mathsf{A}lg^N(\mathbf{F})[x]/\langle x^N+1\rangle$,
- constacyclic codes [12]- ideal of the ring $\mathsf{A}lg^N(\mathbf{F})[x]/\langle x^N-\lambda\rangle$,where $\lambda\in\mathsf{A}lg(\mathbf{F})$,
- polycyclic codes [3]- ideal of the ring $\mathsf{A}lg^N(\mathbf{F})[x]/\langle f(x)\rangle$,where $f(x)\in\mathsf{A}lg(\mathbf{F})[x]$.

The terminology of the cyclic codes theory may be extended to define a larger family ofcodes. We start by introducing vector-induced clockwise and counterclockwise shifts. Given a vector $\mathbf{s}=(s_0,s_1,...,s_{N-2},s_{N-1})\in\mathbf{F}^N$, the $\mathbf{s}$-clockwise and $\mathbf{s}$-counterclockwise shifts of codeword $\mathbf{c}=(c_0,c_1,...,c_{N-2},c_{N-1})\subset\mathbf{F}^N$ are the following correspondences

$$R^{\mathbf{s}}\mathbf{c}=R^{\mathbf{s}}(c_0,c_1,...,c_{N-1})=(0,c_0,c_1,...,c_{N-2})+c_{N-1}(s_0,s_1,s_2,...,s_{N-1})=$$
$$=(c_{N-1}s_0,\ c_0+s_1c_{N-1},\ c_1+s_2c_{N-1},...,c_{N-2}+s_{N-1}c_{N-1}),$$
$$\mathbf{L}^{\mathbf{s}}\mathbf{c}=\mathbf{L}^{\mathbf{s}}(c_0,c_1,...,c_{N-1})=(c_1,c_2,...,c_{N-1},0)+c_0(s_0,s_1,s_2,...,s_{N-1})=$$
$$=(c_1+s_0c_0,\ c_2+s_1c_0,...,c_{N-1}+s_{N-2}c_0,s_{N-1}c_0).$$

Dyadic codes are defined only for length N, a power of 2, say $N=2^n$, as follows.

Definition 2. *For any integer $i\in\{0,1,2,...,N-1\}$, let $i=(i_{n-1},i_{n-2},...,i_1,i_0)$. Denote its radix-2 representation, where $i=i_{n-1}2^{n-1}+i_{n-2}2^{n-1}+...+i_12^1+i_02^0=\sum_{l=0}^{n-1}i_l2^l$ and $i_1\in\{0,1\}$ for $l=0,1,2,...,n-1$. Dyadic addition of two numbers i and j denoted by $i\underset{2}{\oplus}j$ is defined by*

$$k=i\underset{2}{\oplus}j=(i_{n-1},i_{n-2},...,i_1,i_0)\underset{2}{\oplus}(j_{n-1},j_{n-2},...,j_1,j_0)=$$
$$=(i_{n-1}\oplus j_{n-1},i_{n-2}\oplus j_{n-2},...,i_1\oplus j_1,i_0\oplus j_0)=(k_{n-1},k_{n-2},...,k_1,k_0)$$

where $k_l=(i_l\oplus j_l)\bmod 2$, for $l=0,1,2,...,n-1$. The dyadic shift, $m=0,1,2,...,N-1$, of a vector $(c_0,c_1,...,c_{N-1})$ is the vector $\left(c_{0\underset{2}{\oplus}m},c_{1\underset{2}{\oplus}m},...,c_{(N-1)\underset{2}{\oplus}m}\right)$.

Definition 3. *Linear code of length $N=2^n$ is called dyadic code if the m-dyadic shift of every codeword is also a codeword for all $m=0,1,2,...,N-1$.*

The class of dyadic codes is a special case of abelian group codes [13, 14-16] which is briefly discussed in the third. In this paper, we would like tointroduce new linear codes with generalized symmetry. We extend cyclic and group codes in the following way. We introduce codes, invariant with respect to a family of generalized shift operators (GSO). In particle case when this family is a group (cyclic or Abelian), these codes are ordinary cyclic and group codes. They are invariant with respect to this group. We deal with GSO-invariant codes with fast code and encode procedures based on fast generalized Fouriertransforms.The hope is that these more general structures will lead to larger classes of useful codes “good” properties. The rest of the paper isorganized as follows: in Section 2 and 3, the proposed method based on families of generalized shift operators (GSO) is explained.

2. Methods

2.1. Generalized shift operator

The purpose of this subsection is to introduce the mathematical representations of generalized shift operators associated with arbitrary orthogonal (or unitary) Fourier transforms (F-transforms). For illustration, we also particularize our results for many transforms popular in coding and signal theories. The ordinary group shift operators $\left(T_t^{\tau}f\right)(t)=f(t+\tau)$ play the leading role in all the properties and tools of the Fourier transform mentioned above. In order to develop for each orthogonal transform a similar wide set of tools and properties as the Fourier transform has, we associate a family of commutative generalized shift operators (GSO) with each orthogonal (unitary) transform. Such families form *hypergroups*. In 1934 F. Marty [17,18] and H.S. Wall [19,20] independently introduced

the notion of hypergroup. Only in particular cases these families are Abelian groups and hyperharmonic analysis is the classical Fourier harmonic analysis on groups.

Let $f(t):\Omega\to\mathbf{F}$ be a $\mathbf{F}$-valued signal, where $\mathbf{F}$ be a finite field. Usually, $\Omega=[0,N-1]^d$ in coding theory and digital signal processing, where d is the dimension of Ω: $d=\dim(\Omega)$. Let

$$L(\Omega,\mathbf{F}):=\left\{f(t)\middle|f(t):\Omega\to\mathbf{F}\right\}\approx\mathbf{F}^{|\Omega|},$$

be vector space of $\mathbf{F}$-valued functions, where $|\Omega|=\mathrm{card}(\Omega)=N^d$. The theory of generalized shift operators was initiated by Levitan [21]–[22]. According to Levitan the family of generalized shift operators (GSOs) $T_t^\tau[f(t)]:=f(t($ $\tau)$ depending on $\tau\in\Omega$ as a parameter is defined in signal space $L(\Omega,\mathbf{F})$ by the following axioms.

Axiom 1. *For all functions* $f_1(t),f_2(t)\in L(\Omega,\mathbf{F})$ *and any constants* $a,b\in\mathbf{F}$ *the following relation holds*

$$\hat{T}_t^\tau[a\cdot f_1(t)+b\cdot f_2(t)]=a\cdot\hat{T}_t^\tau[f_1(t)]+b\cdot\hat{T}_t^\tau[f_2(t)] \tag{1}$$

Axiom 2. *For an arbitrary function* $f(t)\in L(\Omega,\mathbf{F})$ *and arbitrary* $s,t,r\in\Omega$ *it holds*

$$T_\tau^r\left[T_t^\tau[f(t)]\right]=T_t^r\left[T_t^\tau[f(t)]\right],\ \text{or}\ f(t(\ (\tau(\ r))=f((t(\ \tau)(\ r),\ \textit{i.e.,}\ T_t^{\tau(\ r}=T_{t(\ \tau}^r. \tag{2}$$

i. e., the GSOs are associative.

Axiom 3. *There exists an element* $\tau_0\in\Omega$ *with* $T_x^{\tau_0}[f(t)]\equiv f(t)$ *for all* $t\in\Omega$ *and for all* $f(t)\in L(\Omega,\mathbf{F})$. *This means that the family of GSOs contains identity operator.*

If moreover the following axiom is fulfilled, then the GSOs are called *commutative.*

Axiom 4. *For any elements* $\tau,t\in\Omega$ *and arbitrary* $f(t)\in L(\Omega,\mathbf{F})$ *holds*

$$T_\tau^r\left[T_t^\tau[f(t)]\right]=T_r^\tau\left[T_t^r[f(t)]\right],\ \text{or}\ f(t(\ (\tau(\ r))=f(t(\ (r(\ \tau)),\ \textit{i.e.,}\ T_\tau^rT_t^\tau=T_r^\tau T_t^r \tag{3}$$

We expand notion GSOs on the more complex signal space. Let $f(t):\Omega\to\mathsf{A}lg(\mathbf{F})$ be a $\mathsf{A}lg(\mathbf{F})$-valued signal. The set Ω of the values of the variable t constitutes the *domain* of the signal. Usually, $\Omega=[0,N-1]^d$ in coding theory and digital signal processing, where d is the dimension of Ω: $d=\dim(\Omega)$. The set $\mathsf{A}lg(\mathbf{F})$ of values of the signal $f(t)$ is the *range* of the signal. About the range of the signal we assume, that $\mathsf{A}lg(\mathbf{F})$ is a commutative algebra with aninvolution operation $a\to\bar{a},\ \forall a\in\mathsf{A}lg(\mathbf{F})$. In particular, if $\mathsf{A}lg(\mathbf{F})$ is the complex field then the involution operation is complex conjugate.

Let Ω^* be the space dual to Ω. The first one will be called the *spectral domain*, the second one be called *signal domain* keeping the original notion of $t\in\Omega$ as «time» and $\omega\in\Omega^*$ as «frequency». Let

$$L(\Omega,\mathsf{A}lg(\mathbf{F})):=\left\{f(t)\middle|f(t):\Omega\to\mathsf{A}lg(\mathbf{F})\right\}\approx\mathsf{A}lg^{|\Omega|}(\mathbf{F}),$$

$$L(\Omega^*,\mathsf{A}lg(\mathbf{F})):=\left\{F(\omega)\middle|F(\omega):\Omega^*\to\mathsf{A}lg(\mathbf{F})\right\}\approx\mathsf{A}lg^{|\Omega^*|}(\mathbf{F})$$

be two vector spaces of $\mathsf{A}lg(\mathbf{F})$-valued functions. Here $|\Omega|=|\Omega^*|=N^d$. Let $\{\varphi_\omega(x)\}_{\omega\in\Omega^*}$ be an orthonormal system of functions in $L(\Omega,\mathsf{A}lg(\mathbf{F}))$. Then for any function $f(t)\in L(\Omega,\mathsf{A}lg(\mathbf{F}))$ there exists such a function $F(\omega)\in L(\Omega^*,\mathsf{A}lg(\mathbf{F}))$, for which the following equations hold:

$$F(\omega)=(\mathsf{F}f)(\omega)=\sum_{t\in\Omega}f(t)\overline{\varphi}_\omega(t),\quad f(t)=\left(\mathsf{F}^{-1}F\right)(t)=\sum_{\omega\in\Omega^*}F(\omega)\varphi_\omega(t). \tag{4}$$

The function $F(\omega)\in L(\Omega^*,\mathsf{A}lg(\mathbf{F}))$ is called the Fourier spectrum (F-spectrum) of the $\mathsf{A}lg(\mathbf{F})$-valued signal $f(t)\in L(\Omega,\mathsf{A}lg(\mathbf{F}))$ and expressions (1)-(2) are called the pair of *generalized Fourier transforms* (or F-transforms). In the following we will use the notation $f(t)\xleftrightarrow{\mathsf{F}}F(\omega)$ in order to indicate F-transforms pair.

A fundamental and important tool of coding and signal theories are shift operators in the «time» and «frequency» domains. They are defined as

$$\begin{cases}\left(T_t^\tau f\right)(t) := f(t+\tau), \\ \left(\overline{T}_t^\tau f\right)(t) := f(t-\tau)\end{cases} \text{and} \begin{cases}\left(D_\omega^\nu F\right)(\omega) := F(\omega+\nu), \\ \left(\overline{D}_\omega^\nu F\right)(\omega) := F(\omega-\nu).\end{cases}$$

For $f(t) = e^{j\omega t}$ and $F(\omega) = e^{-j\omega t}$ we have

$$\begin{cases}T_t^\tau e^{j\omega t} = e^{j\omega(t+\tau)} = e^{j\omega\tau}e^{j\omega t} = \lambda_\omega(\tau)e^{j\omega t}, \\ \overline{T}_t^\tau e^{j\omega t} = e^{j\omega(t-\tau)} = e^{-j\omega\tau}e^{j\omega t} = \overline{\lambda}_\omega(\tau)e^{j\omega t}\end{cases} \text{and} \begin{cases}D_\omega^\nu e^{j\omega t} = e^{-j(\omega+\nu)t} = e^{-j\nu t}e^{-j\omega t} = \lambda_\nu(t)e^{-j\omega t}, \\ \overline{D}_\omega^\nu e^{j\omega t} = e^{-j(\omega-\nu)t} = e^{j\nu t}e^{-j\omega t} = \overline{\lambda}_\nu(t)e^{-j\omega t},\end{cases} \quad (5)$$

i.e., harmonic signals $e^{j\omega t}$ and $e^{-j\omega t}$ are eigenfunctions of «time»-shift and «frequency»-shift operators $T_t^\tau, \overline{T}_t^\tau$ and $D_\omega^\nu, \overline{D}_\omega^\nu$, corresponding to eigenvalues $\lambda_\omega(\tau) = e^{j\omega\tau}$, $\overline{\lambda}_\omega(\tau) = e^{-j\omega\tau}$ and $\lambda_\nu(t) = e^{-j\nu t}$, $\overline{\lambda}_\nu(t) = e^{j\nu t}$, respectively.

Definition 4. *The following operators (with respect to which all basis functions are invariant eigenfunctions*

$$\begin{aligned}\left(T_t^\tau \varphi_\omega\right)(t) &:= \varphi_\omega(\tau)\cdot\varphi_\omega(t) = \lambda_\omega(\tau)\varphi_\omega(t), \quad \forall\tau\in\Omega, \\ \left(\overline{T}_t^\tau \varphi_\omega\right)(t) &:= \overline{\varphi}_\omega(\tau)\cdot\varphi_\omega(t) = \overline{\lambda}_\omega(\tau)\varphi_\omega(t), \quad \forall\tau\in\Omega\end{aligned} \quad (6)$$

and

$$\begin{aligned}\left(D_\omega^\nu \overline{\varphi}_\omega\right)(t) &:= \overline{\varphi}_\nu(t)\cdot\overline{\varphi}_\omega(t) = \lambda_\nu(t)\cdot\overline{\varphi}_\omega(t), \quad \forall\nu\in\Omega^*, \\ \left(\overline{D}_\omega^\nu \overline{\varphi}_\omega\right)(t) &:= \varphi_\nu(t)\cdot\overline{\varphi}_\omega(t) = \overline{\lambda}_\nu(t)\cdot\overline{\varphi}_\omega(t), \quad \forall\nu\in\Omega^*\end{aligned} \quad (7)$$

are called commutative F *–generalized "time"–shift and "frequency"–shift operators (GSO's), respectively, where* $\lambda_\omega(\tau) = \varphi_\omega(\tau)$, $\overline{\lambda}_\omega(\tau) = \overline{\varphi}_\omega(\tau)$ *and* $\lambda_\nu(t) = \overline{\varphi}_\nu(t)$, $\overline{\lambda}_\nu(t) = \varphi_\nu(t)$ *are eigenvalues of GSO's* $T_t^\tau, \overline{T}_t^\tau$ *and* $D_\omega^\nu, \overline{D}_\omega^\nu$, *respectively.*

For these operators we introduce the following designations:

$$\begin{aligned}\left(T_t^\tau \varphi_\omega\right)(t) &:= \varphi_\omega(t(\ \tau), \quad \left(\overline{T}_t^\tau \varphi_\omega\right)(t) := \varphi_\omega(t'\ \tau), \quad \forall\tau\in\Omega, \\ \left(D_\omega^\nu \overline{\varphi}_\omega\right)(t) &:= \overline{\varphi}_{\omega\oplus\nu}(t), \quad \left(\overline{D}_\omega^\nu \overline{\varphi}_\omega\right)(t) := \overline{\varphi}_{\omega\$\,\nu}(t), \quad \forall\nu\in\Omega^*,\end{aligned}$$

here, symbols "$(\ ,\oplus$","$'\ ,\$$" denote quasi-sums and quasi-differences, respectively. If $T_{t,\sigma}^\tau, \overline{T}_{t,\sigma}^\tau$ and $D_{\omega,\alpha}^\nu, \overline{D}_{\omega,\alpha}^\nu$ are matrix elements of operators $T_t^\tau = \left[T_{t,\sigma}^\tau\right]$, $\overline{T}_t^\tau = \left[\overline{T}_{t,\sigma}^\tau\right]$ and $D_{\omega,\alpha}^\nu = \left[D_{\omega,\alpha}^\nu\right]$, $\overline{D}_{\omega,\alpha}^\nu = \left[\overline{D}_{\omega,\alpha}^\nu\right]$, then

$$\begin{aligned}\left(T_t^\tau \varphi_\omega\right)(t) &= \varphi_\omega(t(\ \tau) = \varphi_\omega(\tau)\cdot\varphi_\omega(t) = \sum_{\sigma\in\Omega} T_{t,\sigma}^\tau \varphi_\omega(\sigma), \\ \left(\overline{T}_t^\tau \varphi_\omega\right)(t) &= \varphi_\omega(t'\ \tau) = \overline{\varphi}_\omega(\tau)\cdot\varphi_\omega(t) = \sum_{\sigma\in\Omega} \overline{T}_{t,\sigma}^\tau \varphi_\sigma(\sigma),\end{aligned} \quad (8)$$

and

$$\begin{aligned}\left(D_\omega^\nu \overline{\varphi}_\omega\right)(t) &= \overline{\varphi}_{\omega\oplus\nu}(t) = \overline{\varphi}_\nu(t)\cdot\overline{\varphi}_\omega(t) = \sum_{\alpha\in\Omega^*} D_{\omega,\alpha}^\nu \overline{\varphi}_\alpha(t), \\ \left(\overline{D}_\omega^\nu \overline{\varphi}_\omega\right)(t) &= \overline{\varphi}_{\omega\$\,\nu}(t) = \varphi_\nu(t)\cdot\overline{\varphi}_\omega(t) = \sum_{\alpha\in\Omega^*} \overline{D}_{\omega,\alpha}^\nu \overline{\varphi}_\alpha(t)\end{aligned} \quad (9)$$

The expressions (8)–(9) are called *multiplication formulae* for basis functions $\left\{\varphi_\omega(t)\right\}_{\omega\in\Omega^*} \in L\left(\Omega, \mathsf{A}lg(\mathbf{F})\right)$ and $\left\{\varphi_\omega(t)\right\}_{t\in\Omega} \in L\left(\Omega^*, \mathsf{A}lg(\mathbf{F})\right)$. They show that the set of basis functions form two hypergroups with respect to multiplication rules (8) and (9), respectively. Consequently, two

spaces $L\left(\Omega,\mathsf{A}lg(\mathbf{F})\right)$ and $L\left(\Omega^*,\mathsf{A}lg(\mathbf{F})\right)$ form time and frequency algebras with structure constants $T^{\tau}_{t,\sigma}$ and $D^{\nu}_{\omega,\alpha}$, respectively.

From (8) and (9) we easily obtain the matrix elements of the GSOs in time and frequency domains

$$T^{\tau}_{t,\sigma}=\sum_{\omega\in\Omega^*}\varphi_\omega(\tau)\varphi_\omega(t)\overline{\varphi}_\omega(\sigma),\quad \overline{T}^{\tau}_{t,\sigma}=\sum_{\omega\in\Omega^*}\overline{\varphi}_\omega(\tau)\varphi_\omega(t)\overline{\varphi}_\omega(\sigma), \tag{10}$$

$$D^{\nu}_{\omega,\alpha}=\sum_{t\in\Omega}\overline{\varphi}_\nu(t)\overline{\varphi}_\omega(t)\varphi_\alpha(t),\quad D^{\nu}_{\omega,\alpha}=\sum_{t\in\Omega}\varphi_\nu(t)\overline{\varphi}_\omega(t)\varphi_\alpha(t). \tag{11}$$

The expressions (10)–(11) can be compactly written on the operator language

$$\begin{aligned} T^{\tau}_{x}&=\mathsf{F}^{-1}\cdot\mathrm{diag}\left\{\varphi_\omega(\tau)\right\}\cdot\mathsf{F},\quad \overline{T}^{\tau}_{t}=\mathsf{F}^{-1}\cdot\mathrm{diag}\left\{\overline{\varphi}_\omega(\tau)\right\}\cdot\mathsf{F},\\ D^{\nu}_{\omega}&=\mathsf{F}\cdot\mathrm{diag}\left\{\varphi_\nu(t)\right\}\cdot\mathsf{F}^{-1},\quad \overline{D}^{\nu}_{\omega}=\mathsf{F}\cdot\mathrm{diag}\left\{\overline{\varphi}_\nu(t)\right\}\cdot\mathsf{F}^{-1}, \end{aligned} \tag{12}$$

where $\mathrm{diag}\left\{\varphi\right\}$ denotes a diagonal matrix which entries consist of values of the function φ.

If there exist such element t_0 that the equation $\varphi_\omega(t_0)\equiv 1$ for all $\omega\in\Omega^*$ is fulfilled, then there exist the identity GSO in time domain. Indeed, the substitution of t_0 into the expressions (12) gives

$$\begin{aligned} T^{t_0}_{t}&=\mathsf{F}^{-1}\cdot\mathrm{diag}\left\{\varphi_\omega(t_0)\right\}\cdot\mathsf{F}\equiv\mathsf{F}^{-1}\cdot\mathrm{diag}\left\{1\right\}\cdot\mathsf{F}\equiv\mathsf{F}^{-1}\cdot\mathsf{F}\equiv I,\\ \overline{T}^{t_0}_{t}&=\mathsf{F}^{-1}\cdot\mathrm{diag}\left\{\overline{\varphi}_\omega(t_0)\right\}\cdot\mathsf{F}\equiv\mathsf{F}^{-1}\cdot\mathrm{diag}\left\{\overline{1}\right\}\cdot\mathsf{F}\equiv\mathsf{F}^{-1}\cdot\mathsf{F}\equiv I. \end{aligned} \tag{13}$$

If there exist such an element ω_0 that the equation $\varphi_{\omega_0}(x)\equiv 1$ for all $x\in\Omega$ is fulfilled too, then there exist the identity GSO in frequency domain. Indeed, the substitution of ω_0 into the expressions (12) gives

$$\begin{aligned} \hat{D}^{\omega_0}_{\omega}&=\mathsf{F}\cdot\mathrm{diag}\left\{\varphi_{\omega_0}(x)\right\}\cdot\mathsf{F}^{-1}=\mathsf{F}\cdot\mathrm{diag}\left\{1\right\}\cdot\mathsf{F}^{-1}=\mathsf{F}\cdot\mathsf{F}^{-1}=I,\\ \hat{\overline{D}}^{\omega_0}_{\omega}&=\mathsf{F}\cdot\mathrm{diag}\left\{\overline{\varphi}_{\omega_0}(x)\right\}\cdot\mathsf{F}^{-1}=\mathsf{F}\cdot\mathrm{diag}\left\{\overline{1}\right\}\cdot\mathsf{F}^{-1}=\mathsf{F}\cdot\mathsf{F}^{-1}=I. \end{aligned}$$

We see also that two families of time and frequency GSOs form two hypergroups $\mathsf{HG}=\left\{T^{\tau}_{t}\right\}_{t\in\Omega}$ and $\mathsf{HG}^*=\left\{D^{\nu}_{\omega}\right\}_{\nu\in\Omega}$. By definition, functions $\left\{\varphi_\omega(t)\right\}_{\omega\in\Omega^*}$ and $\left\{\varphi_\omega(t)\right\}_{t\in\Omega}$ are eigenfunctions of GSOs. For this reason we can call them hypercharacters of hypergroups. For a signal $f(t)\in L\left(\Omega,\mathsf{A}lg(\mathbf{F})\right)$ we define its shifted copies by

$$\begin{aligned} f(t\,(\,\tau)&=\left(T^{\tau}_{t}f\right)(t)=T^{\tau}_{t}\left(\sum_{\omega\in\Omega^*}F(\omega)\varphi_\omega(t)\right)=\sum_{\omega\in\Omega^*}F(\omega)\left(T^{\tau}_{t}\varphi_\omega\right)(t)=\\ &=\sum_{\omega\in\Omega^*}F(\omega)\varphi_\omega(\tau)\varphi_\omega(t)=\sum_{\omega\in\Omega^*}\left(F(\omega)\varphi_\omega(\tau)\right)\varphi_\omega(t),\\ f(t'\;\;\tau)&=\left(\overline{T}^{\tau}_{t}f\right)(t)=\overline{T}^{\tau}_{t}\left(\sum_{\omega\in\Omega^*}F(\omega)\varphi_\omega(t)\right)=\sum_{\omega\in\Omega^*}F(\omega)\left(\overline{T}^{\tau}_{t}\varphi_\omega\right)(t)=\\ &=\sum_{\omega\in\Omega^*}F(\omega)\overline{\varphi}_\omega(\tau)\varphi_\omega(t)=\sum_{\omega\in\Omega^*}\left(F(\omega)\overline{\varphi}_\omega(\tau)\right)\varphi_\omega(t). \end{aligned} \tag{14}$$

Analogously, for a spectrum $F(\omega)\in L\left(\Omega^*,\mathsf{A}lg(\mathbf{F})\right)$

$$\begin{aligned} F(\omega\oplus\nu)&=\left(D^{\nu}_{\omega}F\right)(\omega)=D^{\nu}_{\omega}\left(\sum_{\omega\in\Omega^*}f(t)\overline{\varphi}_\omega(t)\right)=\sum_{\omega\in\Omega^*}f(t)\left(D^{\nu}_{\omega}\overline{\varphi}_\omega\right)(t)=\\ &=\sum_{\omega\in\Omega^*}f(t)\overline{\varphi}_\nu(t)\overline{\varphi}_\omega(t)=\sum_{\omega\in\Omega^*}\left(f(t)\overline{\varphi}_\nu(t)\right)\overline{\varphi}_\omega(t),\\ F(\omega\,\$\,\nu)&=\left(\overline{D}^{\nu}_{\omega}F\right)(\omega)=\overline{D}^{\nu}_{\omega}\left(\sum_{\omega\in\Omega^*}f(t)\overline{\varphi}_\omega(t)\right)=\sum_{\omega\in\Omega^*}f(t)\left(\overline{D}^{\nu}_{\omega}\overline{\varphi}_\omega\right)(t)=\\ &=\sum_{\omega\in\Omega^*}f(t)\varphi_\nu(t)\overline{\varphi}_\omega(t)=\sum_{\omega\in\Omega^*}\left(f(t)\varphi_\nu(t)\right)\overline{\varphi}_\omega(t). \end{aligned} \tag{15}$$

We will need in the following modulation operators:

$$\left(M_t^{\nu} f\right)(t) := \varphi_{\nu}(t) f(t), \qquad \left(\overline{M}_t^{\nu} f\right)(t) := \overline{\varphi}_{\nu}(t) f(t),$$
$$\left(M_{\omega}^{\tau} F\right)(\omega) := \varphi_{\omega}(\tau) F(\omega), \quad \left(\overline{M}_{\omega}^{\tau} F\right)(\omega) := \overline{\varphi}_{\omega}(\tau) F(\omega).$$

From the GSOs definition it follows the following result (two theorems about shifts and modulations). Shifts and modulations are connected as follows:

$$f(t\,(\ \tau) \xleftrightarrow{\mathsf{F}} F(\omega)\varphi_{\omega}(\tau), \quad f(t\,'\ \tau) \xleftrightarrow{\mathsf{F}} F(\omega)\overline{\varphi}_{\omega}(\tau),$$
$$\left(T_t^{\tau} f\right)(t) \xleftrightarrow{\mathsf{F}} \left(M_{\omega}^{\tau} F\right)(\omega), \quad \left(\overline{T}_t^{\tau} f\right)(t) \xleftrightarrow{\mathsf{F}} \left(\overline{M}_{\omega}^{\tau} F\right)(\omega)$$

and

$$F(\omega \oplus \nu) \xleftrightarrow{\mathsf{F}} f(t)\overline{\varphi}_{\nu}(t), \qquad F(\omega\,\$\ \nu) \xleftrightarrow{\mathsf{F}} f(t)\overline{\varphi}_{\nu}(t)$$
$$\left(D_{\omega}^{\nu} F\right)(\omega) \xleftrightarrow{\mathsf{F}} \left(M_t^{\nu} f\right)(t), \quad \left(\overline{D}_{\omega}^{\nu} F\right)(\omega) \xleftrightarrow{\mathsf{F}} \left(\overline{M}_t^{\nu} f\right)(t).$$

2.2. Generalized convolutions and correlations

Using the notion GSO, we can formally generalize the definitions of convolution and correlation.

Definition 5. *The following functions*

$$y(t) := \left(h \lozenge x\right)(t) = \sum_{\tau \in \Omega} h(\tau) x\left(t\,'\ \tau\right), \ Y(\omega) := \left(H \heartsuit F\right)(\omega) = \sum_{\nu \in \Omega^*} H(\nu) F\left(\omega\,\$\ \nu\right)$$

and

$$c(\tau) = \left(f \clubsuit g\right)(\tau) := \sum_{t \in \Omega} f(t)\overline{g}\left(t\,'\ \tau\right), \ C(\nu) = \left(F \spadesuit G\right)(\nu) := \sum_{\omega \in \Omega^*} F(\omega)\overline{G}\left(\omega\,\$\ \nu\right)$$

are called the ◊- and ♥- convolutions and the cross ♣- and ♠- correlation functions, respectively, associated with a classical Fourier transform F. *If* $f = g$ *and* $F = G$ *then cross correlation functions are called the ♣- and ♠- autocorrelation functions.*

The spaces $L\left(\Omega, \mathsf{A}lg(\mathbf{F})\right)$ and $L\left(\Omega^*, \mathsf{A}lg(\mathbf{F})\right)$ equipped multiplications ◊ and ♥ form commutative signal and spectral convolution algebras $\left\langle L\left(\Omega, \mathsf{A}lg(\mathbf{F})\right), \lozenge \right\rangle$ and $\left\langle L\left(\Omega^*, \mathsf{A}lg(\mathbf{F})\right), \heartsuit \right\rangle$, respectively.

Theorem 1. *Let us take two triplets* $y_1(t), h_1(t), x_1(t) \in L\left(\Omega, \mathsf{A}lg(\mathbf{F})\right)$ *and* $y_2(t), h_2(t), x_2(t) \in L\left(\Omega, \mathsf{A}lg(\mathbf{F})\right)$. *Obviously,* $Y_1(\omega), H_1(\omega), X_1(\omega) \in L\left(\Omega^*, \mathsf{A}lg(\mathbf{F})\right)$ *and* $Y_2(\omega), H_2(\omega), X_2(\omega) \in L\left(\Omega^*, \mathsf{A}lg(\mathbf{F})\right)$. *Let*

$$y_1(t) = \left(h_1 \lozenge x_1\right)(t) = \sum_{\tau \in \Omega} h_1(\tau) x_1\left(t\,'\ \tau\right) \ \text{and} \ Y_2(\omega) = \left(H_2 \heartsuit X_2\right)(\omega) = \sum_{\nu \in \Omega^*} H_2(\nu) F_2\left(\omega\,\$\ \nu\right)$$

then generalized Fourier transforms F *and* F^{-1} *map ◊-and ♥-convolutions into the products of spectra and signals, respectively,*

$$\mathsf{F}\left\{y_1\right\} = \mathsf{F}\left\{h_1 \lozenge x_1\right\} = \mathsf{F}\left\{h_1\right\} \cdot \mathsf{F}\left\{x_1\right\}, \quad \mathsf{F}^{-1}\left\{Y_2\right\} := \mathsf{F}^{-1}\left\{H_2 \heartsuit X_2\right\} = \mathsf{F}^{-1}\left\{H_2\right\} \cdot \mathsf{F}^{-1}\left\{X_2\right\},$$

i.e.,

$$y_1(t) = \left(h_1 \lozenge x_1\right)(t) \xleftrightarrow{\mathsf{F}} Y_1(\omega) = H_1(\omega) \cdot X_1(\omega), \quad y_2(t) = h_2(t)x_2(t) \xleftrightarrow{\mathsf{F}} Y_2(\omega) = \left(H_2 \heartsuit X_2\right)(\omega).$$

Theorem 2. Let us take four triplets $c_1(t), f_1(t), g_1(t) \in L\left(\Omega, \mathsf{A}lg(\mathbf{F})\right)$, $c_2(t), f_2(t), g_2(t) \in L\left(\Omega, \mathsf{A}lg(\mathbf{F})\right)$ and $C_1(\omega), F_1(\omega), G_1(\omega) \in L\left(\Omega^*, \mathsf{A}lg(\mathbf{F})\right), C_2(\omega), F_2(\omega), G_2(\omega) \in L\left(\Omega^*, \mathsf{A}lg(\mathbf{F})\right)$. Let

$$c_1(\tau) = \left(f_1 \clubsuit g_1\right)(\tau) = \sum_{t \in \Omega} f_1(t)\overline{g}_1\left(t\,'\ \tau\right), \ \text{and} \ C_2(\omega) = \left(F_2 \spadesuit G_2\right)(\nu) = \sum_{\omega \in \Omega^*} F_2(\omega)\overline{G}_2\left(\omega\,\$\ \nu\right),$$

then generalized Fourier transforms F *and* F^{-1} *map ♣- and ♠-correlations into the products of spectra and signals, respectively,*

$$\mathsf{F}\left\{c_1\right\} = \mathsf{F}\left\{f_1 \lozenge g_1\right\} = \mathsf{F}\left\{f_1\right\} \cdot \mathsf{F}\left\{g_1\right\}, \quad \mathsf{F}^{-1}\left\{C_2\right\} := \mathsf{F}^{-1}\left\{F_2 \heartsuit G_2\right\} = \mathsf{F}^{-1}\left\{F_2\right\} \cdot \mathsf{F}^{-1}\left\{G_2\right\},$$

i.e.,

$$c_1(\tau)=\left(f_1\clubsuit g_1\right)(\tau)\overset{\mathsf{F}}{\longleftrightarrow}C_1(\omega)=F_1(\omega)\cdot\bar{G}_1(\omega),\ ),\quad c_2(t)=f_2(t)g_2(t)\overset{\mathsf{F}}{\longleftrightarrow}C_2(\omega)=\left(F_2\text{♥}G_2\right)(\omega).$$

2.3. Codes invariant with respect to GSOs

We are going to consider block codes of length N as subsets $\mathsf{C}\subset L\left(\Omega,\mathsf{A}lg(\mathbf{F})\right)$ and $\mathsf{C}^*\subset L\left(\Omega^*,\mathsf{A}lg(\mathbf{F})\right)$, i.e., a collections of N length vectors with components from $Alg(\mathbf{F})$. Let $\{\varphi_\omega(t)\}_{t\in\Omega}$ and $\{\varphi_\omega(t)\}_{\omega\in\Omega^*}$ be orthonormal systems of functions for $L\left(\Omega,\mathsf{A}lg(\mathbf{F})\right)$ and $L\left(\Omega^*,\mathsf{A}lg(\mathbf{F})\right)$, respectively. They generate two hypergroups HG- and HG^*.

Definition 6. HG- *and* HG^*- *invariant block codes* $\mathsf{C}\subset L\left(\Omega,\mathsf{A}lg(\mathbf{F})\right)$ *and* $\mathsf{C}^*\subset L\left(\Omega^*,\mathsf{A}lg(\mathbf{F})\right)$ *are linear block codes with the property that if* $c(t)\in\mathsf{C}$ *and* $C(\omega)\in\mathsf{C}^*$ *then*

$$\left(T_t^\tau c\right)(t)=c(t'\ \tau)\in\mathsf{C},\ \forall T_t^\tau\in\mathsf{HG}\ \text{and}\ \left(D_\omega^\nu C\right)(\omega)=C(\omega\$\ \nu)\in\mathsf{C},\quad\forall D_\omega^\nu\in\mathsf{HG}^*,\ \textit{respectively.}$$

It means that HG- and HG^*- invariant block codes $\mathsf{C}\subset L\left(\Omega,\mathsf{A}lg(\mathbf{F})\right)$ and $\mathsf{C}^*\subset L\left(\Omega^*,\mathsf{A}lg(\mathbf{F})\right)$ have hypergroup symmetries $\mathbf{Symm}\{\mathcal{C}\}\approx\mathcal{HG}$ and $\mathbf{Symm}\{\mathcal{C}\}^*\approx\mathcal{HG}^*$.

Reed-Solomon (RS) codes are nonbinary cyclic codes [23]. The most natural definition of HG- and HG^*- invariant RS codesare in terms of a certain evaluation maps from the subspace $\mathsf{A}lg^k(\mathbf{F})$ of all k-tuples $\mathbf{m}=(m_0,m_1,...,m_{k-1})$ (information symbols = massage) over $\mathsf{A}lg(\mathbf{F})$ to the set of codewords $\mathsf{C}\cong Cod\left[N,k\,|\,\mathsf{A}lg(\mathbf{F})\right]\subset L\left(\Omega,\mathsf{A}lg(\mathbf{F})\right)$

$$\mathbf{m}=(m_0,m_1,...,m_{k-1})\mapsto\mathbf{c}(t)=(c(0),c(1),...,c(N-1)),\quad\mathsf{A}lg^k(\mathbf{F})\ \to L\left(\Omega,\mathsf{A}lg(\mathbf{F})\right)\tag{16}$$

or to the set of codewords $\mathsf{C}^*=Cod^*\left[N,k\,|\,\mathsf{A}lg(\mathbf{F})\right]\subset L\left(\Omega^*,\mathsf{A}lg(\mathbf{F})\right)$

$$\mathbf{m}=(m_0,m_1,...,m_{k-1})\mapsto C(t)=(C(0),C(1),...,C(N-1)),\quad\mathsf{A}lg^k(\mathbf{F})\ \to L\left(\Omega^*,\mathsf{A}lg(\mathbf{F})\right)$$

Definition 7. *We define an encoding function for* HG- *and* HG^*- *invariant Reed-Solomon codes as*

$$\mathsf{HG}\text{-RS: }\mathsf{A}lg^k(\mathbf{F})\ \to L\left(\Omega,\mathsf{A}lg(\mathbf{F})\right),\quad\mathsf{HG}^*\text{-RS: }\mathsf{A}lg^k(\mathbf{F})\ \to L\left(\Omega^*,\mathsf{A}lg(\mathbf{F})\right)$$

in the following forms. A message $\mathbf{m}=(m_0,m_1,...,m_{k-1})$ *with* $m_i\in\mathsf{A}lg(\mathbf{F})$ *are transformed by* F *and* F^{-1}:

$$\begin{bmatrix}C(0)\\C(1)\\C(2)\\....\\...\\C(N-2)\\C(N-1)\end{bmatrix}=\mathsf{F}\begin{bmatrix}m_0\\m_1\\...\\m_{k-1}\\\hdashline 0\\...\\00\end{bmatrix},\qquad\begin{bmatrix}c(0)\\c(1)\\c(2)\\....\\...\\c(N-2)\\c(N-1)\end{bmatrix}=\mathsf{F}^{-1}\begin{bmatrix}m_0\\m_1\\...\\m_{k-1}\\\hdashline 0\\...\\00\end{bmatrix},$$

Hence, generator matrices for HG- and HG^*- invariant Reed-Solomon codesare the generalized Fourier matrices F and F^{-1}.

Convolutional cyclic codes (CC's, for short) form an important class of error-correcting codes in engineering practice. The mathematical theory of these codes has been set off by theseminal papers of Forney [24] and Massey et al. [25].

Definition 8. HG- *and* HG^*- *invariant convolutional codes of length* N *and dimension* k *are ideals* $\langle h(t)\rangle$, $\langle G(\omega)\rangle$ *of* $\left\langle L\left(\Omega,\mathsf{A}lg(\mathbf{F})\right),\lozenge\right\rangle$ *and* $\left\langle L\left(\Omega^*,\mathsf{A}lg(\mathbf{F})\right),\text{♥}\right\rangle$ *having the following forms*

$$c(t)=\left(h\lozenge m\right)(t)=\sum_{\tau\in\Omega}h(t'\ \tau)m(\tau)\ \textit{and}\ C(\omega)=\left(G\text{♥}m\right)(\omega)=\sum_{\nu\in\Omega^*}G\left(\omega\$\ \nu\right)m(\nu)$$

where

$$H(\omega)=\left(\mathsf{F}h\right)(\omega)=\left(H(0),H(1),...,H(k-1),0,...,0\right)\in \mathsf{A}lg^{k}(\mathbf{F}),$$
$$g(\omega)=\left(\mathsf{F}^{-1}G\right)(t)=\left(g(0),g(1),...,g(k-1),0,...,0\right)\in \mathsf{A}lg^{k}(\mathbf{F}).$$

We call matrices $\mathbf{G}=\left[G\left(\omega\,\$\,\nu\right)\right]_{\omega,\nu\in\Omega^{*}}$ *and* $\mathbf{H}=\left[h(t'\ \ \tau)\right]_{t,\tau\in\Omega}$ *encoders.*

It is easy to see that cyclic convolutional codes and group convolutional codes are particular cases of HG- and HG^{*}- invariant convolutional codes.

3. Examples

Let $\mathbf{H}_N$ be a finite Abelian group of order $N=N_1N_2\cdots N_n$.The fundamental structure theorem for finite Abelian group implies that we may write $\mathbf{H}_N$ as the direct sum of cyclic groups, $\mathbf{H}_N=\mathbf{Z}_{N_1}\times\mathbf{Z}_{N_2}\times...\times\mathbf{Z}_{N_n}$, where $\mathbf{Z}_{N_l}$ identified with the ring of integers $\mathbf{Z}_{N_l}$ under with respect to modulo N_l and an element $t\in\mathbf{H}_N$ is identified with a point $t=\left(t_1,t_2,...,t_n\right)$ of nD discrete torus. The addition of two elements $t,\tau\in\mathbf{H}_N$ is defined as $\sigma=t\underset{H_N^{*}}{\oplus}\tau=(\sigma_1,\sigma_2,...,\sigma_n)=(t_1\underset{\mathbf{Z}_{N_1}}{\oplus}\tau_1,t_2\underset{\mathbf{Z}_{N_2}}{\oplus}\tau_2,...,t_n\underset{\mathbf{Z}_{N_n}}{\oplus}\tau_n)$.

The Fourier transforms in the space of all functions, defined on the finite Abelian group $\mathbf{H}_N=\bigoplus_{l=1}^{n}\mathbf{Z}_{N_l}$, and with their values in the finite commutative ring (field) or some finite algebra $\mathbf{A}$ has a great interest for digital signal processing. Denote this space as $L\left(\mathbf{H}_N,\mathsf{A}lg(\mathbf{F})\right)$. Let ε_{N_l} a primitive N_l–th root in the algebra $\mathsf{A}lg(\mathbf{F})$. Let us construct the following functions $\chi_{k_l}(t_l)=\varepsilon_{N_l}^{k_lt_l}$, $k_l=0,1,...,N_l-1$. They form the set of characters of the cyclic group $\mathbf{Z}_{N_l}$. Then theset of all characters of the group $\mathbf{H}_N$ can be describe by the following way

$$\chi_k(t)=\chi_{(k_1,k_2,...,k_n)}(t_1,t_2,...,t_n)=\varepsilon_{N_1}^{k_1t_1}\varepsilon_{N_2}^{k_2t_2}\cdots\varepsilon_{N_n}^{k_nt_n},\tag{16}$$

where $k=(k_1,k_2,...,k_n)$.The set of all characters $\left\{\chi_k(t)\right\}_{k\in\mathbf{H}_N^{*}}$ and the set of all indexes $\mathbf{H}_N^{*}$ forms isomorphic multiplicative and additive groups, respectively, with respect to multiplication of characters and addition of indexes $\chi_k(t)\chi_m(t)=\chi_{k\underset{\mathbf{H}_N^{*}}{\oplus}m}(t)=\chi_l(t)$, where

$$l=k\underset{H_N^{*}}{\oplus}m=(l_1,l_2,...,l_n)=(k_1\underset{\mathbf{Z}_{N_1}}{\oplus}m_1,k_2\underset{\mathbf{Z}_{N_2}}{\oplus}m_2,...,k_n\underset{\mathbf{Z}_{N_n}}{\oplus}m_n).$$

The following matrix $\mathsf{F}\equiv\left[\chi_k(t)\right]_{t\in\mathbf{H}_N,k\in\mathbf{H}_N^{*}}$ forms Fourier transform on $\mathbf{H}_N$.

The set $\mathbf{H}_N^{*}$ is called the dual group. It forms ”frequency” domain. If initial group has the structure $\mathbf{H}_N=\mathbf{Z}_{N_1}\oplus\mathbf{Z}_{N_2}\oplus...\oplus\mathbf{Z}_{N_n}$ then the dual group has the same structure $\mathbf{H}_N^{*}=\mathbf{H}_N$.Let us embed finite groups $\mathbf{H}_N$ and $\mathbf{H}_N^{*}$ into two discrete segments $\Omega=[0,N-1]$ and $\Omega^{*}=[0,N-1]$

$$\mathbf{H}_N\to\Omega=[0,N-1],\qquad \mathbf{H}_N^{*}\to\Omega^{*}=[0,N-1],\tag{17}$$

respectively. For this aim we briefly describe a mixed–radix number system now.

A number system is called a weighted number system if any number t can be uniquely expressed in the following form $t=\sum_i t_iw_i$ for some set of integers t_i, called digits, and w_i ’s, called weights. If the weights are successive powers of the same number (for example, 2 or 10), the number system is called a *fixed–radix number system* (for example, *10–radix* or *2–radix*). Any number t in mixed–radix number system can be expressed in the form $t=\sum_{i=1}^{n}t_i\left(\prod_{j=i+1}^{n+1}N_j\right)$. Let $N_1,N_2,...,N_n,N_{n+1}$, where $N_{n+1}\equiv1$ be a finite set of positive integers. Then, with respect to the mixed radixes above, any nonnegative integer $t\in[0,N-1]$, where $N=N_1N_2\cdots N_n$, can be uniquely expressed as

$$t=(t_1,t_2,...,t_n)=t_1\left(N_2N_3\cdots N_{n-1}N_n\right)+...+t_{n-2}\left(N_{n-1}N_n\right)+t_{n-1}\left(N_n\right)+t_n=\sum_{i=0}^{n-1}t_{n-i}\left(\prod_{j=n+1}^{n-i+1}N_j\right),$$

where $t_1\in[0,N_1-1]$, $t_2\in[0,N_1-1]$, ..., $t_{n-1}\in[0,N_{n-1}-1]$, $t_n\in[0,N_n-1]$. The weights of t_i is $\prod_{j=n+1}^{n-i+1}N_j$. The weight of t_n is unity $(N_{n+1}=1)$. The radix-2 representation is $t=t_{n-1}2^{n-1}+t_{n-2}2^{n-1}+...+t_12^1+t_02^0=$ $=\sum_{i=0}^{n-1}t_{n-i}2^i$. Let $t=(t_1,t_2,...,t_n)\in\mathbf{H}_N$ and $(\omega_1,\omega_2,...,\omega_n)\in\mathbf{H}_N^*$ then expressions $\omega=\sum_{i=0}^{n-1}\omega_{n-i}\left(\prod_{j=n+1}^{n-i+1}N_j\right)$, $k=\sum_{i=0}^{n-1}k_{n-i}\left(\prod_{j=n+1}^{n-i+1}N_j\right)$ define the maps (17).

The following operators (with respect to which all characters are invariant eigenfunctions)

$$\begin{aligned}\left(T_t^{\tau}\chi_{\omega}\right)(t)&:=\chi_{\omega}(\tau)\cdot\chi_{\omega}(t)=\chi_{\omega}(t\underset{\mathbf{H}_N}{\oplus}\tau),\quad\forall\tau\in\Omega,\\ \left(\overline{T}_t^{\tau}\chi_{\omega}\right)(t)&:=\overline{\chi}_{\omega}(\tau)\cdot\chi_{\omega}(t)=\chi_{\omega}(t\underset{\mathbf{H}_N}{\$}\tau),\quad\forall\tau\in\Omega\end{aligned}\tag{18}$$

and

$$\begin{aligned}\left(D_{\omega}^{\nu}\overline{\chi}_{\omega}\right)(t)&:=\overline{\chi}_{\nu}(t)\cdot\overline{\chi}_{\omega}(t)=\overline{\chi}_{\omega\underset{\mathbf{H}_N^*}{\oplus}\nu}(t),\quad\forall\nu\in\Omega^*,\\ \left(\overline{D}_{\omega}^{\nu}\overline{\chi}_{\omega}\right)(t)&:=\chi_{\nu}(t)\cdot\overline{\chi}_{\omega}(t)=\overline{\chi}_{\omega\underset{\mathbf{H}_N^*}{\$}\nu}(t),\quad\forall\nu\in\Omega^*\end{aligned}\tag{19}$$

are called commutative $\mathsf{F}_{\mathbf{H}_N}$–generalized "time"–shift and "frequency"–shift operators, induced an abelian group $\mathbf{H}_N$. Itinduces ”exotic” shifts in segments $\mathbf{H}_N\to\Omega=[0,N-1]$, $\mathbf{H}_N^*\to\Omega^*=[0,N-1]$ too, which we will denote as

$$t\underset{\mathbf{H}_N}{\oplus}\tau=(t_1\underset{N_1}{\oplus}\tau_1,t_2\underset{N_2}{\oplus}\tau_2,...,t_n\underset{N_n}{\oplus}\tau_n)\in\Omega=[0,N-1],$$

$$k\underset{\mathbf{H}_N^*}{\oplus}m=(k_1\underset{N_1}{\oplus}m_1,k_2\underset{N_2}{\oplus}m_2,...,k_n\underset{N_n}{\oplus}m_n)\in\Omega^*=[0,N-1].$$

Instead of spaces $L\left(\mathbf{H}_N,\mathsf{A}lg(\mathbf{F})\right)$ and $L\left(\mathbf{H}_N^*,\mathsf{A}lg(\mathbf{F})\right)$ we will speak aboutspaces $L\left(\Omega,\mathsf{A}lg(\mathbf{F})\right)$ and $L\left(\Omega_N^*,\mathsf{A}lg(\mathbf{F})\right)$ and if necessary, in this designations we will distinguish groups, acting in intervals: $L\left(\Omega,\mathsf{A}lg(\mathbf{F})\,|\,\mathbf{H}_N\right)$ and $L\left(\Omega_N^*,\mathsf{A}lg(\mathbf{F})\,|\,\mathbf{H}_N^*\right)$.

Definition 9. $\mathbf{H}_N$- *and* $\mathbf{H}_N^*$- *invariant block codes* $\mathsf{C}\subset L\left(\Omega,\mathsf{A}lg(\mathbf{F})\,|\,\mathbf{H}_N\right)$ *and* $\mathsf{C}^*\subset L\left(\Omega_N^*,\mathsf{A}lg(\mathbf{F})\,|\,\mathbf{H}_N^*\right)$ *are linear block codes with the property that if* $c(t)\in\mathsf{C}$ *and* $C(\omega)\in\mathsf{C}^*$ *then*

$$\left(T_t^{\tau}c\right)(t)=c(t\underset{\mathbf{H}_N}{\oplus}\tau)\in\mathsf{C},\quad\forall\tau\in\Omega\ \textit{and}\ \left(D_{\omega}^{\nu}C\right)(\omega)=C(\omega\underset{\mathbf{H}_N^*}{\oplus}\nu)\in\mathsf{C},\quad\forall\nu\in\Omega^*,$$

respectively

It means that $\mathbf{H}_N$- *and* $\mathbf{H}_N^*$- *invariant block codes* have ordinarygroup symmetries $\mathbf{HypSym}\{\mathsf{C}\}\supset\mathbf{H}_N$ and $\mathbf{HypSym}\{\mathsf{C}^*\}\supset\mathbf{H}_N^*$. The seclass of codes are called the *abelian group codes* [13, 14-16]. Special cases of abelian group codes are 1) a cyclic code, when $\mathbf{H}_N=\mathbf{Z}_{N_1}\times\mathbf{Z}_{N_2}\times...\times\mathbf{Z}_{N_n}\equiv\mathbf{Z}_N$ is a cyclic group, 2) a dyadic code, when $\mathbf{H}_{2^n}=\mathbf{Z}_2\times\mathbf{Z}_2\times...\times\mathbf{Z}_2$. In the first case $\mathsf{F}\cong\left[\varepsilon_N^{t\omega}\right]_{t\in\mathbf{Z}_N,\omega\in\mathbf{Z}_N^*}$ is the ordinary Fourier transform,where ε_N a primitive N–th root in the algebra $\mathsf{A}lg(\mathbf{F})$ and in the second one $\mathsf{F}\cong\left[(-1)^{\langle t|\omega\rangle}\right]_{t\in\mathbf{H}_{2^n},\omega\in\mathbf{H}_{2^n}^*}$ is the Walsh transform, where $\langle t\,|\,\omega\rangle$ is the scalar products oftwo vectors $t=(t_1,t_2,...,t_n)\in\mathbf{H}_N$ and $(\omega_1,\omega_2,...,\omega_n)\in\mathbf{H}_N^*:\langle t\,|\,\omega\rangle=\sum_{i=1}^{n}t_i\omega_i$.

Let $\mathsf{F} \equiv \left[\varepsilon_{2N}^{t}\varepsilon_{N}^{t\omega}\right]_{t,\omega=0}^{N-1} = \left[\varepsilon_{2N}^{t(2\omega+1)}\right]_{t,\omega=0}^{N-1}$. Then

$$F(\omega) = (\mathsf{F}f)(\omega) = \sum_{t\in\Omega} f(t)\varepsilon_{2N}^{-t}\varepsilon_{N}^{-t\omega}, \quad f(t) = \left(\mathsf{F}^{-1}F\right)(t) = \sum_{\omega\in\Omega^{*}} F(\omega)\varepsilon_{2N}^{t}\varepsilon_{N}^{t\omega}$$

is direct and inverse modulation Fourier transform, where ε_{2N} is a primitive $2N$–th root in the algebra $\mathsf{A}lg(\mathbf{F})$. According to definition 4 for $\varphi_{\omega}(t) = \varepsilon_{2N}^{t}\varepsilon_{N}^{t\omega}$ we have

$$\hat{T}_{t}^{\tau}\left\{\varphi_{\omega}(t)\right\} = \varphi_{\omega}(t(\ \tau) = \varphi_{\omega}(\tau)\cdot\varphi_{\omega}(t) = \varepsilon_{2N}^{(2\omega+1)\tau}\varepsilon_{2N}^{(2\omega+1)t} = \varepsilon_{2N}^{(2\omega+1)(t+\tau)},$$

$$\overline{\hat{T}}_{n}^{\tau}\left\{\varphi_{\omega}(t)\right\} = \varphi_{\omega}(t'\ \tau) = \overline{\varphi}_{\omega}(\tau)\cdot\varphi_{\omega}(t) = \varepsilon_{2N}^{-(2\omega+1)\tau}\varepsilon_{2N}^{(2\omega+1)t} = \varepsilon_{2N}^{(2\omega+1)(t-\tau)}.$$

But $\varepsilon_{2N}^{(2\omega+1)(t+N)} = \varepsilon_{2N}^{(2\omega+1)N}\varepsilon_{2N}^{(2\omega+1)t} = \varepsilon_{2}^{(2\omega+1)}\varepsilon_{2N}^{(2\omega+1)t} = (-1)^{(2\omega+1)}\varepsilon_{2N}^{(2\omega+1)t} = -\varepsilon_{2N}^{(2\omega+1)t}$. Hence, $\hat{T}_{t}^{\tau}$ and $\overline{\hat{T}}_{n}^{\tau}$ are negacyclic (skew-cyclic) GSOs:

$$\hat{T}_{t}^{\tau}\left\{c(t)\right\} = (c_{0(\ \tau}, c_{1(\ \tau}, ..., c_{(N-1)(\ \tau}) = (c_{\tau}, c_{\tau+1}, ..., c_{N-2}, c_{N-1}, \underbrace{-c_{0}, -c_{1}, ..., -c_{\tau-1}}_{\tau}),$$

$$\overline{\hat{T}}_{n}^{\tau}\left\{c(t)\right\} = (c_{0'\ \tau}, c_{1'\ \tau}, ..., c_{(N-1)'\ \tau}) = (\underbrace{-c_{N-\tau}, ..., -c_{N-2}, -c_{N-1}}_{\tau}, c_{0}, c_{1}, ..., c_{N-(\tau-1)}).$$

They generate negacyclic (skew-cyclic) codes. Let $\mathsf{F} \equiv \left[\varepsilon_{mN}^{t}\varepsilon_{N}^{t\omega}\right]_{t,\omega=0}^{N-1} = \left[\varepsilon_{mN}^{(m\omega+1)t}\right]_{t,\omega=0}^{N-1}$. Then

$$F(\omega) = (\mathsf{F}f)(\omega) = \sum_{t\in\Omega} f(t)\varepsilon_{mN}^{-t}\varepsilon_{N}^{-t\omega}, \quad f(t) = \left(\mathsf{F}^{-1}F\right)(t) = \sum_{\omega\in\Omega^{*}} F(\omega)\varepsilon_{mN}^{t}\varepsilon_{N}^{t\omega}$$

is direct and inverse ε_{mN}^{-t}-modulation Fourier transform, where ε_{mN} is a primitive mN–th root in the algebra $\mathsf{A}lg(\mathbf{F})$. According to definition 4 for $\varphi_{\omega}(t) = \varepsilon_{mN}^{t}\varepsilon_{N}^{t\omega}$ we have

$$\hat{T}_{t}^{\tau}\left\{\varphi_{\omega}(t)\right\} = \varphi_{\omega}(t(\ \tau) = \varphi_{\omega}(\tau)\cdot\varphi_{\omega}(t) = \varepsilon_{mN}^{(m\omega+1)\tau}\varepsilon_{mN}^{(2m+1)t} = \varepsilon_{mN}^{(m\omega+1)(t+\tau)},$$

$$\overline{\hat{T}}_{n}^{\tau}\left\{\varphi_{\omega}(t)\right\} = \varphi_{\omega}(t'\ \tau) = \overline{\varphi}_{\omega}(\tau)\cdot\varphi_{\omega}(t) = \varepsilon_{mN}^{-(m\omega+1)\tau}\varepsilon_{mN}^{(m\omega+1)t} = \varepsilon_{mN}^{(m\omega+1)(t-\tau)}.$$

But $\varepsilon_{mN}^{(m\omega+1)(t+N)} = \varepsilon_{mN}^{(m\omega+1)N}\varepsilon_{mN}^{(m\omega+1)t} = \varepsilon_{m}^{(m\omega+1)}\varepsilon_{2N}^{(2\omega+1)t} = \varepsilon_{m}\varepsilon_{mN}^{(m\omega+1)t}$. Hence, $\hat{T}_{t}^{\tau}$ and $\overline{\hat{T}}_{n}^{\tau}$ are constacyclic GSOs:

$$\hat{T}_{t}^{\tau}\left\{c(t)\right\} = (c_{0(\ \tau}, c_{1(\ \tau}, ..., c_{(N-1)(\ \tau}) = (c_{\tau}, c_{\tau+1}, ..., c_{N-2}, c_{N-1}, \underbrace{\varepsilon_{m}c_{0}, \varepsilon_{m}c_{1}, ..., \varepsilon_{m}c_{\tau-1}}_{\tau}),$$

$$\overline{\hat{T}}_{n}^{\tau}\left\{c(t)\right\} = (c_{0'\ \tau}, c_{1'\ \tau}, ..., c_{(N-1)'\ \tau}) = (\underbrace{\varepsilon_{m}c_{N-\tau}, ..., \varepsilon_{m}c_{N-2}, \varepsilon_{m}c_{N-1}}_{\tau}, c_{0}, c_{1}, ..., c_{N-(\tau-1)}).$$

They generate constacyclic codes codes.

4. Conclusion

In this paper we studied a new class of codes with generalized symmetry. They are invariant with respect to a family of generalized shift operators $\mathsf{H}\mathsf{G}$or $\mathsf{H}\mathsf{G}^{*}$. In particle case when this family is a group (cyclic or Abelian), these codes are ordinary cyclic and group codes.We deal with GSO-invariant codes with fast code and encode procedures based on fast generalized Fouriertransforms.

5. References

[1] Berlekamp E R 1969 Negacyclic codes for the Lee metric *Combinatorial Mathematics and its Applications* 298-316
[2] Berlekamp E R 1984 *Algebraic coding theory* (New York: Aegean Park Press)
[3] Sergio R L P and Benigno R 2009 Dual generalizations of the concept of cyclicity of codes *Adv. Math. Commun.* **3(3)** 227-234

[4] Boucher D, Geiselmann W and Ulmer F 2007 Skew-cyclic codes *Appl. Algebra Eng.Comm. Comput*. **18**4 379-389
[5] Delphine B, Patrick S and Felix U 2008 Skew constacyclic codes over Galois rings *Adv. Math. Commun.* **23** 273-292
[6] Delphine B and Felix U 2008 Coding with skew polynomial rings *Adv. Math. Commun*. **33** 254-276
[7] Hai Q D 2005 Negacyclic codes of length 2s over Galois rings *IEEE Trans. Inform. Theory IT* **51(12)** 4252-4262
[8] Hai Q D 2008 On the linear ordering of some classes of negacyclic and cyclic codes and their distance distributions *Finite Fields Appl.* **14**1 22-40
[9] Hai Q D, Sergio R L and Szabo S On the structure of cyclic and negacyclic codes over finite chain rings *Codes over rings* **6**
[10] Hai QD and Sergio R L 2004 Cyclic and negacyclic codes over finite chain rings *IEEE Trans. Inform. Theory* **50(8)** 1728-1744
[11] Hakan O and Ferruh O 2009 A note on negacyclic and cyclic codes of length *p^s* over a finite field of characteristic *Adv. Math. Commun*. **3(3)** 265-271
[12] Hai Q 2006 Structure of some classes of repeated-root constacyclic codes over integers modulo 2m *Groups, rings and group rings Lect. Notes Pure Appl. Math.* **248** 105-117
[13] Sundar R B and Siddigi M U 1994 Transform Domain Characterization of Abelian codes *IEEE Trans. Inform. Theory* **40** 2082-2090
[14] Berman S D 1967 Semi-simple cyclic and abelian codes *Kibernetica* **3** 20-30
[15] Camion P 1970 Abelian codes *Tech. Rep.* 1059
[16] MacWilliams F J 1970 Binary codes which are ideals in the group algebra of an abelian group *BSTJ* **49** 987-1011
[17] Marty F 1934 Sur une generalization de la notion de groupe *Sartryckur Forhandlingar Via Altonde Skandinavioka Matematiker kongresseni* 45-49
[18] Marty F 1937 Role de la notion d'hypergroupedansl'etude des groupes non abelians *Computes Rendus de l'Academie des Sciences* **201** 636-638
[19] Wall H S 1934 Hypergroups *Bulletin of the American Mathematical Socity* **41** 36-40
[20] Wall H S 1937 Hypergroups *American Journal of Mathematics* **59** 77-98
[21] Levitan B M 1949 The application of generalized displacement operators to linear differential equations of second order *Uspechi Math. Nauk* **4(1)** 3-112
[22] Levitan B M 1964 Generalized translation operators *Israel Program for Scientific Translations*
[23] Reed I S and Solomon G 1960 Polynomial Codes Over Certain Finite Fields *SIAM Journal of Applied Math* **3** 300-304
[24] Forney G D 1970 Convolutional codes Part I: Algebraic structure *IEEE Trans. Inform. Theory. It* **16** 720-738
[25] Massey J L and Sain M K 1967 Codes, automata, and continuous systems: Explicit Interconnections *IEEE Trans. Aut. Contr*. *AC* **12** 644-650

Acknowledgments
This work was supported by grants the RFBR № 17-07-00886 and by Ural State Forest Engineering's Center of Excellence in "Quantum and Classical Information Technologies for Remote Sensing Systems".

Combined use of correlation measures for selecting semantically close concepts of the ontology

A Yu Timofeeva[1], T V Avdeenko[1], E S Makarova[1] and M Sh Murtazina[1]

[1]Novosibirsk State Technical University, K. Marks ave. 20, Novosibirsk, Russia, 630073

Abstract. The paper suggests a new approach to the selection of correlated concepts for the ontology. It is based on the principal component analysis, but, unlike the standard approach, not Pearson correlation coefficients, but other correlation measures are used. This is due to the fact that the selection of concepts is based on data on the semantic association between concepts and cases, which are represented in the form of weight coefficients that take discrete values and a significant number of zero values. For such cases, the most appropriate is the polychoric correlation coefficient. It allows one to detect a monotonous dependence on the contingency table. However, for a certain table structure, the coefficient erroneously indicates a close relationship. This problem has been analysed in detail, and it has been suggested to use the correlation ratio in problem cases. Using the example of the problem of selecting concepts for the ontology in the IT consulting practice, the advantages of the proposed approach are shown. The first oneis the increasein the percentage of variance of concepts explained by the principal components. The second one is that more concepts are selected based on unsupervised feature selection using weighted principal components.

1. Introduction

One of the key trends in the development of artificial intelligence is associated with the transition from the storage and processing of data to the accumulation and processing of knowledge. In this process, ontology, as a form of representation of knowledge, plays an important role. The main components of ontology are concepts of the subject domain. It is important to select concepts in such a way as to avoid their redundancy. So the semantically close concepts should be selected. This problem can be considered as one of the tasks of machine learning - the feature selectionor the feature extraction.

There are several approaches to solving the problem of feature selecting [1]: filter methods, wrapping techniques, embedded methods.

Filter methods [2] are the simplest. They evaluate each variable according to individual criteria (information gain, chi-square statistics, etc.). An example is the selection algorithm Relief [2, 3]. The disadvantage of filter methodsis that the correlation between the features is not taken into account, therefore, redundant attributes can be selected.

Embedded methods perform the feature selection as part of the model construction process. An example is the LASSO regression [4], constraining the weights of some features and shrinking of others to zero. Thereby a sparse solution is achieved. It includes only relevant features. Estimates of such regression, however, have no analytical expression. It requires the use of numerical optimization algorithms. In addition, the solution is very sensitive to the regularization parameter, which affects the degree of sparseness of the solution.

This drawback is attempted to be eliminated by using "wrappers", search procedures that include learning and evaluating the model using a potential subset of features. However, such procedures require, ideally, a search from all possible subsets of the feature set. So that algorithms are characterized by exponential complexity in terms of the number of features. This, as a rule, is unacceptable, and one must resort to "greedy" search algorithms, which never revise the earlier choice. For example, forward selection and backward elimination are used. However, they can give a locally optimal solution.

Typically, the described approaches are used in supervised learning. This requires a response variable. Based on the quality of its prediction, the attributes are selected. For example, the selection of ontologyconcepts could be done in order to improve the quality of classification of cases. However, the cases do not always have a class label. In this situation, the unsupervised feature selection is performed. This is a more difficult problem [5]. The approaches used can be categorized into cluster recognition and redundancy minimization.

Methods, that involve clustering, [6] select attributes to group data points (in our example, cases) in the best way. Other approaches are not restricted to clustering problems. Their goal is to select the smallest subset of attributes while preserving the most relevant information about the data [7]. The simplest criterion for selecting such subset can be the data variance. The explained variance can be a criterion for both selection and extraction of variables. The most popular approach here is the principal component analysis [8], which uses the decomposition of the covariance (correlation) matrix. Its results are used in the feature selection on the basis of weighted principal components [9]. These approaches are described in section 2.

However, when using correlation coefficients, it is necessary to take into account that the data, as a rule, are not continuous. Typically, the Pearson correlation coefficient is used, which can give biased results in the case of discrete data. For example, it was shown in [10] that when the validity of constructs is analyzed from ordinal values measured in the Likert scale, the results of factor analysis better reflect the theoretical model when factorization is performed using polychoric correlations rather than Pearson correlation coefficients. Nevertheless, the polychoriccorrelation coefficient has a number of drawbacks; in particular, with a certain structure of the contingency table, it erroneously reveals the presence of a strong relationship. This is a particular problem for sparse tables with a large number of zero values.Further, in Section 3, situations with poor behavior of the polychoriccorrelation are analyzed and other appropriate correlation measures are considered. Section 4 compares the various correlation measures and suggests ways to combine them to select concepts. Section 5 presents the results of applying the proposed approach for the selection of ontology concepts in IT consulting practice. Finally, Section 6 gives an interpretation of the results obtained and discusses the directions for their further application.

2. Dimensionality reduction techniques

Traditionally, dimensional reduction techniques [11] were developed for the analysis of either quantitative (principal component analysis) [12] or categorical data (correspondence analysis). Lately a lot of attention has been paid to approaches to the analysis of discrete data. It is suggested in [10] to use the polychoric correlation coefficients to reduce the dimensionality of such data. In addition, exploratory analysis methods for mixed data are being actively investigated. For example, the French school Analyze des données, founded by Jean-Paul Benzécri, develops factor analysis of mixed data [13]. These approaches differ in the way the correlation matrix is calculated. In general, the procedure of the principal component analysis remains standard, it is described below.

2.1. Principal component analysis

Let M be the correlation matrix of$k$ features. On its basisone can obtain weights, which are the association between the variables and the components, so called loadings. Loadings vector for j-th principal component is calculated as

$$a_j = v_j \sqrt{\lambda_j} \tag{1}$$

where v_j is eigenvector of matrix *M* corresponding to the eigenvalue λ_j. The values of the eigenvector are normalized to the sum of the squares of the values. The matrix *A* of loadings for*q* principal components contains the vectors $a_1,\mathsf{K}\ ,a_q$, $q \le k$, *k* is the number of features. The matrix of values of principal components (factor scores) can therefore be given as

$$F = XA\left(A'A\right)^{-1},$$

where *X* is initial data matrix $n \times k$, *n* is the number of cases. Note that the columns of matrix *X* are normalized, i.e. the sample mean of each variable has been shifted to zero, and the sample variance has been shifted to unit. The choice of number *q*is usually based on a scree plot, which shows the proportion of variance explained by each component.

2.2. Unsupervised feature selection

For the feature selection the results of the principal component analysis described in the previous subsection are used. The approach is based on the calculation of the weighted sum of the loadings for the *i*-th feature [9]:

$$\omega_i = \sum_{j=1}^{q} \left|a_{ij}\right| s_j \tag{2}$$

where a_{ij} is *i*-th element of the vector a_j, i.e. the loading for *j*-th principal component by *i*-th feature, s_j is the fraction of the explained variance calculated as

$$s_j = \frac{\lambda_j}{\sum_{l=1}^{k} \lambda_l}.$$

The ordering of the features in order of decreasing weights ω_i allows us to separate the essential concepts from the irrelevant ones. It is proposed in [9] to determine the threshold of weights on the basis of the ideas of a moving average control chart that has been widely used in quality control [14]. The difference is that the weights are not ordered. Therefore, it is proposed to use their random permutations and calculate the indicator

$$\overline{MR_i} = \frac{\left|\omega_{i_1} - \omega_{i_2}\right| + \left|\omega_{i_2} - \omega_{i_3}\right| + \mathsf{K}\ + \left|\omega_{i_k} - \omega_{i_{k-1}}\right|}{k}$$

where $\omega_{i_1}, \omega_{i_2}, \mathsf{K}\ , \omega_{i_k}$ is *i*-th random permutation. The number of permutations *I* should be taken sufficiently large to obtain stable results, for example, 1000. Further, the results are averaged:

$$\overline{MR}^{*} = \frac{1}{I}\sum_{i=1}^{I} \overline{MR_i}.$$

Finally, the threshold is calculated as follows:

$$\gamma = \bar{\omega} + \Phi^{-1}\left(1-\alpha\right)\frac{\sqrt{\pi}}{2}\overline{MR}^{*}$$

where $\bar{\omega} = \frac{1}{k}\sum_{j=1}^{k} \omega_j$, $\Phi^{-1}\left(1-\alpha\right)$ is a quantile of the standard normal distribution of order $\left(1-\alpha\right)$, α is a given significance level, usually, 0.05. Based on the threshold the indicator of relevance of *i*-th feature can be constructed:

$$P\left(i\right) = \begin{cases} 1, & \omega_i \ge \gamma, \\ 0, & \omega_i < \gamma. \end{cases} \tag{3}$$

All the features for which $P = 1$ are recognized as relevant and selected.

In the article [9], which offers the described approach, it is not specified how the number*q* of extracted principal components is chosen. For a different number of components, different weights

will be obtained, which will affect the ordering of the features and the threshold γ. Further, this problem is investigated using the example of concept selecting.

3. Correlation measures

To analyze the association between attributeswhich are difficult to give objective quantification and whose values are ordered categories, the polychoric correlation coefficient is intended. It can be used also when counting data is analyzed, that is, discrete, taking a limited number of numerical values. It can also be rounded data, as well as data are subjectively and inaccurately, for example, expert ratings.

3.1. Polychoric correlation

Polychoric correlation ρ indicates an association between two theorized normally distributed continuous latent variables, from two observed ordinal variables. Its estimation is usually based on maximum likelihood method [15]. Polychoric correlation has the following properties:

- $-1 \le \rho \le 1$.
- It is symmetrical.
- $\rho = 0$ in the case of independence.
- If $\rho = 1$ then there is a strong monotonic relation.

The latter property is an advantage over the Pearson correlation coefficient, which reveals only a linear relationship. At the same time, the advantage of the polychoriccorrelation coefficient may turn out to be a significant drawback. So let us consider a number of examples of tables of relative frequencies:

$$D_1 = \begin{pmatrix} 0.5 & 0.25 \\ 0.25 & 0 \end{pmatrix},\ D_2 = \begin{pmatrix} 0.74 & 0.01 \\ 0.25 & 0 \end{pmatrix},\ D_3 = \begin{pmatrix} 0.74 & 0.25 \\ 0.01 & 0 \end{pmatrix}.$$

In all three cases, the value of the polychoriccorrelation is -1. Thus, the result does not depend on the frequency of non-zero values, as long as $d_{22} = 0$, and the rest were non-zero frequency. But if in the first case it is still possible to presume the presence of some nonlinear dependence, in other cases the small relative frequency of 0.01 can simply be a consequence of the presence of outliers.

The problem also remains for tables of higher dimension that satisfy conditions:

$$d_{1i} \neq 0\ \forall i,\ d_{j1} \neq 0\ \forall j,\ d_{kl} = 0,\ \forall k,\ l \neq 1. \quad (4)$$

If the matrix is close to such a structure, the coefficient will be close to -1 and erroneously indicate an association. A similar problem is characteristic for the Yule coefficient [16], which reveals the relationship between binary variables. It is noted that it is unstable to small frequencies. However, the scientific literature does not offer approaches to solving this problem, which could be directly applied to the problem of feature selection.

Obviously, if the contingency table has a structure described by relations (4), then the use of the polychoric correlation coefficient leads to incorrect results. For this reason, it is necessary to involve other correlation measures that make it possible to identifynonlinear relationships and be appropriate for analysis discrete variables. In this case, they should be more sensitive to non-zero values of the observed frequencies $d_{1i} \neq 0\ \forall i,\ d_{j1} \neq 0\ \forall j$.

The simplest approach would be to replace the polychoric correlation coefficient by the Pearson correlation coefficient in those cases where the first one falsely indicates a strong relationship. Such a trivial approach will also be analyzed, but it is better to choose a measure that is more suitable for analyzing relationships on discrete data.

3.2. Polyserial correlation

One option is the polyserial correlation coefficient ρ_{XY}. It reveals a latent correlation between a continuous variable X and a ordered categorical variable Y. It has the following properties:

- $-1 \le \rho_{XY} \le 1$.
- It is not symmetric, that is, $\rho_{XY} \neq \rho_{YX}$.

- $\rho_{XY} = 0$ in the case of independence.
- If $\rho_{XY} = 1$ then there is a strong association between *X* and *Y*.

Like polychoric correlation, an estimate of polyserial correlation is the result of maximizing the likelihood function [17]. Since, according to the properties, the coefficient ρ_{XY} is not symmetrical, it is therefore important here which of the variables is assumed to be continuous and which is discrete.

3.3. Correlation ratio

Presumably, the same drawbacks as the polychoric correlation may be inherent in the polyserial correlation coefficient. Therefore, in addition, we consider the correlation ratio of a random variable *Y*at random variable*X* defined as

$$\eta^2_{Y|X} = 1 - \frac{\overline{D}_{Y|X}}{D_Y}, \quad (2)$$

where $\overline{D}_{Y|X}$ is the mean value of the conditional variance of the random variable*Y*under the condition *X*, D_Y is unconditional variance of a random variable *Y*. It is obvious from relation (2) that the correlation ratio is always nonnegative. The correlation ratio is asymmetric, that is, $\eta^2_{Y|X} \neq \eta^2_{X|Y}$. A zero value indicates that there is no association. For comparison with the correlation coefficients, it is better to consider the value $\eta_{Y|X}$ or $\eta_{X|Y}$.

To analyze the possibilities of combined use of correlation measures, the polychoric, polyserial correlation coefficients and the correlation ratio have been calculated using the free software for statistical analysis R. For these purposes, a number of user-functions have been implemented [18].

4. Combined use of correlation measures

The selection of the ontology concepts is based on their semantic relations with the cases.The closeness of the semantic relation is determined by some weights that take values from 0 to 1. As a rule, weights are assigned expertly, so take discrete values (for example, rounded).The values of weight coefficients can be calculated on the basis of associative relationships between the case and the ontology concepts [19]. In this case they take a limited number of rational values as a result of multiplication of simple fractions. Thus, the weights are discrete.

The empirical study used data on the semantic association of cases and ontology concepts in the practice of IT consulting [20]. The data contains 120 cases and 20 concepts. First, a matrix of polychoric correlations between all the concepts was constructed.In total, the matrix (lower triangle) contains 190 correlation coefficients. As a result, it was found that 99 coefficients (about half) are close to -1. It should be noted that in the problemcases the numerical optimization of maximum likelihood function does not always give the estimate values exactly equal to -1, since values, close to -1, give approximately the same value of the objective function.

For comparison, the Pearson correlation coefficients r_{xy} are calculated. Figure 1 shows the results in the form of a scatter plot. Here and below (figures 2-4), the line represents the equality of the correlation coefficients, that is, for figure 1, it is a graph of equation $r_{xy} = \rho$.

It can be clearly seen from figure 1 that in most cases the polychoric correlation coefficient indicates a closer association between the concepts than the Pearson correlation coefficient.However, there are also clearly visible problem points - close to -1 values of the polychoric coefficient. In these cases, the values of the Pearson correlation coefficient range from 0 to -0.2, which indicates a rather weak relationship. Nevertheless, the Pearson correlation coefficient does not reveal a non-linear relationship, and therefore may underestimate the closeness of relation.

Both the correlation ratio and the polyserial coefficient are asymmetric. So, next, the correlation ratio η was calculated as the average value between $\eta_{Y|X}$ and $\eta_{X|Y}$. In the same way the polyserial correlation ρ_s was calculated as the average between ρ_{XY} and ρ_{YX}.

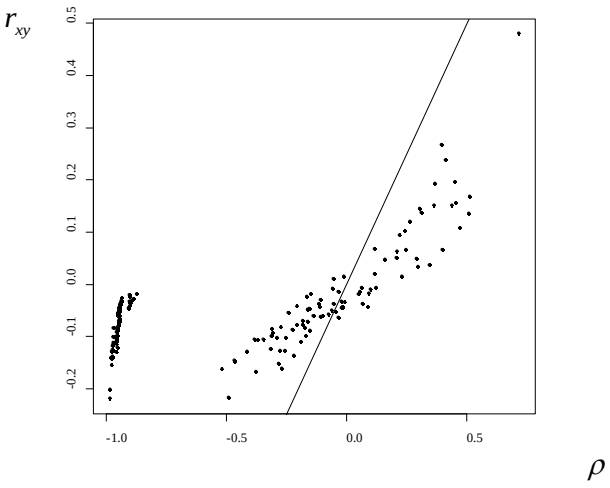


Figure 1. Scatterplot of polychoric and Pearson correlation coefficients.

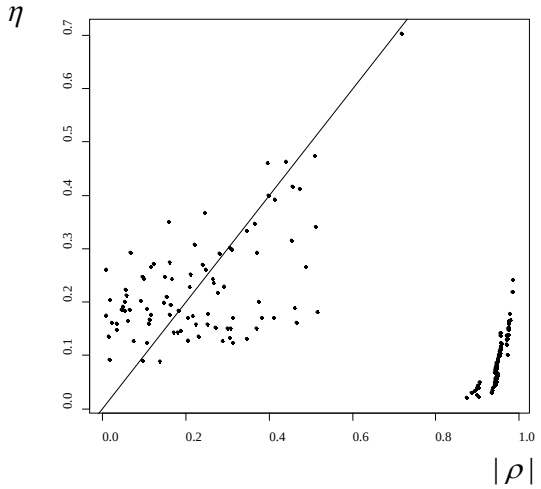


Figure 2. Scatterplot of polychoric coefficient and correlation ratio.

As noted above, the correlation ratio does not indicate the direction of the relationship, since it takes only non-negative values. For this reason, it is more correct to compare it with the absolute values of correlation coefficients. So figure 2 compares its values with the absolute values of the polychoric correlation coefficient. It can be seen that in a number of cases the correlation ratio shows a closer relationship, and in others the polychoric coefficient. There are also problem situations, in these cases the correlation ratio takes values close to the absolute values of the Pearson correlation coefficient, and indicates a weak relationship.But the values of the correlation ratio η, according to its properties, are always greater or equal to $|r_{xy}|$. So it is better to use the correlation ratio. In order to take into account the direction of the relationship, one must take the sign of the polychoric correlation coefficient.

If we compare the polychoric and polyserial correlation coefficients (figure 3), in most cases (77 coefficients out of 91, not related to the problem ones), the polychoric coefficients indicate a closer relationship than the polyserial ones. Thus, polyserial coefficients systematically underestimate the closeness of the relation. What can not be said about the correlation ratio: out of 91 non-problematic coefficients, only 44 polychoric correlation coefficients have the absolute value greater than the correlation ratio.

At the same time, in the problem cases, the polyserial coefficient shows a closer relation between the concepts, since it takes values from -0.6 to -0.2. However, this indicates, rather, that this coefficient also negatively reacts to a certain structure of the contingency tables. This is clearly seen in figure 4, which compares the absolute values of the polyserial coefficient and the correlation ratio.Situations, in which the values of the polychoric coefficient are close to -1, are highlighted in gray. They obviously stand out from the rest of the points on the graph.

Thus, we propose an approach to ontology concepts selection consisting of the following steps.

Step 1. Calculation of polychoric correlations ρ.

Step 2. Identification of problem situations by frequency tables satisfying (4), as well as by the values of the polychoric correlations close to –1.

Step 3. Replacement of the polychoric correlations in the problem situations, revealed at the step 2, by the values of the correlation ratios η, calculated as the mean between $\eta_{Y|X}$ and $\eta_{X|Y}$, multiplied by $\text{sign}(\rho)$.

Step 4. Based on the resulting correlation matrix *M*, consisting of polychoric correlations and correlation ratios, calculation of loadings vector for *j*-th principal component by the formula (1). Interpretation of results allows us to identify blocks of interrelated concepts.

Step 5. For the concept selection, the calculation of the weighted sum of the loadings by the formula (2). The weights $\omega_1,...,\omega_k$ allow us to order concepts by relevance. The calculation of indicator *P* by the formula (3), and the selection of most informative concepts.

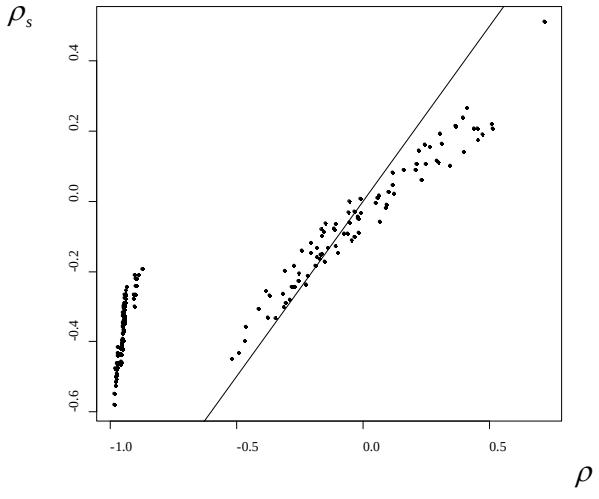


Figure 3. Scatterplot of polychoric and polyserial correlation coefficients.

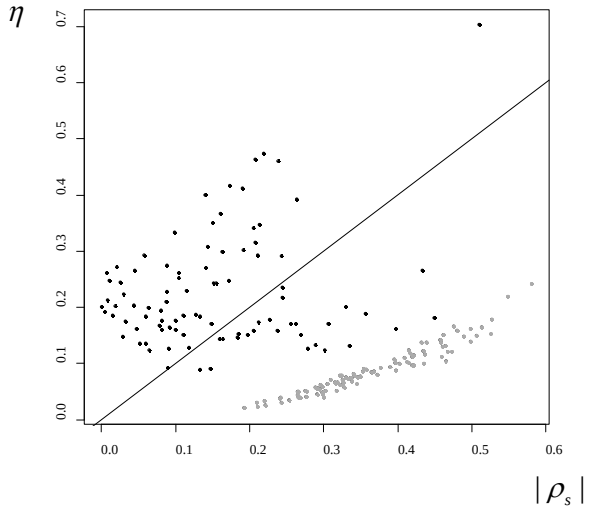


Figure 4. Scatterplot of polyserialcoefficient and correlation ratio.

The advantages of the proposed approach as compared to the standard one (calculation of the Pearson correlation) should consist in increasing the percentage of variance of concepts explained by the extracted components. As a result, this allows us to partitionthe concepts into a smaller number of groups, the interrelation within which are closer.

5. Application in the practice of IT- consulting

Despite the fact that the concepts are carefully organized into the ontology by a domain specialist, the IT problem of the user being solved is often at the junction of various concepts. Therefore, the cases often refer to different hierarchical branches. The application of methods of grouping concepts could identify the most informative groups of concepts, as well as the most frequent combination of concepts describing the user's problems. The latter can be used in the decision support system for intellectual help for the user what additional concepts (in addition to the one already selected) to choose for the link with the current case (user problem).

In the above exampleof ontology in the practice of IT consulting we selected the most relevant concept using the proposed approach. For comparison the standard method based on Pearson correlation coefficients was used. The number *q*of extracted principal components was varied from 1 to 15. Table 1 shows the values of indicator function *P* defined given by (3). The results are presented for a limited number of principal components in order to demonstrate the differences between approaches. It can be seen that number and composition of selected concepts vary depending on *q*. Generally, the number of selected concepts is very small.It is difficult to detect any pattern of changes in a subset of selected concepts with an increase in the number of extracted components. Using the proposed approach, it is possible to increase the number of selected concepts. The maximum number of them is achieved when fiveprincipal components are extracted.Based on this, the optimal number *q* is chosen equal to five.

The loadings matrices for five principal components are presented in table 2. It allows to present concepts of the ontology of IT consulting in a space of small dimension. For clarity, only significant values of the loadings are given in the table. Their absolute values indicate the closeness of the relationship between concepts and the principal components. The weak relation of concepts (for example, "Vacation" in the standard approach) with all five components indicates that such concepts could not be included in the identified groups. The results of the feature selection (table 1) can be

compared with the results of the principal component analysis (table 2). For example, if one uses only the first component, only the attributes that closely correlate with this component remain in the set.

Table 1. The results of unsupervised feature selection.

	The standard approach					The proposed approach				
Concept	*q*=1	*q*=3	*q*=5	*q*=7	*q*=9	*q*=1	*q*=3	*q*=5	*q*=7	*q*=9
Vacation	0	0	0	0	0	1	1	1	0	0
Sick leave	0	1	1	0	0	1	1	1	0	1
Time-keeping	1	1	1	1	1	1	1	1	1	1
Calculation	0	0	0	0	0	0	0	1	0	0
Calculationof deductions	0	0	1	1	1	0	0	1	1	1
2-NDFL	1	1	1	0	0	1	1	1	1	1
6-NDFL	1	1	1	0	0	1	1	1	1	1

Table 2. Loadings on principal components and cumulative percentage of explained variance.

	The standard approach					The proposed approach				
Concept	1	2	3	4	5	1	2	3	4	5
Order on admission		0.445					0.664			
Order of dismissal		0.409			0.375		0.573			
Vacation						-0.629				
Sick leave		-0.469				-0.514				
Time-keeping	0.514	-0.440				-0.549				
Reporting		0.673					0.806			
Calculation prepayment				-0.457						0.573
Calculation					-0.399				0.748	
Payment at the average wage					0.396				-0.543	
Calculation of deductions			0.446	-0.422						0.476
Salary				-0.569						0.576
Recalculation			0.645					-0.607		
2-NDFL	-0.744					0.838				
6-NDFL	-0.681		-0.459			0.727				
Insurance payment			0.425	0.404	-0.337			-0.677		
Other taxes								-0.491		
Wirings								-0.461	0.415	
Cumulative explained variance, %	9.7	18.1	25.9	32.7	**38.9**	14.7	27.0	37.7	46.6	**55.1**

As can be seen, from the results of table 2, due to the proposed approach, the percentage of the explained variance significantly increases.So, with the standard approach, the five extracted components sum up only 38.9% of the initial variation of the concepts, whereas the proposed approach allows to explain 55.1% of the variance. In addition, five identified groups were able to include more concepts, additionally included "Vacation", "Other taxes" and "Wirings". Thus, the desired effect is achieved.

6. Results and discussion

The obtained results can be interpreted from the point of view of IT consulting practice.

The concepts, combined the first principal component, reflect the most common user errors in the calculations. If there is an incorrect calculation, then as a rule the error arises either in the incorrect

formulation of vacation or sick leave, and the problem with the time-keeping. At the same time, problems with vacation and sick leave can lead to the errors in reporting on taxes (2-NDFL and / or 6-NDFL). Reports on personal income tax are also interrelated, if there is an error or a question on one report, then the second one most likely will also have an error.

The second group of concepts deals with problems in personnel reporting. If there is a question on the admission / dismissal orders, there will be a problem with personnel reporting, and vice versa, if there is an error in the report, then it is worth checking the personnel orders (admission, dismissal).
The concept "Recalculation" is connected with the third principal component. When recalculating, as a rule, users forget to remake taxes, so there are errors in taxes, insurance payment and wirings as a consequence.

Wirings also fell into the fourth group. The problem with wirings also arises when the calculation is incorrectly. These are interdisciplinary issues. The calculation and the payment at the average wage are mutually exclusive types, that is, at the same time a sick leave (payment at the average wage) and calculation (salary payment) can not meet together, this is a mistake. So, the user needs to make changes.

The fifth principal component associates with calculation prepayment, calculation of deductions and salary. In the payment documents, it is always necessary to check the calculation of deductions, so that everything is reflected correctly in the 6-NDFL statements. Also through salary payment documents a prepayment is formed. The prepayment is usually a fixed amount, sometimes as half of the salary, then in the payment document deductions are reflected. But such questions are rare.

Thus, concepts are combined into the groups by how often the errors occur when working with the software products. The first group of concepts is the most frequently encountered problematic situation, since the calculation errors are usually more frequent. The second most popular are the problems with personnel documents (the errors of the second group). The problems with taxes and the average wage are not very frequent operations and this part is fairly well implemented in the programs. Therefore, there are fewer questions on this part. The prepayment, deductions and salary are, as a rule, the most recent operations in the general list of all operations, and if everything was done correctly in the previous steps, then there are very few errors in this part.

As a result, concepts of different hierarchical branches were combined, i.e. errors often arise at the junction of the concepts that fall into groups. As a recommendation to improve the decision support system, this can be used, for example, if the consultant chooses one concept for linking the case, the system may recommend him choosing other concepts from the group identified on the basis of the principal components.

Thus, during the ontology concept selection based on their semantic relationships with cases, the use of the principal component analysis requires the choice of appropriate correlation measures. Due to the characteristics of the data representing the weight coefficients and taking discrete values, it is suggested to use the polychoric correlation. However, as it turned out, it gives incorrect results for a certain structure of the contingency tables. At the same time, in the conducted empirical study of the ontology of the IT consulting domain, this structure occurs quite often (in half the cases). Therefore, it is suggested in the problem situations to replace the polychoric correlation coefficient by the correlation ratio, which reveals nonlinear relationships and is appropriate for discrete data. As a result, such a combined use of correlation measures makes it possible to increase the percentage of the explained variance of principal components. So it allowsto increasethe number of selected concepts based on unsupervised feature selection using weighted principal components.

7. References

[1] Flach P 2012 *Machine learning: the art and science of algorithms that make sense of data* (Cambridge University Press)
[2] Kira K and Rendell L A 1992 *Proceedings of the ninth international workshop on Machine learning* 249-256
[3] Robnik-Sikonja M and Kononenko I 2003 *Machine learning* **53** 23-69
[4] Tibshirani R1996 *Journal of the Royal Statistical Society. Series B (Methodological)* **58** 267-288

[5] Dy J G and Brodley C E 2004 *Journal of machine learning research* **5** 845-889
[6] He X, Cai D, and Niyogi P 2006 *Advances in neural information processing systems* **18** 507-514
[7] Golay J and Kanevski M 2017 *Knowledge-Based Systems* **135** 125-134
[8] Parveen A N, Nisthana H, Inbarani H H and Kumar E N S 2012 *Proc. Int. Conf. on Computing, Communication and Applications (ICCCA)* 1-7
[9] Kim S B and Rattakorn P 2011 *Expert systems with applications* **38** 5704-5710
[10] Holgado–Tello F P, Chacón-Moscoso S, Barbero–García I and Vila-Abad E 2010 *Quality & Quantity* **44** 153-166
[11] Myasnikov E V 2017 *Computer Optics* **41(4)** 564-572 DOI: 10.18287/2412-6179-2017-41-4-564-572
[12] Spitsyn V G, Bolotova Yu A, Phan N H and Bui T T T 2016 *Computer Optics* **40(2)** 249-257 DOI: 10.18287/2412-6179-2016-40-2-249-257
[13] Pagès J 2014 *Multiple Factor Analysis by Example Using R* (London, Chapman & Hall/CRC The R Series)
[14] Vermaat M B, Ion R A, Does R J M M and Klaassen C A J 2003 *Quality and Reliability Engineering International* **19** 337-353
[15] Olsson U 1979 *Psychometrica* **44** 443-460
[16] Kendall M and Stuart A 1961 *The Advanced Theory of Statistics: Inference and relationship* (London: Charles Griffin and Co., Ltd.) p 676
[17] Drasgow F 1988 *Encyclopedia of statistical sciences* (John Wiley & Sons)
[18] Timofeeva A Y 2017 *CEUR Workshop Proceedings* **1837** 188-194
[19] Avdeenko T V and Makarova E S 2017 *CEUR Workshop Proceedings* **2005** 11-20
[20] Avdeenko T V and Makarova E S 2017 *Journal of Physics: Conference Series* **803** 012008

Acknowledgments
The reported study was funded by Russian Ministry of Education and Science, according to the research project No. 2.2327.2017/4.6.

Algorithms of automation of 3D modeling of urban environment objects using attributive information from a digital map

M P Osipov[1] and O A Chekodaev[1]

[1]Lobachevsky State University of Nizhni Novgorod, Prospekt Gagarina 23, Nizhni Novgorod, Russia, 603950

Abstract. A method for generating three-dimensional models of urban environment objects using attributive information from a digital map is described. Algorithms for automating the process of modeling objects of the urban environment are presented. A flexible and expanding structure is proposed that allows generating both types of objects already available in the classifier, as well as new ones added by the user.

1. Introduction

3D models of urban environment are widely used in various fields, such as civil engineering and architecture, culture and education, advertising and entertainment. Unlike a two-dimensional schematic representation in maps, three-dimensional models of objects allows a person to perceive information in a familiar spatial form. No special technical knowledge is required for the analysis of such information, so decision-making based on the information thus provided is much quicker and easier. Real-time visualization of a 3D model allows for a total immersion into the model space and thus gives the user the most complete information about the object. With this approach, it is possible to gain some experience in the operation of objects even at an early stage when their concept is being developed. In particular, a 3D model of urban environment facilitates the interaction of designers and investors. By incorporating the model of the future object (building) into the model of the surrounding development, it is possible to perform a visual and landscape analysis, to assess the influence of the future object on the appearance of the historical urban environment, to analyze the visibility, taking into account the height of the buildings. In urban management, a 3D model helps to carry out engineering calculations and to plan social infrastructure, to study possible consequences of emergencies and to simulate the actions of emergency and other special services [1-4].

Urban environment models can be created manually, semi-automatically or automatically. Automatic modeling is performed by software based on the results of airborne laser scanning and aerial photography [5, 6]. This approach allows for rapid model generation, but the geometry of such a model can be distorted due to overlap of objects and imperfection of existing equipment and algorithms. With the manual and semiautomatic modeling of the urban environment, distortion of the geometry can be avoided, but this process is quite time-consuming, because the modeling is either completely or partially performed by an operator [7, 8]. The source of the data for creating the model is a digital terrain map. The aim of automation of modeling methods is to minimize the proportion of manual labor.

2. Problem statement

Automation of methods for modeling three-dimensional objects of the urban environment is achieved through the development of "procedural modeling" algorithms using both ready-made 3D object models and those that were completely generated with the use of attributive information from digital maps. The term "procedural modeling" refers to the automatic creation of three-dimensional models of objects using certain rules of spatial modeling based on various information contained in a digital map. The main advantages of this approach include the high speed of creating an urban environment model, the model's metric accuracy, and sufficient flexibility in editing. One drawback of the approach is the insufficient photorealism of the model obtained. Methods using "procedural modeling" distribute all objects according to the types of construction of their 3D model [9, 10]. The rules for generating 3D models of each type of construction are formed, as a rule, separately and independently of each other, which does not allow for the flexibility and extensibility of the generation process of the urban environment model. For example, when creating a new type of object model construction, one has to develop an algorithm for generating a 3D model of this type "from scratch".

The aim of this work is to develop algorithms for automating the process of modeling urban environment objects using attributive information from a digital map. The objective is to identify the general principles for constructing the geometry and model texturing (construction types), to develop algorithms for constructing and creating a hierarchical structure of construction types for a 3D model of the object, which facilitates the editing of the construction rules and automates the process of creating new types of a 3D object model construction.

3. Methods for generating 3D models of the urban environment

The urban environment model includes a digital model of the terrain and models of the urban environment objects.

The digital model of the terrain is built on the basis of elevation information (horizontals, elevation marks, etc.) contained in the digital map (figure 1). Fast algorithms for the automatic generation of such a model are developed by means of Delaunay triangulation [11] and its subsequent restructuring in accordance with the shape of the terrain described by a set of structural lines from the digital map.

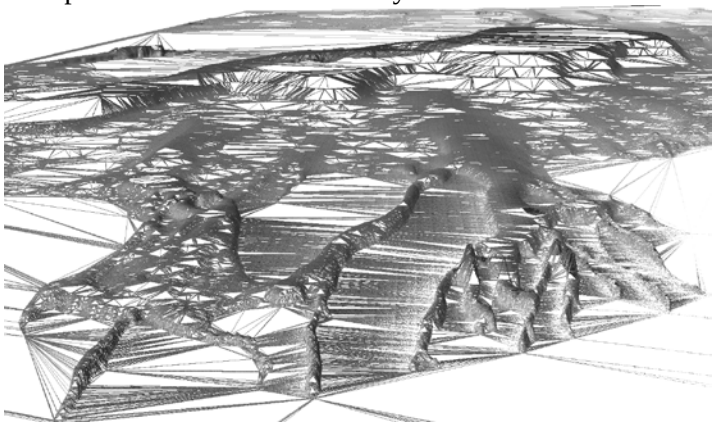


Figure 1. The digital model of the terrain.

As a basis for constructing Delaunay triangulation, a dynamic caching algorithm is adopted. It is characterized by high speed (average complexity is O (N)) and ease of implementation [11]. It makes it possible to complicate and rearrange the existing triangulation in accordance with the new altitude information (points). In this way, it locally changes the parts of the terrain model according to the updated high-altitude data without rearranging the entire model. Approaches to increase the speed of the dynamic caching algorithm was proposed and implemented. These include the choice of data storage structures, the use of integer arithmetic, the minimization of conditional transitions, the unfolding of a part of the cycles, the simplification of the functions, etc. The optimal step to perform the verification of compliance with the Delaunay condition was determined experimentally. This step provides a gain in the running time of the algorithm on large amounts of data. A special structure for

working with a list of verifiable triangles was proposed and implemented. It allows you to minimize the use of procedures for allocating and freeing dynamic memory, by effectively managing the allocated memory. The conducted experiments showed the effectiveness of the proposed approaches.

Objects of the urban environment are buildings and structures, roads, lawns, sidewalks, bodies of water, infrastructure elements, etc. Most models of such objects can be generated manually beforehand using a 3D modeling editor. The information about the ready 3D model and its position can be stored in a digital map. In the absence of a ready model, it is generated automatically based on the attributive description of the object from the digital map of the local area. For storing information about the digital map, a distributed database management system is used, which is based on an object-oriented approach to the representation and processing of complex structured spatially distributed information [12]. The human brain can easily compare a real object with its unified 3D model, which allows the urban environment objects to be typified [13]. Therefore, it is proposed to divide all objects of the urban environment into types of their representation and to develop algorithms for generating a 3D model of objects of each type. Objects can be divided by the type of presentation into two large groups: those containing metric information and discrete objects.

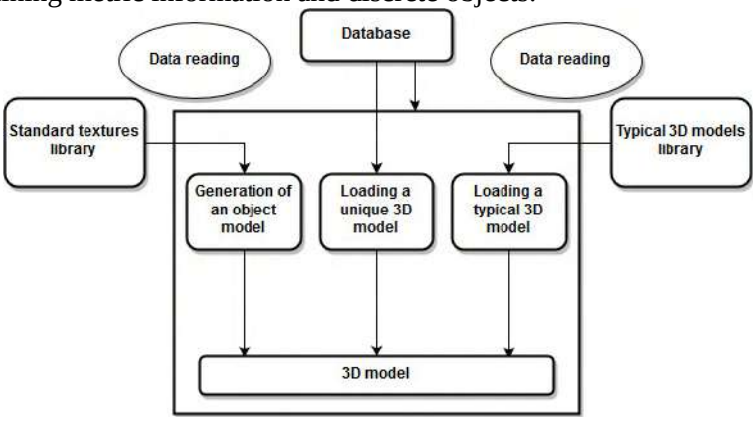


Figure 2. The flowchart of model generation.

Discrete objects are typical objects of the urban environment that have no metric description: trees, stops, traffic lights, etc. Each of such types of objects can have a different description represented by a list of characteristics. This can be the type of material, age, dimensions, etc. Since the list of possible characteristics of this type of object in a digital map is known and only a set of characteristics is changed for a particular object, it is proposed to form beforehand a database of ready-made 3D models of typical objects and textures representing different combinations of their characteristics. Attribute information of discrete objects in the digital map also contains information about the object's location on the ground. In this case, discrete objects are not generated, but are represented by ready 3D models that are loaded into the program from the library of typical 3D models. If discrete objects have characteristics that enable automatic generation of their geometry, they are allocated to separate generation classes.

Objects that have metric characteristics will be generated automatically. Such objects include buildings, fences, road network, water objects, etc. Each such object should have a description in the form of a representation of its contour and a list of characteristics. For example, for a building, this is the material, the purpose of the building, construction year, the number of floors, etc. It is proposed to form beforehand a library of standard textures characterizing possible sets of characteristics of a particular object. The geometric model of such objects is generated automatically by the contour representation and the list of characteristics stored in the digital map. Also, based on these characteristics, the desired texture for this model will be selected from the library of standard textures. The general scheme of this principle is shown in figure 2.

All the automatically generated objects of the urban environment are divided into classes, depending on the methods used to build their geometry. Such classes are a hierarchical structure, where each next step of the hierarchy is a modification of a more general geometric construction of the

model at the previous step. Thus, to draw an object added to the classifier, it is sufficient to indicate to which class this object belongs. Figure 3 shows a block diagram of the distribution of objects from the database by the initial classes.

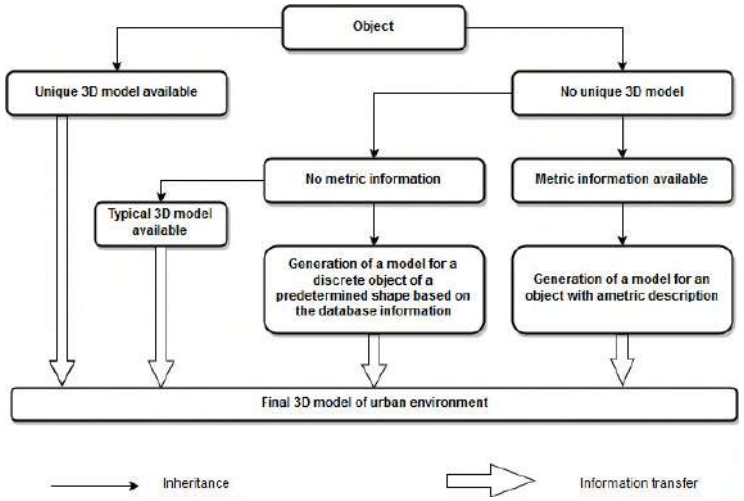


Figure 3. Distribution of objects by initial classes.

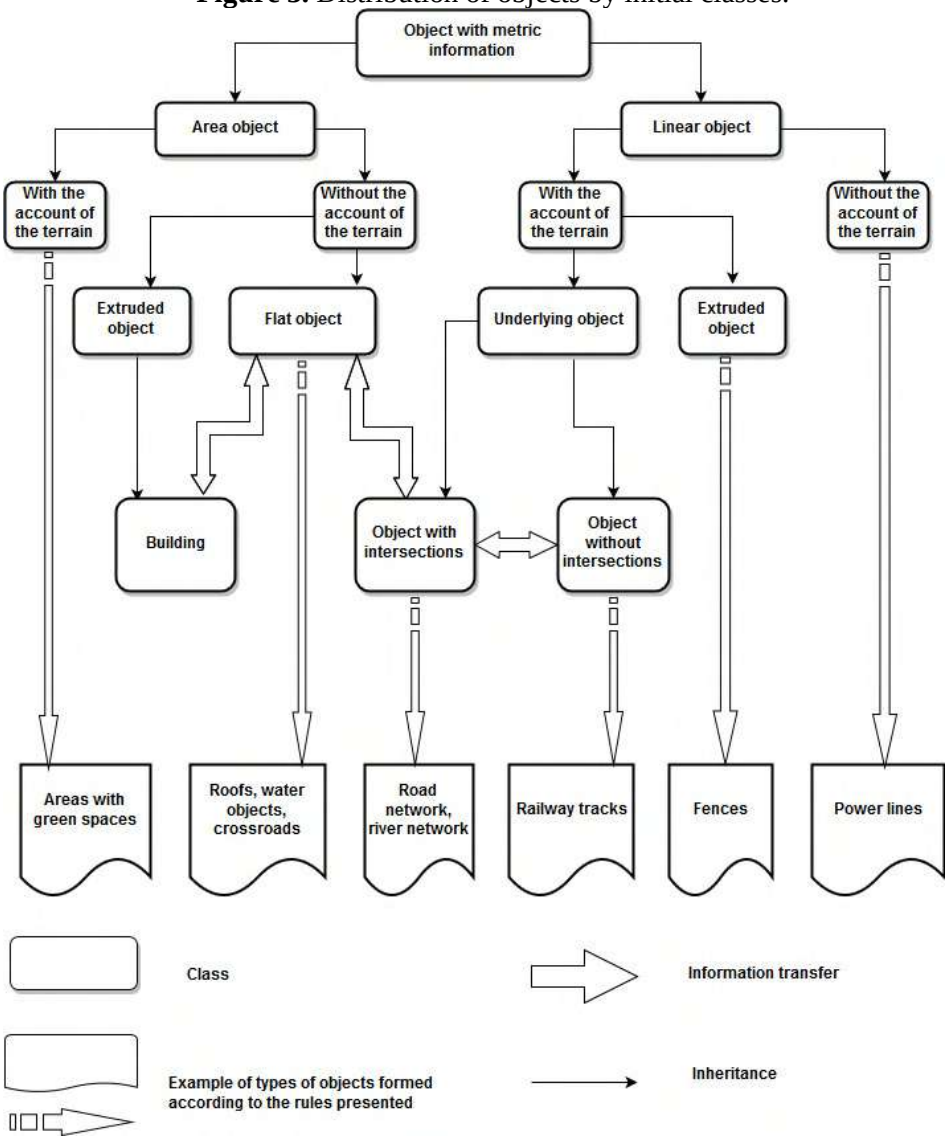


Figure 4. Schematic of objects subdivision into types of geometric model construction.

In the case of an object of complex shape, the object should be divided into composite objects whose shape can be generated by the existing construction classes. These composite objects must have characteristic links, which allow storing information about the relations between composite objects. For example, the object "fence with columns" is divided into several objects: a linear one ("fence") and some discrete "columns". For the "fence", interrupts are stored at the metric points, where "columns" adjoin the fence.

Let us consider in more detail the blocks defined in figure 3 as "Model generation by a metric description". The breakdown by classes is based on various characteristics of objects, which make it possible to identify common algorithms for constructing the geometry and texturing the model.

The general scheme of objects subdivision into types of geometric model construction is shown in figure 4.

The subdivision into the types of the model construction is carried out according to the following principle. Objects containing metric information are divided into area and linear objects. In turn, area and linear objects may or may not be linked to terrain features. For example, a linear object that is not linked to terrain includes power lines hanging in the air, elevated pipes, etc.

An area object that is not linked to terrain features can be flat or extruded. A flat object is constructed by triangulation within a specified contour. Some examples of such objects are a lake or a flat roof of a house. The geometry of an extruded object is represented by a set of vertical rectangles obtained by extruding the segments of a metric description to a given height. Examples of such objects are buildings.

The geometry of an area object linked to terrain features is a surface consisting of a set of triangles (parts of triangles) of the terrain that falls inside a given contour of this object. Objects of such type include elements of the underlying surface, such as a covered area, a green area, a square, etc.

Linear objects linked to terrain features are divided into two types: extruded or underlying objects (being part of the earth's surface). The geometry of extruded linear objects is similar to the geometry of an extruded area object. Various types of fences can be named as examples of such objects.
Due to the different configuration of the location of the underlying linear objects relative to one another, it may be required to form intersections when such objects are modeled.

The geometry of underlying linear objects that do not require intersections to be formed is constructed in the form of a set of bound quadrilaterals obtained by stretching the segments of a metric description to a given width. In this case, the height of each point of the obtained quadrilateral should correspond to the height of the terrain at this point. An example of this type of objects are road sections without intersections.

Underlining linear objects that require to form intersections are divided into several parts of the following type: underlying linear objects whose geometry was described above, and the intersection area proper. To form a geometric model that describes the intersection of a set of linear objects is a difficult task and will be described separately. This type of objects includes a road network, a river network, and the like.

The texturing of the constructed models is performed based on the information from a set of characteristics describing the object being modeled. The desired texture is automatically extracted from the library of standard textures that characterize possible sets of object characteristics. The principle of texturing will be considered below using a building model as an example.

4. Forming the areas of linear objects' intersections

Linear objects such as the road network and the river network are extended and have crossings and ramifications. If a usual algorithm for forming a linear object model is applied, overlapping of objects will occur in such places, which will result in a poor quality of rendering of these areas. To avoid this, it was decided to divide the algorithm for constructing a model of linear objects with intersections into three stages: forming the sections far from the intersection, forming the sections near the intersection, and forming the intersection areas themselves.

The sections of the road far from the intersections will be assigned to the class of underlying linear objects without intersections, taking into account the terrain, while crossroads will be assigned to flat area objects. The sections of the road near the crossroads will be moved to the next class in the

hierarchy, in the branch of linear objects taking into account the terrain and the presence of intersections. This division is due to the fact that the crossroads model in its origin is a set of linear objects around the crossroads point, and to construct them, information on each of them is required for the whole group. To build a section of the road near the intersection, information is required about the adjacent linear sections of the road near the intersection. From the database point of view, these will be another object. Therefore, in order to optimize the process of acquiring all the necessary information, these road sections will be sorted clockwise.

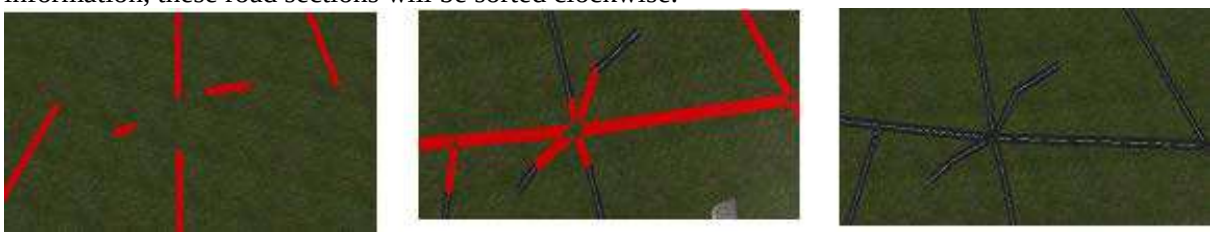


Figure 5. Stages of forming an intersection of roads (from left to right).

After all necessary geometries of roads have been formed, all three classes are combined into one and will represent a model of the road network. Figure 5 shows all 3 stages of the process of forming a crossroads.

5. Producing a building model

A building is a combination of two objects: the sidewalls and the roof. The building's flat roof is an area object without height. The walls of the building remain in the class of extruded area objects. One floor of the building is formed and the information about the roof contour is transferred to the class of flat area objects without taking the terrain into account. Next, in the class "Building walls", the building walls are formed in full from all floors (each floor being obtained by using a previous stage algorithm). This class will be the final class for the walls, since this model has a specific algorithm for calculating the geometry and texture coordinates. The reason for this is that, to make the building more realistic, we decided to create a set of standard textures according to the following parameters of the building: number of storeys, building type (residential or office buildings), the material from which the buildings are made.

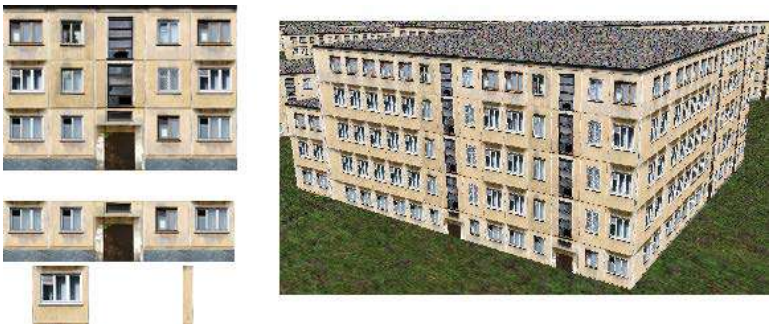


Figure 6. Standard texture for the building wall; division of the texture into sections: the entrance, the window, the wall and the resulting 3D model.

The texture is a combination of the lower, middle and upper storeys of the building. This texture will be superimposed on the walls according to the following principle: the walls of the building are formed storey-by-storey and for each storey a respective part of the given texture will be assigned, namely, the lower part of the texture for the first storey, the upper part for the upper storey, and the medium part for all other storeys. To avoid distortion of the texture on narrow and wide walls, we decided to divide the sidewall not only into storeys, but also into entrances, window apertures where the wall is too small for the entrance and simply into pieces of the wall that are too small even for a window. The geometry is calculated for each section of the wall described above.

Figure 6 gives an example of a standard texture, and shows how this texture is split into the sections of the same storey described above.

6. Conclusions

The proposed structure allows for the flexibility and extensibility of the urban environment model generation process. When adding a new type of objects to the classifier, there is no need to create an algorithm for forming a model of this type. It is sufficient to only specify in its description a set of characteristics by which the program will determine which graphic class of objects it belongs to and, based on the particular characteristics of the object indicated in the digital map, it will build the desired graphic model. Figure 7 shows a fragment of the electronic map and a 3D model obtained on its basis.

Figure 7. A fragment of the electronic map and a 3D model built on the basis of the attributive information of the digital map.

7. References

[1] Vasin Yu G, Osipov M P and Tomchinskaya T N 2011 Development of Interactive Virtual Models of the Urban Landscape of the Historical Center of Nizhni Novgorod *Pattern Recognition and Image Analysis* **21(2)** 351-353 DOI: 10.1134/s1054661811021100

[2] Kim D H and Bejleri I 2005 Using 3D GIS simulation for urban design *ESRI Users Conf.* 1-14

[3] Bourdakis V 2008 Low Tech Approach to 3D Urban Modeling *Architecture in Computro* 959-964

[4] Guoqing Z, Zhenyu T, Penggen C, and Weirong C 2017 *Modelling and visualizing 3D urban environment via internet for urban planning and monitoring* (Access mode: http://www.isprs.org/proceedings/xxxv/congress/comm2/papers/151.pdf)

[5] Bondarets A 2010 Main strategies for creating 3D city models *GisLab Geographic information systems and remote sensing* (Access mode: http://gis-lab.info/qa/3dcities.html)

[6] ArcReview: Three-dimensional modeling and photorealistic visualization of urban areas *Dataplus Geoinformation systems for business and society* (Access mode: https://www.dataplus.ru/news/arcreview/ detail.php)

[7] *C3 - a precursor of the 3D revolution in computer mapping* (Access mode: http://www.ixbt.com/td/c3-technologies-3d-maps.shtml)

[8] Sholomitsky A A, Lunev A A and Tarasova O M 2011 A technology for three-dimensional modeling of cities *Scientific Bulletin of National Mining University* **2** 57-62

[9] Carrozzino M, Tecchia F and Bergamasco M 2017 *Urban procedural modeling for real-time rendering* (Access mode: http://www.isprs.org/proceedings/XXXVIII/5-W1/pdf/carrozzino_ etal_ 2.pdf)

[10] Rautenbach V, Bevis Y, Coetzee S and Combrinck C 2015 Evaluating procedural modeling for 3D models of informal settlements in urban design activities *South African Journal of Science (SAJS)* **111(11/12)** 11-10 DOI: 10.17159/sajs.2015/20150100

[11] Skvortsov A V 2002 *Delaunay Triangulation and its Application* (Tomsk: Tomsk University Publ.) p 128

[12] Vasin Yu G and Yasakov Yu V 2016 Distributed database management system for integrated processing of spatial data in a GIS *Computer Optics* **40(6)** 919-928 DOI: 10.18287/24126179-2016-40-6-919-928

[13] Vasin Yu G, Osipov M P, Muntyan S V and Kustov E A 2015 Procedural Modeling and Interactive 3D Visualization of Objects of the Internal Structure of Buildings and Facilities *Pattern Recognition and Image Analysis* **25(2)** 278-280 DOI: 10.1134/s105466181502025

Acknowledgements

The study was supported by a grant from the Russian Science Foundation (project No. 16-11-00068).

On improving the quality of VoIP connections

A A Bukatov[1], D Y Polukarov[2], N D Zaitsev[3] and A M Sukhov[2]

[1]Southern Federal University, Stachki str. 200/1, k.213, Rostov-on-Don, Russia, 344090
[2]Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086
[3]Don State Technical University, Lenina str. 69, Rostov-on-Don, Russia, 344079

Abstract. Improving the quality of VoIP connections is a very important goal in the area of telecommunications. The proportion of multimedia traffic in relation to the total traffic supported by providers is constantly increasing. To identify and troubleshoot issues with VoIP connections, network providers need both criteria and a methodology for assessing connection quality. We offer a methodology for assessing the quality of VoIP connections. A comparative analysis of VoIP codecs is also made.

1. Introduction

The proportion of multimedia traffic as compared to the total traffic supported by providers is constantly increasing [1]. Hence, much attention is paid to the quality of VoIP services. There is a need to define criteria for assessing the quality of VoIP connections in order to improve the quality of telecommunications services.

The existing methods for assessing speech quality across VoIP connections are analyzed in [2]. The authors of [2] chose what they considered to be the best audio codec in this context.

The following types of voice signal distortion are considered in [2]:

- voice stream delay;
- intermittent and illegible speech;
- extraneous noise;
- echoes;
- unnatural voice (robotic voice effects);
- abnormal (too low) signal volume.

Further, in [3], four characteristics of network connections are considered: i.e., bandwidth, delay, loss, and the variation of the delay (jitter). All these characteristics are measured in terms of three grades: Good, Acceptable, Poor. Thus, the compound metric is comprised of four values which each take one of the grades G-A-P (Good, Acceptable, or Poor).

This approach is more general and formal than others which have been proposed, and therefore more universal and suitable for automation.

However, the approach implemented in [3] is specifically for audiovisual traffic (VVoIP connections). We see the advisability of using this approach for VoIP codecs, as well, discussed in [2].

2. Related work

Objective methods for assessing the quality of a voice signal are widely studied in [4, 5, 6, 7, 8]. In these works, various methods are offered: Emodel, PSQM / PSQM +, PESQ, P.563. These will be described in more detail below.

In [2] methods are proposed for constructing tools for monitoring the quality of voice flow transmissions. The development of these tools were based on the use of VoIP telephony systems to support the activities of educational institutions using distance learning forms to enhance the distance learning process.

In [3] a new mechanism for providing online assessments of VVoIP quality of service is introduced. This operates on network paths without the participation of users. The mechanism uses the "GAP-model", which is a model for measuring the QoE in terms of measurable network factors such as bandwidth, delay, jitter and loss (see above).

3. Overview of assessment methods

Speech quality assessment methods for VoIP systems are subdivided into subjective and objective metrics [9]. Subjective methods require that an expert evaluates the situation in question, and therefore are unacceptable in relation to automatic evaluation. Objective methods for assessing the quality of voice transmission are divided into two groups. The first group assesses the quality of transmission of primary data streams. This group includes Emodel [4]. The second group evaluates the quality of the audio stream transmission itself. The second group includes the PSQM/PSQM+ method (Perceptual Speech Quality Measure) [5], which has been further developed into the PESQ method (Perceptual Evaluation of Speech Quality) [6]; this second group also includes the method P.563 [7, 8].

The results of a comparative analysis of voice quality estimation methods for VoIP systems are given in Table 1 [2].

Table 1. Comparison of objective speech quality analysis methods.

	Emodel	PESQ	P.563
The absence of excess traffic	yes	no	yes
The possibility of one-way data flow analysis	yes	no	yes
The possibility of analyze the types of distortion of the received speech stream	no	yes	yes

4. Comparing codecs and choosing the best one

There are network routes that are comparatively long but connect, end-to-end, quite closely spaced nodes.

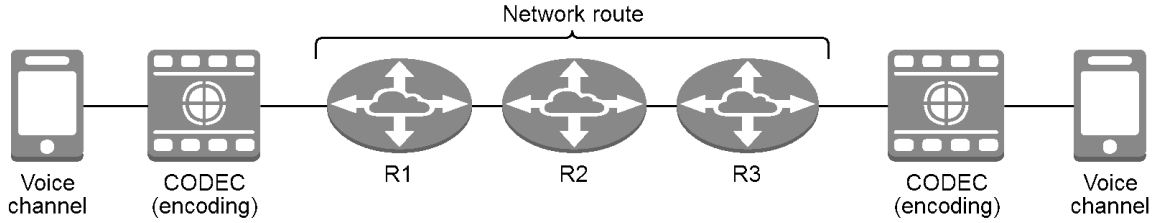


Figure 1. Scheme of transmission voice signal between VoIP-phones.

For example, consider these cases.

Case 1: a subscriber VoIP phone is connected to the corporate telecommunications network of the Southern Federal University (SFU) through the LTE (mobile 4G network) of the mobile operator MTS. The VoIP subscriber is located in the city of Rostov-on-Don.

The route shown in Figure 2 can be described by the following sequence of cities: Rostov-on-Don - Moscow - St. Petersburg - Helsinki - Stockholm - Amsterdam - St. Petersburg - Moscow – Rostov-on-Don. It should be noted that the shorter return route from Amsterdam to St. Petersburg (which reduces total length of the route) is due to the highly developed infrastructure of the international channels of the branch network of the Ministry of Education and Science of the Russian Federation RUNNet [10].

It is known that the transmission of information flows through "long" routes subjects such streams to certain kinds of distortion [11], such as those caused by delays in packet delivery (tempo distortions) and those caused by packet loss — in the case of data transfer protocols based on RTP protocols / RTTP data transfer for VoIP telephony, and over the transport protocol UDP. Regarding

delay, we note that the most "destructive" distortion imposed on the transmitted real-time signal (which is the VoIP telephony signal) is not so much the delay per-se, as the variation in this value (or jitter) [12]. In Figure 2, the maximum delay value reaches 200 ms (in the 18th line), and the jitter value is 76 ms. Note that the packet loss level is not displayed by the route trace command.

Table 2. Tracing options.

Parameter	Value
Maximum delay, ms	200
Jitter, ms	76

```
C:\Program Files\Far Manager>tracert  www.sfedu.ru

Трассировка маршрута к www-sfedu.r61.net [195.208.245.171]
с максимальным числом прыжков 30:

  1     1 ms     1 ms     1 ms  www.huaweimobilewifi.com [192.168.8.1]
  2    48 ms    38 ms    34 ms  198.18.8.1
  3    55 ms    48 ms    46 ms  10.250.245.201
  4    59 ms    54 ms    40 ms  10.249.12.115
  5    52 ms    57 ms    48 ms  10.249.126.253
  6    48 ms    44 ms    40 ms  ler-cr01-ae23.100.rnd.stream-internet.net [195.34.38.174]
  7    80 ms    60 ms    74 ms  mag9-cr01-be4.61.msk.stream-internet.net [212.188.29.5]
  8    80 ms    76 ms    70 ms  oct-cr03-be1.78.spb.stream-internet.net [212.188.2.37]
  9    91 ms    79 ms    99 ms  kivi-cr01-ae8.78.hel.stream-internet.net [212.188.54.2]
 10    99 ms    80 ms    87 ms  bro-cr01-be7.135.stk.stream-internet.net [195.34.50.146]
 11   110 ms    99 ms    88 ms  se-fre.nordu.net [194.68.128.24]
 12   115 ms    94 ms    90 ms  fi-csc.nordu.net [109.105.102.57]
 13   115 ms    88 ms    94 ms  ndn-gw2.runnet.ru [109.105.102.58]
 14   123 ms   101 ms   107 ms  b57-1-gw.spb.runnet.ru [194.85.40.186]
 15   115 ms   107 ms    99 ms  m9-3-gw.msk.runnet.ru [194.85.40.229]
 16   117 ms   100 ms   117 ms  m9-2-gw.msk.runnet.ru [194.85.40.214]
 17   133 ms   128 ms   127 ms  rsu.rostov-don.runnet.ru [194.190.254.58]
 18   125 ms   124 ms   200 ms  c1-uginfo-vl4.r61.net [195.208.248.141]
 19   148 ms   158 ms   158 ms  bg1-uginfo-gi0-1-130.r61.net [195.208.248.150]
 20   118 ms   121 ms   125 ms  c1-uginfo-vl156.r61.net [195.208.248.154]
 21   127 ms   124 ms   131 ms  www-sfedu.r61.net [195.208.245.171]

Трассировка завершена.
```

Figure 2. Example of an access route from an "external" VoIP device to a corporate network.

Case 2: a subscriber VoIP phone is connected to the corporate telecommunications network of the Southern Federal University (SFU) through the home Wi-Fi network. The VoIP subscriber is located in the city of Rostov-on-Don. The sequence of cities, given in Figure 3, is the same as that of the previous case: Rostov-on-Don - Moscow - St. Petersburg - Helsinki - Stockholm - Amsterdam - St. Petersburg - Moscow – Rostov-on-Don.

Table 3. Tracing options (second case).

Trace number	1	2
Maximum delay, ms	208	113
Jitter, ms	65	2

The parameters of these distortions of the data stream transmitted through the network can be used directly to estimate the quality of voice transmission, but such estimates are very approximate. However, these general distortions of the data stream are the primary sources of the secondary distortions which are directly perceived by VoIP subscribers — specific to speech transmission. Such distortions include, for example, delays in the arrival of a speech signal, intermittency and illegibility in speech, the presence of extraneous noise, the presence of an echo signal, and others which are considered in this paper. The evaluation of speech quality by measuring these secondary types of distortion is more accurate.

As we shall see, the jitter value found via the above analysis significantly exceeds the threshold level: i.e., that which is acceptable for high-quality speech transmission. Therefore, the problem of analyzing the quality of voice streams sent within the framework of connections with remote

subscribers is of great importance for the purpose of further improving the parameters affecting the quality of voice transmission.

```
C:\Users\Диимон>tracert www.sfedu.ru

Трассировка маршрута к www-sfedu.r61.net [195.208.245.171]
с максимальным числом прыжков 30:

  1     2 ms     2 ms     1 ms  myhome.ru [192.168.1.1]
  2     3 ms     3 ms     3 ms  10.255.61.251
  3     5 ms     3 ms     6 ms  10.255.61.18
  4     5 ms     4 ms     3 ms  10.255.61.14
  5     4 ms     4 ms     3 ms  ler-cr01-ae23.200.rnd.stream-internet.net [195.34.36.112]
  6    25 ms    26 ms    25 ms  mag9-cr01-be4.61.msk.stream-internet.net [212.188.29.5]
  7    38 ms    37 ms    37 ms  oct-cr03-be1.78.spb.stream-internet.net [212.188.2.37]
  8    43 ms    45 ms    43 ms  kivi-cr01-ae8.78.hel.stream-internet.net [212.188.54.2]
  9    50 ms    49 ms    49 ms  bro-cr01-be7.135.stk.stream-internet.net [195.34.50.146]
 10    49 ms    51 ms    50 ms  se-tug.nordu.net [194.68.123.24]
 11    55 ms    56 ms    56 ms  fi-csc.nordu.net [109.105.102.57]
 12    81 ms    79 ms    79 ms  ndn-gw2.runnet.ru [109.105.102.58]
 13   143 ms   206 ms   208 ms  185.141.124.150.runnet.ru [185.141.124.150]
 14    85 ms    84 ms    84 ms  m9-2-gw.msk.runnet.ru [194.85.40.53]
 15   109 ms   109 ms   109 ms  rsu.rostov-don.runnet.ru [194.190.254.58]
 16   107 ms   107 ms   107 ms  c1-uginfo-vl4.r61.net [195.208.248.141]
 17   104 ms   103 ms   103 ms  bg1-uginfo-gi0-1-130.r61.net [195.208.248.150]
 18   123 ms   107 ms   107 ms  c1-uginfo-vl156.r61.net [195.208.248.154]
 19   111 ms   103 ms   104 ms  www-sfedu.r61.net [195.208.245.171]
```

```
C:\Users\Диимон>tracert www.sfedu.ru

Трассировка маршрута к www-sfedu.r61.net [195.208.245.171]
с максимальным числом прыжков 30:

  1     6 ms     4 ms     3 ms  myhome.ru [192.168.1.1]
  2     9 ms     4 ms     4 ms  10.255.61.251
  3     4 ms     4 ms     5 ms  10.255.61.18
  4     4 ms     4 ms     4 ms  10.255.61.14
  5     5 ms     6 ms     7 ms  ler-cr01-ae23.200.rnd.stream-internet.net [195.34.36.112]
  6    27 ms    29 ms    27 ms  mag9-cr01-be4.61.msk.stream-internet.net [212.188.29.5]
  7    40 ms    40 ms    42 ms  oct-cr03-be1.78.spb.stream-internet.net [212.188.2.37]
  8    47 ms    45 ms    63 ms  kivi-cr01-ae8.78.hel.stream-internet.net [212.188.54.2]
  9    51 ms    52 ms    51 ms  bro-cr01-be7.135.stk.stream-internet.net [195.34.50.146]
 10    50 ms    50 ms    55 ms  se-tug.nordu.net [194.68.123.24]
 11    76 ms    75 ms    58 ms  fi-csc.nordu.net [109.105.102.57]
 12    80 ms    84 ms    82 ms  ndn-gw2.runnet.ru [109.105.102.58]
 13    92 ms    90 ms    91 ms  185.141.124.150
 14    91 ms    86 ms    85 ms  m9-2-gw.msk.runnet.ru [194.85.40.53]
 15   111 ms   111 ms   113 ms  rsu.rostov-don.runnet.ru [194.190.254.58]
 16   110 ms   111 ms   110 ms  c1-uginfo-vl4.r61.net [195.208.248.141]
 17   105 ms   105 ms   108 ms  bg1-uginfo-gi0-1-130.r61.net [195.208.248.150]
 18   108 ms   110 ms   109 ms  c1-uginfo-vl156.r61.net [195.208.248.154]
 19   106 ms   105 ms   105 ms  www-sfedu.r61.net [195.208.245.171]

Трассировка завершена.
```

Figure 3. Second example of an access route from an "external" VoIP device to a corporate network.

There are known methods for assessing the quality of voice transmission, based both on the analysis of primary distortions of the transmitted data stream, and on the analysis of secondary distortions characteristic of voice transmission specifically. This article discusses the causes and types of such distortions; compares the methods for assessing the presence of such distortions in order to choose the method most suitable for use in monitoring the quality of speech; and suggests methods for improving voice quality across corporate VoIP telephony systems by improving both the system VoIP-telephony, and its environment. The article ends with general conclusions concerning the results obtained.

One of the most promising methods for improving the quality of VoIP connections is associated with the use of more efficient codecs [13].

The use of the more modern Speex codec [14] allows for a significantly weakening of the requirements, i.e., the threshold values for the primary data transmission quality indicators in terms of what is needed in order to provide acceptable voice quality. Specifically, when using this codec, a satisfactory quality of voice transmission is ensured even with delays of up to 150 ms, a jitter value of up to 15 ms, and losses of up to 10% of packets. Thus, in comparison with the recommendations of ITU-T G.712 [15], the threshold value for the permissible variation in delays is increased by a factor of 1.5, and the percentage of admissible data loss is increased tenfold. [2, 9].

Note, however, that the Speex codec is not the best currently available. The electronic resource devoted to this codec [14] provides the information that the Speex codec is surpassed, according to all indicators, by the new freely distributed Opus codec [16]. Note also that the Opus codec, developed in 2011 (the latest version of this codec was released in July 2016), has already been standardized by the IETF (Internet Engineering Task Force) as standard RFC 6716 [17]. This standard combines the technologies of such well-known codecs as Skype SILK [18] and Xipn.Org CELT [19]. Because of the noted advantages of the codec, Opus decided to implement it in the VoIP-PBX IP4Tel system, which is the basis of the system, proposed here, for monitoring the quality of voice transmission across corporate VoIP-telephony systems.

5. Conclusions and future work

For this study, a comparative analysis and a development of the methods for assessing the quality of voice transmission across VoIP telephony networks was carried out. Also, a comparative analysis of the following VoIP codecs was performed: Speex, Opus, Skype SILK and Xipn.Org CELT. This was a comparison by the following criteria: allowable delay, allowable losses and allowable jitter (delay variation). This comparison showed that the Opus codec is the best choice at present. Thus, the Opus codec was chosen for the further development of the VoIP-telephony system of the Southern Federal University.

We believe that once our program of implementation of the methods proposed in this work has been completed, the results will be of considerable interest to many organizations using corporate VoIP telephony systems that allow remote access of subscribers through networks of third-party telecommunications operators.

6. References

[1] Sukhov A M, Sultanov T G and Polukarov D Y 2012 The method of available bandwidth estimation of IP-connection based on variable packet size *Elektrosvyaz* **11** 39-42
[2] Bukatov A A, Zaitsev N D and Berezovsky A N 2017 Methods for monitoring and improving the quality of voice transmission for the system of corporate VoIP telephony SFU *Modern information technologies: trends and development prospects. Materials of the XXIV scientific conference* 47-52
[3] Calyam P, Ekici E, Lee C G, Haffner M and Howes N 2007 A "GAP-Model" based framework for online VVoIP QoE measurement *Journal of Communications and Networks* **9(4)** 446-456
[4] G.107: E-model: a computational model for use in transmission planning (Access mode: http://www.itu.int/rec/T-REC-G.107/) (2017-21-04)
[5] P.861: Objective quality measurement of telephone-band (300-3400 Hz) speech codecs (Access mode: https://www.itu.int/rec/T-REC-P.861/en/) (2017-21-04)
[6] P.862: Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs (Access mode: http://www.itu.int/rec/T-REC-P.862/en/) (2017-21-04)
[7] P.563: Single-ended method for objective speech quality assessment in narrow-band telephony applications (Access mode: http://www.itu.int/rec/T-REC-P.563-200405-I/) (2017-21-04)
[8] Kurittu A 2006 Validation of ITU-T P. 563 single-ended objective speech quality measurement *Journal of the Audio Engineering Society* **54(11)** 1092-1098
[9] Lukinskikh K S 2015 Unbiased Assessment Method For Voice Communication in Cloud VoIP-Telephony *Journal of Theoretical and Applied Information Technology* **80(1)** 75
[10] Gugel Yu V, Izhvanov Yu L, Kurakin D V 2010 Development of international connectivity of the federal computer network RUNNet *Proceedings of the XVII All-Russian Scientific and Methodical Conference "Telematics'2010"* 273-275
[11] Bolot J C 1993 Characterizing end-to-end packet delay and loss in the internet *Journal of High Speed Networks* **2(3)** 305-323
[12] Karam M J and Tobagi F A 2001 Analysis of the delay and jitter of voice traffic over the internet *Twentieth annual joint conference of the IEEE Computer and Communications Societies* **2** 824-833

[13] Singh H P, Singh S and Singh J 2011 Real time analysis of VoIP system under pervasive environment through spectral parameters *International Journal of Computer Applications* **31(2)** 1-8
[14] *Speex: A Free Codec For Free Speech* (Access mode: https://speex.org/) (2017-21-04)
[15] *G.712: Transmission performance characteristics of pulse code modulation channels* (Access mode: https://www.itu.int/rec/T-REC-G.712-199209-S/en/) (2017-21-04)
[16] *Opus Codec* (Access mode: http://opus-codec.org/) (2017-21-04)
[17] *RFC 6716 – Definition of the Opus Audio Codec* (Access mode: https://tools. ietf.org/html/rfc6716/) (2017-21-04)
[18] *SILK – Wikipedia* (Access mode: https://en.wikipedia.org/wiki/SILK/) (2017-21-04)
[19] *CELT: new audio codec from xipn.org* (Access mode: https://www.linux.org.ru/ news/ opensource/2601080/) (2017-21-04)

Acknowledgements
This work falls within the public tasks allotted to the Ministry of Education and Science of the Russian Federation (2.974.2017/4.6) and was carried out with the support of grant RFBR 16-07-00218a.

Is there an optimum in ad-hoc networks?

D Y Polukarov [1] and P O Chursin[1]

[1]Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. One of the most promising modern data transmission technologies is ad-hoc wireless peer-to-peer networks. A distinctive feature of such networks is the absence of a central node, which allows you to create connections "on the fly" with arbitrary connections between nodes. Ad-hoc networks are designed to create mobile computer networks and can be used, for example, to coordinate transport, in search and rescue operations and similar situations in which the use of infrastructure nodes is ineffective. When using Wi-Fi as a transport for self-organising sensor networks, attention should be paid to details. Wi-Fi technology is used in ad-hoc mode in self-organising sensor networks with non-hierarchical nodes. An experiment was conducted to measure the available bandwidth of a segment of the Wi-Fi network. This parameter was measured, and the dependence of the available bandwidth on the segment load was determined. We tried to create a tool for testing the Wi-Fi segment while also testing the tool's own performance in various configurations of the Wi-Fi segment.

1. Introduction

There is a large number of studies devoted to measurements in the Ethernet segment [1]. Also, a lot of work was done on the topic of measuring the parameters of the infrastructural mode of the Wi-Fi segment.

However, there is not enough attention paid to measuring the parameters of the Wi-Fi segment in the ad-hoc mode, yet this mode is important for organising mesh networks and other self-organising networks. Therefore, it makes sense to pay more attention to measuring the ad-hoc parameters of the Wi-Fi segment.

Ad-hoc mode allows network nodes to establish connections directly between devices without using a base station. Such networks are decentralised, and the definition of the node to which you want to send data is dynamic based on the connectivity of the network.

On the other hand, a network in infrastructure mode always consists of at least one access point and several wireless client stations connected to it. At the same time, data exchange between the client stations takes place through their forwarding through the access point.

When testing a network, you can't do without special software tools. There are many free and commercial products that solve this problem [2,3]. The comparative analysis of common tools for testing networks showed that, for all their advantages, the proposed programs have shortcomings that do not allow for fully performed functional testing of computer networks.

2. Related works

Nping [4] is an open source tool for generating network packets, analysing responses and measuring response times. Nping allows users to generate network packets from a wide range of protocols such as TCP, UDP, ICMP, and ARP, allowing them to configure virtually any area of protocol headers. While Nping can be used as a simple ping utility to detect active nodes in the network, it can also be

used as generator packages for stress tests of the network stack, sending ARP requests, attacks such as "denial of service", route tracking and other purposes. The program has a command-line interface.

Iperf3 [5] - cross-platform console client-server program-generator of TCP, UDP and SCTP traffic to test network bandwidth. The program allows you to perform load testing of a network segment, has a convenient command line interface, but does not have the ability to configure the package headers. The "iperf3-s"command is used to start the server. There is also a graphical shell called jperf [6].

Ostinato [7] is a multithreaded traffic generator designed to test services that enable the network to run at different levels of the network protocol stack. The user is given the opportunity to create data packages of arbitrary content, defining both the package title and the content of all its fields. In addition to the content of the packages, you can choose the interface and the frequency of traffic generation. Ostinato is a commercial product, though the cost of a single-user license is, however, relatively low. The disadvantage of the product is the heavy weight and lack of command line interface, which prevents it from installing on embedded systems.

Table 1. Comparative analysis of software products.

Feature	Nping	Iperf3	Ostinato	Our software
Protocols	TCP, UDP, ICMP, ARP	TCP, UDP, SCTP	TCP, UDP, ICMP, ARP	UDP
Configure packet headers	partially	no	yes	yes
Setting the packet size	yes	yes	yes	yes
Custom send interval (between packets)	yes	no	yes	yes
Select target speed	no	yes	yes	no
Changing traffic profiles “on the fly”	no	no	yes	yes

3. Features of ad-hoc networks

Wi-Fi segments commonly use CSMA/CA [8] for media access control — modification of pure Carrier Sense Multiple Access (CSMA). CSMA/CA differs from CSMA/CD [9] in that only jam-signals are vulnerable to collisions, but not the data packets. That is what “collision avoidance” stands for.

Collision avoidance is used to improve the performance of the CSMA method by attempting to divide the channel somewhat equally among all transmitting nodes within the collision domain. This functionality is assigned to the «jamming signal». Decrease in collision probability and the number of retransmissions leads to performance improvement, but waiting for the jam signal creates additional delays, thus making other methods able to achieve better results. Collision avoidance is quite useful in situations, when it is impossible to detect collisions in real-time – for example, when using radio transmitters [8].

The hidden node problem occurs when several nodes are visible to an access point (AP), but not to other nodes, communicating with this AP, thus making them unable to physically receive signals from each other (e.g. because of distance, signal spreading conditions, etc.). This leads to difficulties in the media access control sublayer, as the majority of existing client-side methods for accessing digital networks use carrier sensing to detect channel utilisation (CSMA/CD, CSMA/CA, etc.) [10].

4. Measurement of parameters

4.1. Fully connected ad-hoc Wi-Fi segment

The scheme of the experiment is shown in figure 1. The segment of the wireless ad-hoc network consisted of four nodes: two wireless Wi-Fi routers with special software installed and two computers used to measure the available bandwidth.

Wireless routers were used as sources of spurious traffic. Devices ran operating system OpenWRT [11], as well as specially developed software to generate spurious traffic. This software has the ability

to change the intensity of generated traffic. Changing the intensity of spurious traffic is achieved by adjusting the time interval between the packets being sent.

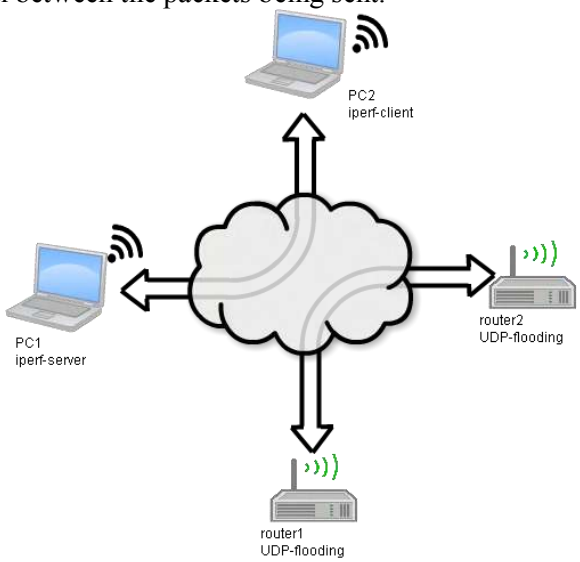


Figure 1. Scheme of the experiment.

Nodes PC1 and PC2 are set up for connection in ad-hoc mode and utilise iperf3 [5] to measure available bandwidth.

The obtained measurement results are presented in Table 2.

Table 2. Results of the experiment.

Delay, µs	Available bandwidth Bav, Mbit/s		
	Packet size V=500 bytes	Packet size V=1000 bytes	Packet size V=1500 bytes
10	2.97	2.35	1.36
500	7.30	6.48	6.30
2000	10.72	9.74	9.94
5000	11.85	11.00	11.34

The results of the measurements are also presented on the graphs shown in Figure 2.

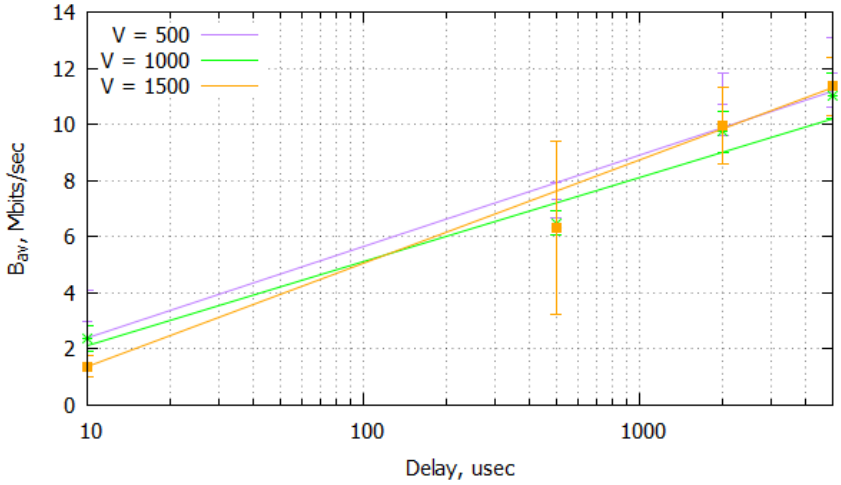


Figure 2. Dependence of bandwidth on segment load.

4.2. Ad-hoc Wi-Fi segment with hidden node problem

Figure 3 shows a diagram of an experiment that implements the hidden node problem. Special waveguides were used to limit the spread of the Wi-Fi signal.

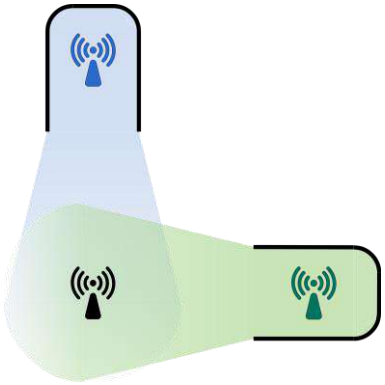


Figure 3. Scheme of the experiment with hidden node problem.

Portable Wi-Fi routers of type TL-MR3020 with an OpenWRT operating system installed on them were used as ad-hoc segment nodes. To prevent spurious signal leakage, the nodes were powered from autonomous power sources.

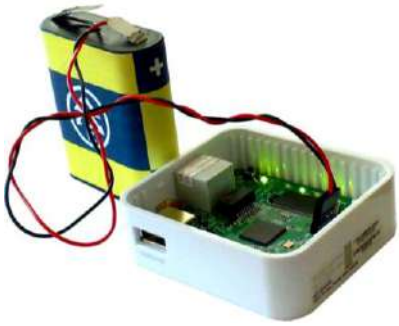


Figure 4. Portable Wi-Fi router TL-MR3020 as ad-hoc segment node.

Special waveguides are cardboard boxes covered with aluminum foil. This allowed for limiting the area of the Wi-Fi signal and significantly reducing the size of the experimental setup.

Figure 5. Photo of the experiment with hidden node problem.

Unfortunately, we were unable to capture the noticeable effect of the hidden node. However, this experiment allowed us to test the performance of our software in this mode.

5. Conclusion

In this paper, the bandwidth of the ad-hoc Wi-Fi network segment was measured under various load conditions. The measurements did not reveal a significant qualitative difference in the throughput of the Wi-Fi network segment in the ad-hoc mode from the infrastructure mode. The use of the developed tools allows us to determine the rational mode of operation of the Wi-Fi segment. Also, these results can be used to solve the optimal path problem [12] and other related works [13].

6. References

[1] Boggs D R, Mogul J C and Kent C A 1988 *Measured capacity of an ethernet: Myths and reality* **18(4)** 222-234

[2] Chursin P O and Polukarov D Yu 2017 Some features of the implementation of the source of test traffic *Perspective information technologies* 1033-1035
[3] Chursin P O and Polukarov D Yu 2018 Designing an automated information system for testing the functionality of network interaction *Perspective information technologies* 1154-1155
[4] *Nping - Network packet generation tool* (Access mode: https://nmap.org/nping/) (21.11.2017)
[5] *iPerf - Download iPerf3 and original iPerf pre-compiled binaries* (Access mode: https://iperf.fr/iperf-download.php/) (21.11.2017)
[6] *JPerf – graphical interface wrapper for Iperf* (Access mode: https://www.rarst.net/software/jperf/) (21.11.2017)
[7] *Ostinato Network Traffic Generator* (Access mode: https://ostinato.org/) (21.11.2017)
[8] *Carrier-sense multiple access with collision avoidance* (Access mode: https://en.wikipedia.org/wiki/Carrier-sense_multiple_access_with_collision_avoidance/) (21.11.2017)
[9] *Carrier-sense multiple access with collision detection* (Access mode: https://en.wikipedia.org/wiki/Carrier-sense_multiple_access_with_collision_detection/) (21.11.2017)
[10] *Hidden node problem* (Access mode: https://en.wikipedia.org/wiki/Hidden_node_problem/) (21.11.2017)
[11] *OpenWrt Project: Welcome to the OpenWrt Project* (Access mode: https://openwrt.org/) (21.11.2017)
[12] Agafonov A A and Myasnikov V V 2016 Method for the reliable shortest path search in time-dependent stochastic networks and its application to GIS-based traffic control *Computer Optics* **40(2)** 275-283 DOI: 10.18287/2412-6179-2016-40-2-275-283
[13] Nikitin V S, Semenov E I, Solostin A V, Sharov V G and Chayka S V 2016 Modeling the"smartlink connection" performance *Computer Optics* **40(1)** 64-73 DOI: 10.18287/2412-6179-2016-40-1-64-73

Acknowledgements
This work falls within the public tasks allotted to the Ministry of Education and Science of the Russian Federation (2.974.2017/4.6) and was carried out with the support of grant RFBR 16-07-00218a.

Improvement of resource efficiency of the catalytic isomerization process by mathematical modeling

A G Faskhutdinov[1], I V Akhmetov[2], A E Musina[2] and I M Gubaydullin[1,2]

[1]Institute of Petrochemistry and Catalysis of the Russian Academy of Sciences, Oktyabrya St. 141, Ufa, Russia, 450075
[2]Ufa State Petroleum Technological University, Kosmonavtov St. 1, Ufa, Russia, 450062

Abstract. The article examines main schemes of reaction networks of the catalytic isomerization of the pentane-hexane fraction. The selection of the kinetic model for the most complete description of the basic chemical processes based on the mechanism of the reactions proceeding is substantiated. The mathematical model is constructed on the basis of kinetic model. The foundations for the subsequent modeling of the entire chemical-technological process system are laid.

1. Introduction

In connection with the change-over of the domestic oil refining industry to the production of motor gasoline compliant with the Euro-5 and Euro-6 standards, the current task is to reduce the content of aromatic hydrocarbons and, in particular, benzene in motor fuel while preserving the octane value. It is known that the main amount of aromatics in motor gasolines is formed in the process of catalytic reforming. Catalytic isomerization of light paraffins, in its turn, makes it possible to obtain a high-octane component of motor gasoline with a minimum content of aromatics. The purpose of the process of catalytic isomerization of the pentane-hexane fraction is the production of high-octane components of automobile gasolines. This process has a high efficiency, because such low-octane oil components as fractions 62-70 ^{0}C, as well as catalytic reforming raffinates are used as raw materials. The process is carried out in a hydrogen environment in the presence of bifunctional catalysts [1, 2].

It should be noted that in the postwar years, during the design and construction of most installations, including reactors and devices, inadequate attention has been given to an intensive study of the mechanism of processes. The intensive development of such modern information technologies as parallel computing technology, artificial neural networks (ANN), 3D modeling, simulation modeling, etc., allow us to develop detailed kinetic models of complex processes. Then, based on these models, it is possible to increase significantly the yield of the target product, while minimizing energy and material resources [3].

There are often problems associated with the description of kinetics when studying refining processes. First of all it can be associated with a large component composition of raw materials, as well as the flow of a huge number of chemical reactions taking place on the acid and metal centers of bifunctional catalysts [4, 5]. In this connection, to simplify the description of the kinetics, the stage of formation of the intermediates that are not present in the composition of the target products is usually not considered.

2. Experimental

The object of the study is the reactor section of the catalytic isomerization unit of the pentane-hexane fraction, consisting of three series reactors. Hydrotreated gasoline fraction with a boiling range of 62-70 ° C, containing mainly pentane and hexane was used as a raw material.

The raw material load was 47.6 m3 / h. Consumption of hydrogen bearing gas (HBG) – 3400 nm^3/h. The composition of the fresh HBG is shown in Table. 1. The pressure in the reaction zone was 3.2 MPa. The temperatures at the inlet of the first, second and third reactors were 147, 160 and 145 ° C, respectively. The temperatures at the outlet of the first, second and third reactors were 160, 170 и 146 ^{0}C, respectively. The bifunctional catalyst of the domestic production SI-2 was used as the catalyst of the process. As noted, the reactor unit of the research object consists of three reactors in which the catalyst was distributed at 9,000 kg.

Table 1. The composition of fresh HBG.

Component	Mass fraction of components, % (GOST 14920)
Nitrogen	1.44
Hydrogen	83.61
Isobutane	0.4
Isopentane	0.99
Oxygen	0.58
Methane	8.18
n-butane	0.16
n-pentane	0.14
Propane	1.06
C_5H_{10} (sum)	0.43
Ethane	3.01

Characteristics of raw materials and product (stable isomerizate) are given in Table. 2 (operation data of the industrial unit for the catalytic isomerization of the pentane-hexane fraction).

Table 2. Characteristics of raw materials and product.

Analyzed indicator	Method	Value	
		Raw material	Stable isomerate
Density at 20 ° C, kg / m3	GOST R 51069	653.6	643.2
Octane number (research method)	GOST 8226	67	81
Fractional composition, ^{0}C	GOST 2177		
boiling point		42	36
90% boils away		62	59
97% boils away		64	63
end of boiling		65	64
Yield, %		98	98

The composition of raw material and product of the catalytic isomerization process of the pentane-hexane fraction is given in Table 3 shows (industrial unit data).

It is obvious that for developing a detailed kinetic model it is enough to model only a cascade of reactors. The composition of the gas product mixture after the reactor unit is calculated from the product data after the stabilization unit, that is available. Here it should be noted that with such a calculation, the indicators of productive flows have a large effect on the final result.

In such cases, the instruments error can reach 5% for flows in the gaseous state and 3% for the flows in the liquid state. Therefore, the use of such data without preliminary verification (balance on carbon and hydrogen) is absolutely unacceptable.

Table 3. The composition of raw material and product.

Component	Raw material, % weight	Isomerizate, % weight
2,2- DMB	0.36	21.13
2,3- DMB	2.15	6.57
2-MP	19.32	19.58
3-MP	14.73	11.23
Benzene	1.25	-
Isopentane	10.75	23.56
Methylcyclopentane	2.7	0.12
N-hexane	18.19	7.27
N-pentane	30.81	7.80
Isobutane	-	1.04
N-butane	-	0.87
Propane	-	0.61
Ethane	-	0.13
Methane	-	0.09

where 2-MP – 2-methylpentane, 3-MP – 3-methylpentane, 2.2-DMB – 2.2-dymethylbutane, 2.3-DMB – 2.3-methylbutane.

3. Scheme of reaction networks of the catalytic isomerization of the pentane-hexane fraction

The first stage of the mathematical description of the object is the scheming of hydrocarbon reactions during the process. It should be noted that the accuracy of calculations and the adequacy of the mathematical model to the actual process directly depend on the degree of detail of chemical transformations. Therefore, the transformation scheme should sufficiently reflect the physico-chemical nature of the process [6, 7].

In this article there was an attempt to make more detailed analysis of the scheme of reaction networks proceeding from the mechanism.

It is known that reactions in the process of isomerization in the presence of bifunctional catalysts proceed along the carbon-ion mechanism [8-10]. The thermodynamics of the process is the basis. The result of thermodynamic analysis is a list of reactions that probably proceed under given conditions [11, 12]. The probability of reactions under the process conditions of the catalytic isomerization of the pentane-hexane fraction (temperature-150 ° C (423 K) and pressure-3.2 MPa) was estimated from the value of Gibbs energy ΔG (Table 4).

From the Table 4 we can see that isomerization, dehydrocyclization, naphthene conversion, dehydrogenating reactions are reversible. Whereas hydrocracking reactions are irreversible [13].

According to the equations of reactions taking place in the process of catalytic isomerization of the pentane-hexane fraction, we write the kinetic equations as follows (Table 5).

Table 4. The reaction equations for the catalytic isomerization of the pentane-hexane fraction and the values of the thermodynamic characteristics.

	Reaction	ΔH, kJ/mol	ΔG, kJ/mol
1.	n-C_5H_{12}↔i-C_5H_{12}	-8.22	-5.76
2.	n-C_6H_{14}↔2-MP	-6.98	-3.77
3.	n-C_6H_{14}↔3-MP	-4.44	-7.60
4.	n-C_6H_{14}↔2.2-DMB	-18.53	-5.60
5.	n-C_6H_{14}↔2.3-DMB	-10.93	-0.99
6.	2-MP ↔3- MP	-2.54	-3.02
7.	2.2-DMB ↔2.3-DMB	-7.61	-4.61
8.	BZ +H_2↔ CH	-213.94	-52.65
9.	BZ +H_2↔ MCP	-197.03	-53.83
10.	CH ↔ MCP	-112.34	49.10
11.	n-C_6H_{14}↔ CH+H_2	8.57	7.82
12.	2-MP ↔ MCP + H_2	-66.96	-29.59

13.	3-MP ↔ MCP + H_2	-64.42	-26.58
14.	2.2-DMB ↔ MCP + H_2	-78.51	-31.43
15.	2.3-DMB ↔ MCP + H_2	-70.91	-26.82
16.	MCP + H_2→3-MP	-64.42	-26.58
17.	n-C_5H_{12}+ H_2→C_3H_8+C_2H_6	-44.49	-49.96
18.	n-C_5H_{12}+ H_2→n-C_4H_{10}+CH_4	-56.57	-61.66
19.	n-C_6H_{14}+ H_2→2C_3H_8	-37.94	-48.88
20.	n-C_6H_{14}+ H_2→n-C_4H_{10}+C_2H_6	-45.91	-51.80
21.	n-C_6H_{14}+ H_2→n-C_5H_{12}+CH_4	-56.14	-47.34
22.	n-C_6H_{14}+ H_2→i-C_5H_{12}+CH_4	-64.35	-66.71
23.	2-MP + H_2→i-C_4H_{10}+C_2H_6	-47.38	-49.82
24.	3-MP + H_2→i-C_5H_{12}+CH_4	-59.91	-65.95
25.	2.3-DMB +H_2→i-C_4H_{10}+C_2H_6	-43.44	-52.63
26.	2.3-DMB +H_2→i-C_5H_{12}+CH_4	-53.43	-65.71
27.	i-C_5H_{12}+ H_2 →i-C_4H_{10}+ CH_4	-56.82	-57.71
28.	2.2-DMB +H_2→i-C_4H_{10}+C_2H_6	-35.83	-48.05
29.	2.2-DMB +H_2→i-C_5H_{12}+CH_4	-45.82	-61.11

where 2-MP – 2-methylpentane, 3-MP – 3-methylpentane, 2.2-DMB – 2.2-dymethylbutane, 2.3-DMB – 2.3-methylbutane, CH – cyclohexane, MCP – methylcyclopentane, BZ – benzene.

Table 5. Kinetic equations of the process of catalytic isomerization of pentane-hexane fraction.

	Reaction	Kinetic equations
1.	n-C_5H_{12}↔i-C_5H_{12}	$W_1=k_1*x_1-k_2*x_2$
2.	n-C_6H_{14}↔2-MP	$W_2=k_3*x_3-k_4*x_4$
3.	n-C_6H_{14}↔3-MP	$W_3=k_5*x_3-k_6*x_5$
4.	n-C_6H_{14}↔2,2-DMB	$W_4=k_7*x_3-k_8*x_6$
5.	n-C_6H_{14}↔2,3-DMB	$W_5=k_9*x_3-k_{10}*x_7$
6.	2-MP ↔3-MP	$W_6=k_{11}*x_4-k_{12}*x_5$
7.	2.2-DMB ↔2.3-DMB	$W_7=k_{13}*x_6-k_{14}*x_7$
8.	BZ +H_2↔ CH	$W_8=k_{15}*x_8*x_9^3-k_{16}*x_{10}$
9.	BZ +H_2↔ MCP	$W_9=k_{17}*x_8*x_9^3-k_{18}*x_{11}$
10.	CH ↔ MCP	$W_{10}=k_{19}*x_{10}-k_{20}*x_{11}$
11.	n-C_6H_{14}↔ CH+H_2	$W_{11}=k_{21}*x_3-k_{22}*x_{10}*x_9$
12.	2-MP ↔ MCP + H_2	$W_{12}=k_{23}*x_4-k_{23}*x_{11}*x_9$
13.	3-MP ↔ MCP + H_2	$W_{13}=k_{25}*x_5-k_{26}*x_{11}*x_9$
14.	2.2-DMB ↔ MCP + H_2	$W_{14}=k_{27}*x_6-k_{28}*x_{11}*x_9$
15.	2.3-DMB ↔ MCP + H_2	$W_{15}=k_{29}*x_7-k_{30}*x_{11}*x_9$
16.	MCP + H_2→3-MP	$W_{16}=k_{31}*x_{11}*x_9$
17.	n-C_5H_{12}+ H_2→C_3H_8+C_2H_6	$W_{17}=k_{32}*x_1*x_9$
18.	n-C_5H_{12}+ H_2→n-C_4H_{10}+CH_4	$W_{18}=k_{33}*x_1*x_9$
19.	n-C_6H_{14}+ H_2→2C_3H_8	$W_{19}=k_{34}*x_3*x_9$
20.	n-C_6H_{14}+ H_2→n-C_4H_{10}+C_2H_6	$W_{20}=k_{35}*x_3*x_9$
21.	n-C_6H_{14}+ H_2→n-C_5H_{12}+CH_4	$W_{21}=k_{36}*x_3*x_9$
22.	n-C_6H_{14}+ H_2→i-C_5H_{12}+CH_4	$W_{22}=k_{37}*x_3*x_9$
23.	2-MP + H_2→i-C_4H_{10}+C_2H_6	$W_{23}=k_{38}*x_4*x_9$
24.	3-MP + H_2→i-C_5H_{12}+CH_4	$W_{24}=k_{39}*x_5*x_9$
25.	2.3-DMB +H_2→i-C_4H_{10}+C_2H_6	$W_{25}=k_{40}*x_7*x_9$
26.	2.3-DMB +H_2→i-C_5H_{12}+CH_4	$W_{26}=k_{41}*x_7*x_9$
27.	i-C_5H_{12}+ H_2 →i-C_4H_{10}+ CH_4	$W_{27}=k_{42}*x_2*x_9$
28.	2.2-DMB +H_2→i-C_4H_{10}+C_2H_6	$W_{28}=k_{43}*x_6*x_9$
29.	2.2-DMB +H_2→i-C_5H_{12}+CH_4	$W_{29}=k_{44}*x_6*x_9$

where x_i – concentration of components in mole fractions, x_1 – n-pentane, x_2 – isopentane, x_3 – n-hexane, x_4 – 2-MP, x_5 – 3-MP, x_6 – 2,2-DMB, x_7 – 2,3-DMB, x_8 – benzene, x_9 – hydrogen, x_{10} – cyclohexane, x_{11} – methylcyclopentane, x_{12} – propane, x_{13} – ethane, x_{14} – н-butane, x_{15} – methane, x_{16}

– isobutane, k_j – kinetic constant of the j-th reaction, $L \cdot mol^{-1} \cdot h^{-1}$ (j=8, 9, 11-29), h^{-1} (j=1-7, 10), W_j – rate of the j-th reaction, mol/(l·h).

The reaction rates entering into the kinetic model (Table 5) were written according to the law of mass action [14]. The mathematical model of the process is a system of nonlinear differential equations [15, 16]:

$$\begin{cases}
\frac{dx_1}{dt} = -W_1 - W_{17} - W_{18} + W_{21} \\
\frac{dx_2}{dt} = W_1 + W_{22} + W_{23} + W_{24} + W_{26} - W_{27} + W_{29} \\
\frac{dx_3}{dt} = -W_2 - W_3 - W_4 - W_5 - W_{11} - W_{19} - W_{20} - W_{21} - W_{22} \\
\frac{dx_4}{dt} = W_2 - W_6 - W_{12} - W_{23} \\
\frac{dx_5}{dt} = W_3 + W_6 - W_{13} + W_{16} - W_{24} \\
\frac{dx_6}{dt} = W_4 - W_7 - W_{14} - W_{28} - W_{29} \\
\frac{dx_7}{dt} = W_5 + W_7 - W_{15} - W_{25} - W_{26} \\
\frac{dx_8}{dt} = -W_8 - W_9 \\
\frac{dx_9}{dt} = -3W_8 - 3W_9 + W_{11} + W_{12} + W_{13} + W_{14} + W_{15} - W_{16} - W_{17} - W_{18} - W_{19} - W_{20} - \\
- W_{21} - W_{22} - W_{23} - W_{24} - W_{25} - W_{26} - W_{27} - W_{28} - W_{29} \\
\frac{dx_{10}}{dt} = W_8 - W_{10} + W_{11} \\
\frac{dx_{11}}{dt} = W_9 + W_{10} + W_{12} + W_{13} + W_{14} + W_{15} - W_{16} \\
\frac{dx_{12}}{dt} = W_{17} + 2W_{19} \\
\frac{dx_{13}}{dt} = W_{17} + W_{25} + W_{27} + W_{28} \\
\frac{dx_{14}}{dt} = W_{18} \\
\frac{dx_{15}}{dt} = W_{18} + W_{21} + W_{22} + W_{23} + W_{24} + W_{26} + W_{29} \\
\frac{dx_{16}}{dt} = W_{25} + W_{27} + W_{28}
\end{cases}$$

With initial conditions t=0, $x_i(0)=x_{i,0}$, where i – corresponding hydrocarbon.

Thus, the dimension of the system of a mathematical model equations is coincides with the number of substances whose concentration is determined in calculations.

To apply the obtained model, it is first necessary to make a comparative assessment of the kinetic parameters.

When solving the inverse kinetic task, it is planned to use the method of parabolic descent. It is also planned to use a modified Kutta-Merson method of the fifth order of accuracy to solve the system of ordinary nonlinear differential equations [17].

The solution algorithm of a system of differential equations $\frac{dx}{dt} = f(x)$ requires a fivefold calculation of the right-hand sides f(x) [18]:

$$F_1=f(x_k)\cdot h,$$
$$F_2=f(x_k+F_1/3)\cdot h,$$
$$F_3=f(x_k+(F_1+F_2)/6)\cdot h,$$
$$F_4=f(x_k+(F_1+3\cdot F_2)/8)\cdot h,$$
$$x_{k+1}=x_k+(F_1-3\cdot F_2+4\cdot F_4)/2,$$
$$F_5=f(x_{k+1})\cdot h,$$
$$x^*_{k+1} = x_k + (F_1 + 4\cdot F_4 + F_5),$$

where h – step of integration over a spatial or temporal coordinate, x_k – values of variables at the beginning of the integration interval. At the end of the integration interval, two solutions are obtained: x_{k+1} – by the Runge-Kutta method with accuracy $O(h^4)$ and x^*_{k+1} – by the Kutta-Merson method with accuracy $O(h^5)$.

4. Findings

- With the help of calculations, it has been shown that all reactions occurring in the transformation scheme is probable from the point of view of thermodynamics under the conditions of the process.
- The detailed mathematical model is proposed, consisting in consideration of all components of the raw material fraction with the boiling range of 62-70 ° C (pentane-hexane fraction).

5. Conclusion

Development on the basis of the reaction mechanism of an adequate kinetic model is an important stage in the creating of a mathematical model. Due to the fact that by solving inverse kinetic tasks, the parameters that will serve as the basis for the mathematical model are determined, and also taking into account physicochemical aspects of the process, the mathematical model will allow to carry out predicted calculations, to select the optimal technological conditions for increasing the resource efficiency of the catalytic isomerization process of pentane-hexane faction.

6. References

[1] Ahmetov S A 2002 *Technology of advanced oil and gas refining* (Ufa: Gilem) p 672
[2] Koncsa C I, Tutun I and Safta C 2011 Study of C5/C6 isomerization on Pt/H-zeolite catalyst in industrial conditions *Ovidius Univ. Annal. Chem.* **22** 102-106
[3] Bursian N R 1985 *Technology of isomerization of paraffinic hydrocarbons* (Moscow: Chemistry) p 192
[4] Stijepovic M Z, Ostojic A V, Milenkovic I and Linke P 2009 Development of a kinetic model for catalytic reforming of naphtha and parameter estimation using industrial plant data *Energy Fuels* **23** 979-983
[5] Ismagilov I Z, Matus E V, Nefedova D V, Kuznetcov V V, Yashnik S A, Kerzhentcev M A and Ismagilov Z R 2015 Effect of support modification on the physicochemical properties of a $NiPd/Al_2O_3$ catalyst for the autothermal reforming of methane *Kinetics and Catalysis* **56** 397
[6] Belinskaya N S, Silko G U, Frantcina E V, Ivashkina E N and Ivanchina E D 2013 Development of a formalized scheme for the transformation of hydrocarbons and the kinetic model of the process of hydrodewaxing diesel fuels *Bulletin of the Tomsk Polytechnic University* **322** 129-133
[7] Chuzlov V A, Ivanchina E D, Dolganov I M and Molotov K V 2015 Simulation of light naphtha isomerization process *Procedia Chemistry* **15** 282-287
[8] Barron Y, Maire G, Muller J M and Gault F G 1966 The mechanisms of hydrogenolysis and isomerization of hydrocarbons on metals *Journal of Catalysis* **5** 428-445

[9] Ono Y 2003 A survey of the mechanism in catalytic isomerization of alkanes *Catal. Today* **81** 3-16
[10] Liu H, Meng X, Zhang R, Liu Z, Meng J and Xu C 2010 Reaction performance and disappearance kinetics of n-pentane isomerization catalyzed by chloroaluminate ionic liquid *Catal. Commun* **12** 180-183
[11] Dymond J H, Young K J and Isdale J D 1979 *Chem. Thermodyn.* p 887
[12] Maginn E J, Theodorou D N and Bell A T 1995 Sorption thermodynamics, siting and conformation of long n-alkanes in silicalite as predicted by configurational-bias Monte Carlo integration *Phys. Chem.* **99** 2057-2079
[13] Triwahyono S, Jalil A A, Mukti R R, Musthofa M, Razali N A M and Aziz M A A 2011 Hydrogen spillover behavior of Zn/HZSM-5 showing catalytically active protonic acid sites in the isomerization of n-pentane *Appl. Catal.* (A: Gen) **407** 91-99
[14] Khusnutdinov R I, Baiguzina A R, Mukminov R R, Akhmetov I V, Gubaidullin I M, Spivak S I and Dzhemilev U M 2010 New synthesis of pyrrole-2-carboxylic and pyrrole-2,5- dicarboxylic acid esters in the presence of iron-containing catalysts *Russian Journal of Organic Chemistry* **46** 1053-1059
[15] Koledina K F and Gubaydullin I M 2016 Kinetics and mechanism of olefin catalytic hydroalumination by organoaluminum compounds *Zhurnal fizicheskoi khimii* **90** 671-678
[16] Gubaydullin I M, Koledina K F and Lind Yu B 2011 Modern high-performance computing technologies for modeling the detailed mechanism for the reaction of catalytic hydroalumination of olefins *Nauka I obrazovanie: nauchnoe izdanie MGTU im. N.E. Bauman* **6** 10
[17] Akhmetov I V and Gubaydullin I M 2016 Analysis of methods for solving inverse problems of chemical kinetics with the use of parallel computing *CEUR Workshop Proceedings* **1576** 402-410
[18] Akhmetov I V and Gubaydullin I M 2017 Parallel calculations in the construction of the kinetic model of benzylidene benzylamine synthesis *CEUR Workshop Proceedings* **1902** 1-4

Methods of signal processing and construction of diagnostic matrixes onset by sleep apnea treatment equipment

N V Ivakhno[1], S I Zykin[1] and S V Antsibor[1]

[1]Tula State University, Lenina ave. 92, Tula, Russia, 300012

Abstract. The paper deals with the problem of data processing in adaptive detection of inspiration / expiration by machines for treating sleep apnea using statistical decision theory and based on preliminary analysis of human respiration. The authors establish a law of noise distribution and develop an inspiration and expiration detection algorithm which envisages calculation of the likelihood ratio, which is compared with the threshold values, at each step. As a result, a conclusion is drawn, and a decision is taken on the need to initiate the treatment of sleep apnea with the machine. The use of this algorithm reduces the detection time by 2-3 times, making it possible to carry out preliminary adjustment of parameters for each patient. The problem of definition of person s respiratory system condition with the use of the methods based on a task of the dosed values of resistance/pressure of switching in a respiratory contour with the subsequent creation of diagnostic matrixes state, in which every line characterizes parameters value at a certain loading for the throttle and relay modes of complexes correcting influence is solved.

1. Introduction

Sleep apnea is the cessation of pulmonary ventilation (respiratory arrest) during sleep for more than 10 seconds. Usually it lasts for 20–30 seconds, though it may be as long as 2–3 minutes [1,2].

Apnea can be classified as central, obstructive, or mixed. In the first case, the respiratory arrests during sleep are caused by the disorder of brain function caused by congenital abnormalities or brain injuries. Obstructive sleep apnea (OSA) is a sleep disorder that occurs when the soft tissues in the back of the throat (upper airway) become narrow, and the muscles naturally relax [1,3,4], which results in reduced supply of oxygen to the lungs.

Various devices are used to detect apnea [1,2], which make it possible to identify respiratory pauses, episodes of oxygen starvation, snoring, and other symptoms. For the treatment of sleep apnea [1, 2], they mainly use the respiratory ventilation mode known as positive airway pressure (PAP). A CPAP machine is a small compressor supplying a constant flow of air into the respiratory tract at a predetermined pressure through a flexible tube and an airtight nasal mask [2,3], which prevents the respiratory tract from occluding and blocking the flow of air (and thus of oxygen the organism needs).

A crucial role during treatment is played by CPAP equipment. Speedy recovery of the patient depends to a great extent on the effectiveness and reliability of medical equipment. The advisability of real-time pressure adjustment is predetermined by the fact that therapeutic pressure changes depending on the body position and the sleep stage. During the stage of deep sleep, as well as when the person is sleeping on the back, a significantly greater pressure is required to open the airway as compared with the stage of surface sleep and sleeping on the side, respectively. Thus, in order to control the motor of

the CPAP machine, it is necessary to provide adaptive functioning of the algorithms adjusting to each individual patient.

Analysis suggests that the existing models of CPAP machines do not take into account changes in the human condition, as well as the interaction processes taking place in the biotechnical system "machine – patient". The treatment of sleep apnea syndrome requires high precision adjustment of the initial parameters and synchronization of the functioning of the CPAP machine and the patient.

One of the modules providing the determination of the initial parameters of the system and affecting the accuracy of synchronization is the detector of the onset of inspiratory and expiratory activity functioning as part of the operation modes of a sleep apnea treatment machine.

2. Model of the process of recognizing the onset and end of respiratory activity

Recognition of the beginning and the end of respiratory activity is usually carried out based on a given pressure value; however, at a fixed pressure, due to noises, a delay is observed in determining the onset of inspiration and expiration, and with weak respiration the decision to switch the CPAP machine motor on for treating sleep apnea is made almost in the middle of the respiratory cycle, which leads to significant desynchronization of operation [5,6,7].

Therefore, in order to establish a decision-making criterion, we have developed an adaptive data processing algorithm based on modern statistical decision theory and taking into account the distribution law and the parameters of the useful signal and noise measured during automatic adjustment of the system [5,8].

In order to obtain the mathematical model of signal processing in anticipation of inspiratory / expiratory activity, experimental studies have been carried out of the breathing of various patients with the purpose of determining the law of noise distribution. The empirical data obtained were used to build histograms and to put forward the hypothesis about the normal noise distribution law, the likelihood of which was confirmed using chi-squared goodness-of-fit test.

According to research conducted in [2,4,6], a mathematical model of the recognition process was established in which the likelihood ratio is obtained at each step of observation:

$$\ln \Lambda^{(m)} = \rho_0 \left(\frac{1}{a} \sum_{i=1}^{m} z_i - \frac{m}{2}\right). \quad (1)$$

where $\rho_0 = \frac{a^2}{\sigma^2}$ is signal / noise ratio, m is the count number, $z_i = y_i - a0$, $a = as - a0$, as is the input signal amplitude, $\sigma^2, a0$ are the variance and the average value of noise, y_i is the measured value of the pressure in the breathing circuit [2,9]. The value $\ln \Lambda^{(m)}$ is compared with two constant threshold values: A and B (A> B), which are found based on the predetermined conditional probabilities α^* and β^* of detection ahead of time (false alarm) and of delayed inspiration / expiration detection (signal omission) [1,2].

If the decision is taken that there is no signal, the analysis is repeated until the onset of inspiration / expiration is detected in accordance with the limitation interval.

For the purposes of both the simulation modeling and the subsequent implementation of the developed model of inspiratory / expiratory activity onset detection, the probability of signal omission was found based on the analysis of respiratory rate and efficiency of sleep apnea treatment procedure $\alpha^* = 0.01$, and the probability of false alarm, given the architecture and the function of CPAP machines, was set to $\beta^* = 10^{-3}$. In this case, $\ln A = 4.605$, $\ln B = -6.89$ [1,10].

3. Algorithm of data processing in detecting the inspiration onset

According to the obtained formula (1), the sleep apnea treatment complex comprising a pressure measurement unit was supplemented by additional elements making it possible to implement the adaptive method of detecting the onset of inspiration / expiration taking into account random external disturbances characteristic for each patient [4,5,11].

Given the data obtained in experiments, we established the time of parameter adjustment, which amounts to 1-1.5 minutes, on the basis of the predetermined confidence level and the variance [2,3].

Automatic parameter adjustment allows implementation of the mathematical model of signal processing in detecting the onset of inspiration / expiration. The functioning of the algorithm can be represented as a block diagram (Fig. 1).

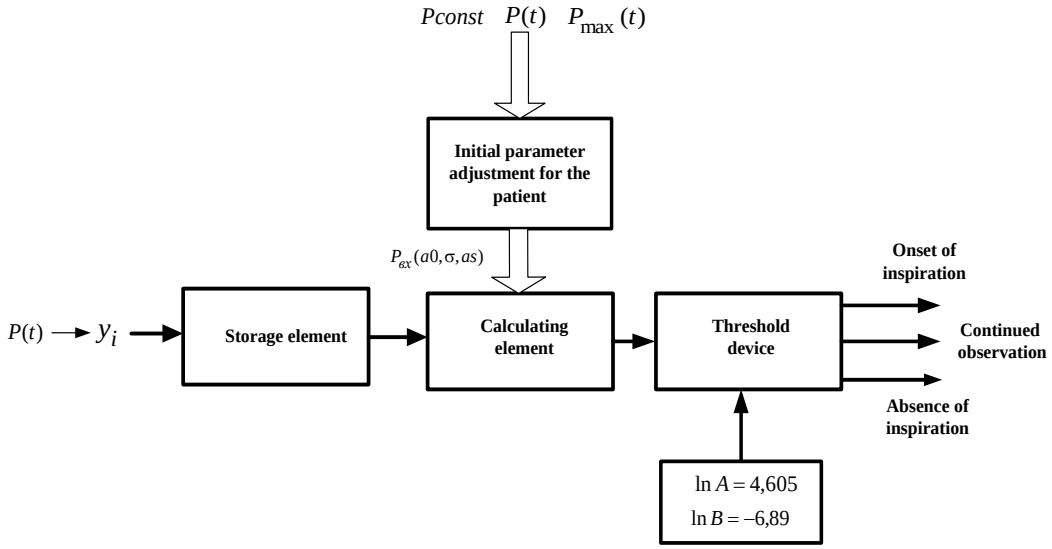


Figure 1. A block diagram of the method of data processing when detecting the onset of inspiration.

The initial parameters are as follows: the values *lnA* and *lnB* – constants for various categories of patients; $\sigma, a0$ - the standard deviation and the average noise value (found during the automatic adjustment of parameters, with the patient breathing independently); as – the value of the useful signal, found during the automatic parameter adjustment as

$$q1 \cdot P_{\max}(t) + Pconst ,$$

where $q1$ is a variable coefficient, the value of which, as a rule, may amount to 0.01÷0.2 depending on the maximum signal value (the ratio decreases with increasing amplitude), $P_{\max}$ is the average maximum value of inspiratory / expiratory pressure found during the adjustment of parameters for each person, *Pconst* is the pressure level against which the measurements are taken, relative zero.

The device for detecting the beginning of the patient's inspiration / expiration operates as follows (Figure 2): y_i - the measured value of the pressure in the breathing circuit is summed with the next count obtained (storage element), and the calculating element calculates the likelihood ratio $\ln \Lambda^{(m)}$ at each stage of observation, taking into account $a0, \sigma$, and as. The threshold device generates a conclusion about whether the beginning of the inspiration / expiration has been detected, or whether observation should be continued. Signal processing is carried out in the same way when detecting the end of inspiration / expiration.

The results of experimental studies on the calculation of likelihood ratio value are presented in Fig. 2 and Fig. 3.

4. Diagnostic matrixes of the respiratory system condition for the correcting influence complexes

After recognition of the beginning and the end of respiratory activity hardware has a controlling influence, at the same time should be diagnosis of the state of the human respiratory system, result processing in real time and adjusting loads.

A comprehensive approach based on parametric analysis of the dependence of the pressure at various modes of functioning of the complexes has a correcting effect on the respiratory system, includes specifying a load in two modes – the throttle and relay that determines the change in the resistance of breathing circuit [3]. Under throttle type refers to the task of resistance in the breathing circuit in the form of restriction of the cross-sectional area of the breathing tube under the relay –

when the initial complete overlap of the respiratory tube at the beginning of the inhalation/exhalation – the full opening of the valve when it reaches a certain pressure (pressure switch).

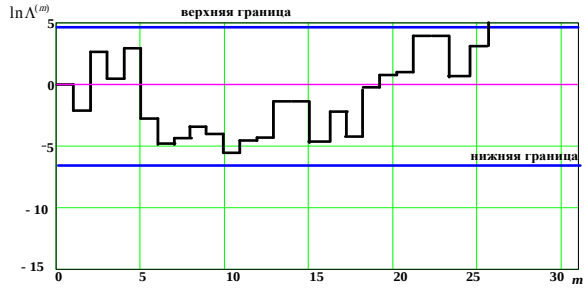


Figure 2. The values of likelihood ratio at each step of observation with respect to the decision limits (onset of inspiration registered at count 26).

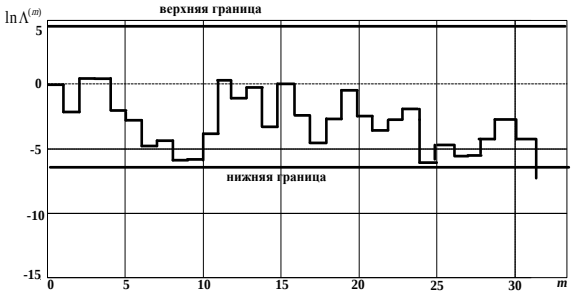


Figure 3. The values of likelihood ratio at each step of observation with respect to the decision limits (onset of inspiration registered at count 31).

The pressure values were measured in the breathing tube of a model according to the structural scheme presented in [2,3,4,11].

To conduct a parametric analysis of the dependence of the pressure measurement results were normalized relative to the maximum possible value of pressure measured during the initial studies of the respiration of each person,

$$P(t) = \frac{P_{meas}(t)}{P_{\max}},$$

where $P(t)$ - normalized characteristic pressure; $P_{meas}(t)$ - measured characteristic pressure in the respiratory tube according to the structural scheme presented in [3,4,5,11,12]; $P_{\max}$ - maximum pressure of inhalation/exhalation fixed at the time of the initial investigation.

To select the type of exposure variable is introduced which takes two values $j = 0, j = 1$ depending on the relay or the throttle operation. To build a matrix of conditions of the respiratory system enter the number of levels of load impacts N. Sequential changes in the load/pressure is indicated by $i = 0,1...N$.

Every exposure causes (table 1): at throttle type – change of cross-sectional area relative to the original in $\frac{100}{N+1} i$ %, the relay type (load change pressure switch) – change the pressure switch on $\frac{0,7P_{\max}}{N} i$.

Table 1. An example of setting exposure levels when $N = 4$

i = 0,1…4	0	1	2	3	4
The establishment of resistance (the area of overlap)	$R_0 = 0$	$R_1 = 20$	$R_2 = 40$	$R_3 = 60$	$R_4 = 80$
Pressure switch	P_0	$P_1 = \frac{0,7P_{\max}}{4}$	$P_2 = \frac{0,7P_{\max}}{2}$	$P_3 = \frac{0,7P_{\max}}{4}3$	$P_4 = 0,7P_{\max}$
			min load		*max* load

View graphs of the normalized pressure obtained at different load resistances, shown in Fig.4. As a result of analysis of numerous experimental data revealed that the relationship between the specified resistance and measured characteristics is not linear.

As can be seen from Fig. 4, the phase of inhalation/exhalation can be divided into two sections and approximated by linear functions, the slope of which carries information about the state of the respiratory system of the patient.

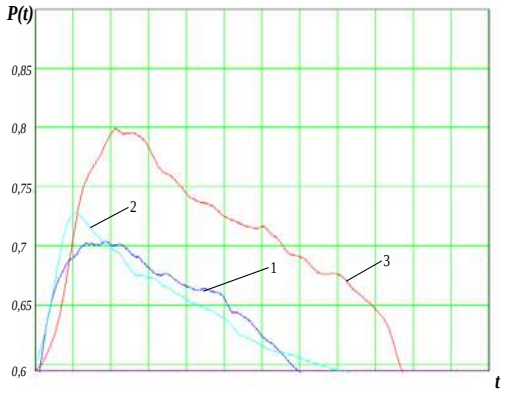


Figure 4. Examples of the different characteristics of the pressure in the breathing circuit at the throttle type of impact (expiratory phase): 1 - easy breathing $R_0 = 0$; 2 – breathing through a resistance $R_1 = 0{,}2R$; 3 – breathing through the resistance - $R_3 = 0{,}6R$

Thus, a generalized mathematical description when the throttle type of impact is represented by the combination of two linear functions:

$$P_i(t) = \begin{cases} \alpha_i t + b_i, at \;\; 0 < t \le t_{ri}, \\ \beta_i t + c_i, at \; t_{Hi} < t \le T_i - t_{ri}, \end{cases}$$

where α_i, β_i - angles when increasing and decreasing the approximating function of the i-th load; t_{ri} - the rise time of the pressure curve up to the maximum; b_i, c_i - coefficients; T_i - the duration of the inspiratory phase/expiratory (Fig.4), $P_i(t)$ - normalized pressure characteristic.

A generalized mathematical description of the resulting characteristics of the pressure at relay type of impact (Fig. 5) is expressed by the following functional characteristics:

$$P_i(t) = \begin{cases} \alpha_i \cdot t + b_i \quad at \;\; t_1 < t < t_2, \\ \alpha 1_i \cdot t + b1_i \quad at \;\; t_2 < t < t_3 \\ \beta p_i \cdot \ln(t) + b2_i \quad at \;\; t_3 < t < t_4, \end{cases},$$

where $\alpha_i, \alpha 1_i, \beta p_i, b_i, b1_i, b2_i$ - the coefficients of the approximating functions of the i-th load; $t_1 - t_2, t_2 - t_3, t_3 - t_4$ - intervals partitioning functions.

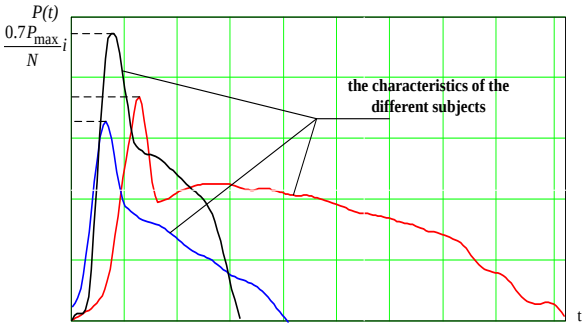


Figure 5. Examples of the different characteristics of the pressure in the breathing circuit at the relay type of impact (expiratory phase).

Thus, the variable parameters characterizing the state of the human respiratory system when exposed to different levels of resistance and different pressure switch $R_0, ..., R_N$, $P_0, ..., P_N$ are: the duration of the inspiratory phase/expiratory $T_1, ..., T_N$, the inclination angle of the approximating

curve on the first section of the observation $\alpha_1,..,\alpha_i..,\alpha_N$, angle during the descending of the approximating function β_i (with the load in the form of resistance), the coefficient of the approximating function in the third monitoring interval $\beta p_1,...,\beta p_N$ (when the load in the form of a pressure switch), the rise time of the pressure curve to the maximum $t_{r1},...,t_{rN}$, the characteristics of the pressure curve under free breathing $\alpha_0,\beta p_0,t_{r0},T_0$ (load shift pressure), $\alpha_0,\beta_0,t_{r0},T_0$ (load resistance) [6,7].

Then the generalized matrix characterizing the state of a person, his level of fitness when exposed to resistance and pressures, described so:

$$M=\begin{bmatrix} \alpha_0 & \beta_0 & t_{r0} & T_0 \\ . & . & . & . \\ \alpha_i & \beta_i & t_{ri} & T_i \\ \alpha_N & \beta_N & t_{rN} & T_N \end{bmatrix}, \qquad M1=\begin{bmatrix} \alpha_0 & \beta p_0 & t_{r0} & T_0 \\ . & . & . & . \\ \alpha_i & \beta p_i & t_{ri} & T_i \\ \alpha_N & \beta p_N & t_{rN} & T_N \end{bmatrix}.$$

The General criterion to establish the type of control action is formed by the conjunction of a number of parameters that make up the matrix of States of M and M1 and characterizes the variability of the condition of the human respiratory system with different types of impact, which is determined by the ratio:

$$K1_j^2=\begin{cases} \frac{1}{N}\sum_{k=0}^{s-1}\sum_{i=1}^{N} c_k\cdot\left(\frac{M1_{ik}-M1_{i^*k}}{M1_{0k}}\right)^2 \text{ at } j=1, \\ \frac{1}{N}\sum_{k=0}^{s-1}\sum_{i=1}^{N} c_k\cdot\left(\frac{M_{ik}-M_{i^*k}}{M_{0k}}\right)^2 \text{ at } j=0. \end{cases}$$

where c_1,c_2,c_3,c_4 - weights characterizing the importance of each indicator determined by expert; j - is the variable that determines the type of impact, k - is the parameter number ($k=0,...,S-1$, S – number of parameters, i - is the level of exposure $i=1,...,N$, $i^*=i-1$ - previous exposure).

Thus, the total set of informative features from different types of impacts (throttle and relay) is characterized by two matrices M and M1, which are necessary to identify the state of the respiratory system and formation of control actions [8,9]. These characteristics can be applied to the evaluation of corrective action. The General criterion to establish the type of control action, characterizes the variability of the condition of the human respiratory system with different types of exposure, is formed on the basis of indicators:

$$K1_j^2=\begin{cases} \frac{c_4}{N}\sum_{i=1}^{N}\frac{(T_i-T_{i^*})^2}{T_0^2}+\frac{c_3}{N}\sum_{i=1}^{N}\frac{(t_{ri}-t_{ri^*})^2}{t_{r0}^2}+\frac{c_2}{N}\sum_{i=1}^{N}\frac{(\beta p_i-\beta p_{i^*})^2}{\beta p_0^2}+ \\ +\frac{c_1}{N}\sum_{i=1}^{N}\frac{(\alpha_i-\alpha_{i^*})^2}{\alpha_0^2}, \text{ at } j=1, \\ \frac{c_4}{N}\sum_{i=1}^{N}\frac{(T_i-T_{i^*})^2}{T_0^2}+\frac{c_3}{N}\sum_{i=1}^{N}\frac{(t_{ri}-t_{ri^*})^2}{t_{r0}^2}+\frac{c_2}{N}\sum_{i=1}^{N}\frac{(\beta_i-\beta_{i^*})^2}{\beta_0^2}+ \\ +\frac{c_1}{N}\sum_{i=1}^{N}\frac{(\alpha_i-\alpha_{i^*})^2}{\alpha_0^2}, \text{ at } j=0, \end{cases}$$

where c_1, c_2, c_3, c_4 - the weights characterize the importance of each indicator, j - is the variable that determinesthe type of impact, i - impact level ($i = 1,...,N$), $i^* = i - 1$ - previous exposure.

If we denote the number of the parameter k ($k = 0,...,S-1$, S - is the number of parameters) General calculation specified ratio can be obtained based on the components of the matrices of the States of the respiratory system M and M1 for different numbers of parameters and different numbers of levels of effects [10,11,12,13].

Then, the calculation of this coefficient will describe the formulas [9,10,14,15]:

$$K1_j^2 = \begin{cases} \frac{1}{N} \sum_{k=0}^{s-1} \sum_{i=1}^{N} c_k \cdot \left(\frac{M1_{ik} - M1_{i^*k}}{M1_{0k}} \right)^2 \text{ at } j = 1, \\ \\ \frac{1}{N} \sum_{k=0}^{s-1} \sum_{i=1}^{N} c_k \cdot \left(\frac{M_{ik} - M_{i^*k}}{M_{0k}} \right)^2 \text{ at } j = 0. \end{cases}$$

After the preliminary diagnostics of the condition of the respiratory system throttle type (load – resistance), set ratio $K1$, the initial effect of throttle $(j = 0)$ and compared with the coefficient, determined experimentally for a group of patients whose characteristics change when holding the throttle pressure, which is expressed combined ratio $K1_{0\,\det}$ (at peak effect of this factor - $K1_{1\,\det}$). When $K1_0 > K1_{0\,\det}$ the throttle is applied the method of the impact of adaptive selection of a load, when $K1_0 < K1_{0\,\det}$ applied to the relay effect of shifting the pressure, or additionally is a factor $K1_1$ $(j = 1)$ and by comparing the proximity found in the process of diagnosis of coefficients to be determined $K1_{1\,\det}$ and $K1_{0\,\det}$, select the effect, which accounts for the maximum human reaction.

Example of finding the coefficients $K1_0$ of equal weights of 0.25, are shown in table 2 at $j = 0$, $N = 4$.

Table 2. Data to find the coefficient of $K1_0$

Load characteristic	i = 0	i = 1	i = 2	i = 3	i = 4
	R0=0	R1=0.2R	R2=0.4R	R3=0.6R	R4=0.8R
α_i	0.139	0.186	0.186	0.187	0.188
β_i	-0.003	-0.003	-0.003	-0.003	-0.003
t_{r2}	0.5	0.6	0.61	0.62	0.62
T_i	1	1.2	1.3	1.5	1.6

The calculation results show the value of the parameter $K1_0 = 0{,}187$, when the input values obtained the following values of the coefficient: $K1_0 = 0{,}095$, $K1_0 = 0{,}121$, $K1_0 = 0{,}079$.

So, if you receive two of the coefficient $K1_0 = 0{,}295$, $K1_1 = 0{,}329$ the direction of impact will be selected at maximum $K1$[9,10].

4. Conclusions

Experimental studies of the work of a patient with sleep apnea treatment equipment show that with the signal-to-noise ratio = 0.8 and using the developed mathematical model and the signal processing algorithm the beginning of expiration is detected at count 31 on average, that is after 0.062 seconds (at sampling frequency of 500 Hz), and that of inspiration – at count 23 (0.046 seconds).

Application of this method reduces the recognition time by 2.5-3 times, making it possible to carry out preliminary parameter adjustment for each patient and to select optimal detection threshold values to be used when operating sleep apnea treatment equipment.

For a detailed description of the functionality of the respiratory system formation of diagnostic matrices at successive levels of loads. Developed methods for constructing matrices of state for various types of exposure apparatus, based on the allocation of the four main informative parameters in the analysis of the pressure characteristics is the angle of inclination of the pressure curve, the duration of the inspiratory phase/expiratory, the rise time of the pressure curve to the maximum angle during the descending pressure curve under load in the form of a resistance, the coefficient of the approximating function at the third site of observation under a load in the form of a pressure switch.

5. References

[1] *CPAP therapy machines* (Access mode: http://sleepmedicine.popmed.ru/cpap/cpap_ comparative_analysis) (31.08.2017)
[2] Ivakhno N V, Prohortsov A V, Senina E N and Fedorov S S 2014 Method of registration of chest movement in the diagnosis of sleep apnea *Bulletin of New Medical Technologies* **21(4)** 133-136
[3] Ivakhno N V and Merkulov O V Utility model patent No. 115668. Respiratory muscles simulator *Russian Federation. Priority date 20.09.2011*
[4] Ivakhno N V and Tyagin A D Invention patent No. 2375034. Artificial pulmonary ventilation machine *Russian Federation. Priority date 04.06.2008*
[5] Antsiperov V E 2016 Automatic target recognition for low-count terahertz images *Computer Optics* **40(5)** 746-751DOI: 10.18287/2412-6179-2016-40-5-746-751
[6] Ivakhno N V and Fedorov S S 2014 Principle of construction of mathematical model of process of signal processing in the detection of respiratory activity in systems of intellectual fitness effects *Biotechnosphere* **5(35)** 19-22
[7] Ivakhno N V 2015 The structure and algorithm circuit self-diagnosis of the intellectual equipment of respiratory muscles *Biotechnosphere* **3(39)** 40-44
[8] Gaidel A V, Khramov A G, Kapishnikov A V, Kolsanov A V and Pyshkina Yu S 2017 A method for digital renal scintigram analysis based on brightness and geometric features *Computer Optics* **41(1)** 103-109 DOI: 10.18287/2412-6179-2017-41-1-103-109
[9] Ivakhno N V and Antsibor S V 2015 Parametric analysis of respiration characteristics under the relay action *Proceedings of the Tula State University. Technical sciences* **5(2)** 78-84
[10] Ivakhno N V 2015 The method of diagnosis of the respiratory system under the choking action *Proceedings of the Tula State University. Technical sciences* **5(2)** 92-97
[11] Smelkina N A, Kosarev RN, Nikonorov A V, Bairikov I M, Ryabov K N, Avdeev A V and Kazanskiy N L 2017 Reconstruction of anatomical structures using statistical shape modeling *Computer Optics* **41(6)** 897-904 DOI: 10.18287/2412-6179-2017-41-6-897-904
[12] Gaidel A V 2016 Matched polynomial features for the analysis of grayscale biomedical images *Computer Optics* **40(2)** 232-239 DOI: 10.18287/2412-6179-2016-40-2-232-239
[13] Ivakhno N V 2015 Formation of a common set of informative signs to identify the condition of the respiratory system *Proceedings of the Tula State University. Technical sciences* **5(2)** 106-111
[14] Ivakhno N and Fedorov S 2016 Adaptive stress control method used in remedial action complexes for human respiratory system *Proceedings of the 12th Russian-German Conference on Biomedical Engineering* 243-246
[15] Gaidel A V 2015 A method for adjusting directed texture features in biomedical image analysis problems *Computer Optics* **39(2)** 287-293 DOI: 10.18287/0134-2452-2015-39-2-287-293

Asknowledgments

The results of the research project are published with the financial support of Tula State University within the framework of the scientific project № 8718.

Using component-wise functions in cryptographical transformation algorithm from Russian National Standard GOST R 34.12-2015

I I Vasilishin[1] and S Y Korabelshchikova[1]

[1]Northern (Arctic) Federal University named after M.V. Lomonosov, Severnaya Dvina Emb. 17, Arkhangelsk, Russia, 163007

Abstract. The paper presents the general approach to selecting functions, keeping the first argument field, in the process of symmetric encryption of a plaintext. Quantitative estimation and general characteristics of ordinate vector for such functions are given. Ten component-wise functions of binary logic algebra of three arguments, replacing one function of digit-wise addition modulo two in the cryptographictransformation algorithm from Russian National Standard GOST R 34.12-2015, are presented in the paper. Using component-wise functions widens the range of intermediate options of round transformations in block encryption, which complicates the decryption (cracking) algorithm for the cipher.

1. Introduction

The Russian National Standard GOST R 34.12-2015[1]is a symmetric cipher in which plaintext is converted by blocks of fixed length of 128 or 64 digits and a 256-digitlong key, and the encryption/decryption algorithms are reverse procedures using multi-round operations of substitutions and transformations the sequence of which isdeveloped by bitwise operation "addition modulo two" (AMT) for the source text and the first iterative key. Consequently, the multi-variant substitution of the AMT operation proposed by the authors leads to a greater variety of intermediate variants of the bit addition, which changes the results of the final operations, and in general complicates the algorithm of decryption (hacking) of the ciphertext.

2. General analysis of the existing algorithm

Cryptographic transformation of information used in GOST R 34.12-2015 is based on the principles of block data encryption [2, 3, 4] and contains a combination of operators that ensures the implementation of the properties of symmetric encryption in the standard [1]:

- *bitwise addition is formed by the AMT operation on the current transformation (a) and the round key (k), which corresponds to the bit transformation* $X[k]$: $V_{128} \rightarrow V_{128}$, where the result determines the equality of blocks before and after the AMT operation and is defined by the formula

$$X[k](a) = k \oplus a, \quad (1)$$

where: $k, a \in V_{128}$;

- *mixing information* is formed with a nonlinear bijective transformation over replacement block S performing both byte-substitution operation ($a_{15}, \ldots, a_0$) of128-bit value S: $V_{128} \rightarrow V_{128}$, where the value a_i

determines the index of the replacement array π, and the result $V_{128} \rightarrow V_{128}$ determines the equality of blocks before and after the replacement, and the formation of analytical complexity of dependencies between the key and the encrypted text, and is provided by the conversion

$$S(a) = S(a_{15}||...||a_0) = \pi(a_{15})||...||\pi(a_0), \quad (2)$$

where: $a = a_{15}||...||a_0 \in V_{128}$, $a_i \in V_8$, $i = 0, 1, ..., 15$;

- *dispersion of information* is achieved both by making nine-round consecutive embedding *F(a)* for sixteen times byte-wise conversion in each round *L(a)* over a 128-bit value of the replacement block *S(π(a))*, and by spreading of influence of each symbol in the plaintext to all the characters of the ciphertext, which is provided with transformations

$$F[k](a_1, a_0) = LSX[k](a_1) \oplus a_0, a_1, \quad (3)$$

where: $L(a) = R^{16}(a)$; $R(a) = R(a_{15}||...||a_0) = \ell(a_{15}, ..., a_0)||a_{15}||...||a_1$; $k, a_i \in V_{128}$.

Generalizing the transformations given in equations (1), (2) and (3), let us form a complete encoding algorithm $E_{Ki}(a)$ that performs transformations of 128-bit source block of information, where the direct numbering substitution of iterative keys is used

$$E_{K1, ..., K10}(a) = (k_1 \oplus a)F(a)(k_2 \oplus a)F(a)...(k_9 \oplus a)F(a)(k_{10} \oplus a). \quad (4)$$

It should be noted that the $D_{Ki}(a)$ full information decryption algorithm uses inverse transformations $S^{-1}(a)$, $R^{-1}(a)$ and $L^{-1}(a)$, and the numbering of iterative key substitution is done in the reverse order

$$D_{K10, ..., K1}(a) = (k_{10} \oplus a)F^{-1}(a)(k_9 \oplus a)F^{-1}(a)... (k_2 \oplus a)F^{-1}(a)(k_1 \oplus a). \quad (5)$$

In accordance with the transformations given in equations (1) – (5), the complete algorithmic sequence of information encryption/decryption contains both repetitive AMT operation over the iterative key (k_i) and the current transformation *(a)* and round wise mixing and dispersion of information for direct substitutions $S(a)$, $R(a)$, $L(a)$ and inverse substitutions $S^{-1}(a)$, $R^{-1}(a)$, $L^{-1}(a)$. Consequently, the entire periodic process of encryption / decryption of information should be divided, for the current presentation, into two groups of operations: AMT and substitutions.

3. Introduction of component-wise functions

The bitwise operation AMT used in cryptographic information conversion algorithms D(a) belongs to the functions of Boolean algebra [3], is generated by a combination of logical values of two variables and is oneof the operations that has the property of "restoring the input value of one of the variables" in sequential application of the operation in the process of encryption, and then in the process of decryption

$$D(a) = (k \oplus a) \oplus k, \quad (6)$$

where k is an iterative key and a is information for encryption.

Taking into account that the AMT operation belongs to the functions of Boolean algebra of two variables, we would like to point out that other functions generated by combinations of logical values of three [5], four or more variables possess the same property, but the generated functions do not perform the "classical" AMT operation. Therefore, let us present a new function for restoration of the input values, as an operation equivalent to the "classical" AMT operation and call it by a new termcomponent-wise function *M(a)*.

Let us present (see figure) the location of the component functions$M_j(a)$ in the process of encrypting 128-bit block of information in the form of an algorithm based on the transformation (4), for the standard transformation and the oneproposed by the authors.

As shown in the figure, the differences in the authors' encryption algorithm component-wise functions $M_j(a)$are used, replacing the only AMT operation used in the transformation (4), as well as the multiple use of iterative keys, while the production and the number of iterative keys remains the same as the standard and equals ten. It should be noted that the decryption process remains similar to the encryption process, except for changing the order numbers of the iterative keys to the opposite, similar to the transformation (5).

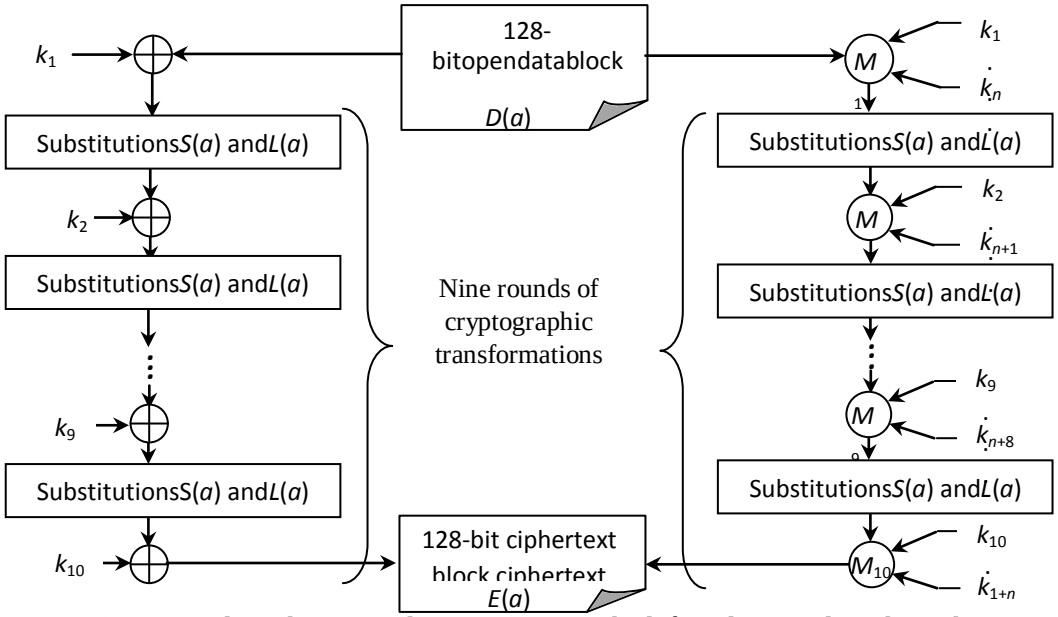


Figure 1. Algorithms encoding: existing on the left and original on the right.

4. The formation of the component-wise functions

The essence of the forming component functions consists in theoretical determination of the characteristics of functions that have the property of restoring the input value for one of the arguments when performing only the logical operation of the function over the operations of forward and inverse transformation, similar to the use of the AMT operation in equation (1). Further, the symbol "◈" defines a logical operation of the component function.

In the theory of abstract algebra, the proofs of existence of Boolean algebras for any number of variables are presented, indexing of Boolean functions is introduced, and belonging of the set of indexed functions of Boolean algebras to systems of normal forms is established [6]. Using the terminology of [5], we will present the principles of formation of the component functions in the form of statements, which are valid for any number of variables. For certainty we use component-wise functions of three variables.

Statement 1. On the distribution of meaning of variables of potentially suitable component-wise functions.

Direct conversion $M(a)[k]$ the argument field A contains the data to be converted, and the arguments B and C contain the iterative key values in direct numbering

$$M_j(a)[k]= M_j(A, B, C) = a◈k_i◈k_{i+1}≡X[k](a), \quad (7)$$

where j is the sequence number of component-wise functions; i= 1, 2, …, 10 isthe sequence number of the iteration key if $i> 10 \Rightarrow i := i \bmod 10$; k is the logical values of bits of the iteration key.

Inverse transformation $M^{-1}{}_j(a)[k]$:the argument A field contains the data to restore, and the arguments B and C contain the iterative key values in reverse numbering

$$M^{-1}{}_j(a)[k]= M^{-1}{}_j(A, B, C) = a◈k_{i+1}◈k_i≡X^{-1}[k](a), \quad (8)$$

where: i = 10, 9, …, 1 is the sequence number of the iteration key if $i< 1 \Rightarrow i := i \bmod 10$.

Let's apply the direct transformation specified in equation (7) in the process of encryption and the inverse equation (8) in the process of decryption

$$D(a) = M_j{}^{-1}M_j(a)[k], \quad (9)$$

consequently, such a double transformation results in restoration of the input data, similar to the transformation specified in equation (6).

Statement 2. On the distribution of the field of logical zeros and ones of the truth table of potentially suitable component functions.

The forward and reverse conversion functions, potentially suitable for use, contain an equal number of zeros and ones, as confirmed by the condition

$$M^{-1}{}_j M_j(a)[k] = M(M(A, B, C), B, C) = A, \quad (10)$$

which gives us:

$$\begin{cases} M(M(0,B,C),B,C) = 0 \\ M(M(1,B,C),B,C) = 1 \end{cases}.$$

When performing substitution for all combinations of arguments, we choose from potentially suitable functions, which restore the field of argument A. Let's not consider here a function identical to the first argument, and its negation as not depending significantly on other arguments and, therefore, unsuitable for encryption. In accordance with the statements above and using the theory of abstract algebra and combinatorial techniques, let us perform quantitative calculations for the functions of two, three, and four arguments on the condition of equation (10), the results of which are given in table 1.

Table 1. Quantitative characteristics of component-wise functions.

№	Caracteristics	Numberofvariables		
		2	3	4
1	TotalBooleanfunctions	16	256	65536
2	Potentiallysuitablefunctions	6	70	17920
3	All functions that restore the variable A	2	14	254
4	Usefulfunctions	2	10	218

As follows from condition (10), the functions, keeping the argument A field, contain equal amount of zeros and ones in the ordinate vector. We consider such functions potentially applicable in the encryption algorithm. The number of such functions of n variables is described with the formula:

$$C(2^n, 2^{n-1}), \quad (11)$$

where C is the number of combinations.

In particular, for n=4 we get C(16, 8)=17920

But not all the potentially suitable Boolean functions restore the argument A field. Another condition can be derived from equality (10): for inverse sets of argument values with A variable the function takes on inverse values. This condition can be easily confirmed with the following algorithm.

Algorithm 1.

Step 1. Let us divide the field of logical values of Boolean function F in half. Wegettwovectors: F_1andF_2.

Step 2. If $F_2 = \neg F_1$, then function F restores the first argument. Otherwise it doesn't.

Performing the final selection of component functions, the number of which is specified inTable 1, let us form their perfect and then the minimal disjunctive normal form (MDNF). The results of the MDNF function formation are given in Table 2.

It should be clarified that the symmetric arrangement of the same transition probabilities in Table 2, with a natural increase in weight coefficients of the field of zeros and ones, characterizes the completeness and correctness of the representations of component-wise functions.

Unfortunately, some of the presented functions are interdependent. It is obvious, that if function F keeps the argument A field, function ¬F obtains the same property. Such dependence between cryptographic functions is unacceptable, that is why only the 5 functions, listed above, are to be used in the encryption/decryption algorithm, while F or ¬F can be picked randomly.

There are other functions to restore arguments B and C, MDNF for which are not given in Table 2. Functions, keeping the argument B field, meet the following condition: for sets of argument values, inverse on variable B, i.e. for sets (*A,B,C*) and (*A,¬B,C*), they take on inverse values. This condition can be easily confirmed with the following algorithm.

Table 2. Component-wise functions used to transform data.

№ (*j*)	Component-wise functions		Transitionprobabilities, %	
	FunctionBooleanvalues	function MDNF	0→0 1→1	0→1 1→0
1.	00011110	¬ABC∨A¬B∨A¬C	75	25
2.	00101101	¬AB¬C∨A¬B∨AC	75	25
3.	01001011	AB∨A¬C∨¬A¬BC	75	25
4.	01101001	AB¬C∨A¬BC∨A¬B¬C∨ABC	50	50
5.	01111000	¬AB∨¬AC∨A¬B¬C	25	75
6.	10000111	AB∨AC∨¬A¬B¬C	75	25
7.	10010110	A¬BC∨AB¬C∨¬A¬B¬C∨¬ABC	50	50
8.	10110100	¬AB∨¬A¬C∨A¬BC	25	75
9.	11010010	¬A¬B∨¬AC∨AB¬C	25	75
10.	11100001	¬A¬B∨¬A¬C∨ABC	25	75

Algorithm 2.

Step 1. Let us divide the field of logical values of Boolean function F in 4 equal parts: F_1, F_2, F_3, and F_4.

Step 2. If $F_2 = \neg F_1$ and $F_4 = \neg F_3$, then function F restores the second argument. Otherwiseitdoesn't.

Similar algorithm can be used to find functions, restoring the third argument. The only difference is that the logical values field of Boolean function F is divided into 8 equal parts (for n=3 we get one component in each part), and set them equal pairwise. To define a component-wise function by an ordinate vector, we can randomly pick half of the values, and the other half is constructed inversely to the first one. So the number of Boolean functions of n variables that restore second, third, …, n^{th} argument is described with formula:

$$2^{2\wedge(n-1)}. \quad (12)$$

In the table 3 we present component-wise functions of three variables, restoring the second argument and significantly dependent on all the variables.

Table 3. Component-wise functions used to restore the second argument.

№ (*j*)	Component-wise functions		Transitionprobabilities, %	
	FunctionBooleanvalues	function MDNF	0→0 1→1	0→1 1→0
1.	00110110	¬AB∨A¬BC∨AB¬C	75	25
2.	00111001	¬AB∨A¬B¬C∨BC	75	25
3.	01100011	AB∨¬AB¬C∨¬A¬BC	75	25
4.	01101001	¬A¬BC∨¬AB¬C∨A¬B¬C∨ABC	50	50
5.	01101100	A¬B∨¬BC∨¬AB¬C	25	75
6.	10010011	AB∨BC∨¬A¬B¬C	75	25
7.	10010110	A¬BC∨AB¬C∨¬A¬B¬C∨¬ABC	50	50
8.	10011100	¬AB∨¬B¬C∨¬ABC	25	75
9.	11000110	¬A¬B∨¬B¬C∨ABC	25	75
10.	11001001	¬A¬B∨A¬B¬C∨ABC	25	75

In addition, there are other functions for restoring variables C that do not exist in Table 2 and 3.Show them in the Table 4.

Table 4. Component-wise functions used to restore the third argument.

1.	01010110	6.	10010101
2.	01011001	7.	10010110
3.	01100101	8.	10011010
4.	01101001	9.	10100110
5.	01101010	10.	10101001

We have eliminated 6 functions with fiction variables. When we calculated the total number of Boolean functions from the four variables used to transform the data, we excluded from the total number 38 functions with fiction variables.

Statement 3. On the use of component-wise functions in cryptographic transformation.

The existing algorithm of cryptographic transformation GOST R 34.12-2015 uses only one function of two variables, the AMT function, to be exact, (see the left algorithm in the Figure). According to Table 2, there are ten different functions of the three variables that satisfy the condition of restoring the input value of the field of argument A, therefore, in cryptographic information conversion can be used one of the functions listed in the table as well as any combination of these functions. When you use the required number of three-argument functions, the encryption algorithm takes the form shown on the right side of the Figure.

5. Analysis of transformations by component-wise functions

One way to demonstrate the practical capabilities of component functions is to perform transformations of functions from Table 2 and compare the results with the results given in GOST R 34.12-2015 [1].

To use component-wise functions, we extend the general appearance of the bitwise addition property, as shown in transformation (1), to apply the binary logic functions of three arguments [4]

$$X[k](a) = M_j\{(a)[k_1][k_2]\}, \quad (13)$$

where: M_j is the sequence number of the functions in Table 2; k_1 and k_2 are the iterative keys.

Extending properties of bitwise addition in the conversion (1) to the form of transformation (13) makes other combinations of operators, which gives the whole cryptographic transformation a new property, i.e. *component-wise transformation*.

We are going to present the results of a component-wise transformation for direct transformation $M_j(a)[k]$, corresponding to the application of equation (7), taking into account transformation (13), for functions the j = 1 and 2 from Table 2:

$$M_1(a)[k] = a \blacklozenge k_1 \blacklozenge k_2 = 99ba99dc51325510ffefddefbbabddef; \quad (13, a)$$

$$M_2(a)[k] = a \blacklozenge k_1 \blacklozenge k_2 = 676623226766670 0fecc988832221000, \quad (13, б)$$

where: a = 1122344556677 00ffeeddccbbaa9988 [1, p. 14]; k_1 = 8899aabbccddeeff0011223344556677 [1,p. 13]; k_2 = fedcba98765432100123456789abcdef [1, p. 13].

It should be noted that the result of the bitwise addition (1) given in [1, p. 14] has the form

$$X[k](a) = k_1 \oplus a = 99bb99ff99bb99ffffffffffffffff, \quad (14)$$

expectedly different from transforming equation (13), the results of which are given in transformation (13, a and b) proposed by the authors method of component conversion.

6. Conclusion

The three-argument binary logic algebra functions extending the cryptographic transformation modes of Russian National Standard GOST R 34.12-2015 are the beginning of a series of similar functions for four, five and more arguments, and their use will undoubtedly increase the possibilities of a bit-wise algorithm transformation. It is also possible to extend the approach to the considered question using the method of work [7], provided that the values of the arguments of the component-wise functions are read from the

file. Further research will be aimed at the development of a new type of functional transformation, followed by its implementation in the software and hardware complex, which will enable introduction of the next version of GOST R 34.12-2015.

7. References

[1] *GOSTR 34.12-2015 2015 Information Technology. Cryptosecurity. Block Ciphers* (Moscow: Standartinform) p 25
[2] Shannon C E 1963 *Works on Communication Theory and Cybernetics* (Moscow: Publishing house of foreign literature) p 830
[3] Shannon C E 1948 A Mathematical Theory of Communication *Bell System Technical Journal* **27** 379-423, 623-656
[4] Fomichev M I 2003 *Discrete Mathematics and Cryptology* (Moscow: Dialogue-MEPhI) p 400
[5] Sultanov D M, Vasilischin I I and Pugin M S 2016 Architecture Concept of the Encryption Unit, Performing the Russian National Standard GOST R 34.12-2015 Cryptographic Transformation Based on PAC. Expanding the Modes of the GOST R 34.12-2015 Cryptographic Transformation Based on a Specific Microcontroller *Proceedings of the international scientific conf. Parallel Computing Technology* 797
[6] Lidl R and Pilz G 1984 *Applied Abstract Algebra* (Springer-Verlag New York Inc.) p 743
[7] Yumaganov A S and Myasnikov V V 2017 A method of searching for similar code sequences in executable binary files using a featureless approach *Computer Optics* **41(5)** 756-764 DOI: 10.18287/2412-6179-2017-41-5-756-764

Automated sound generation based on image colour spectrum with using the recurrent neural network

N A Nikitin[1], V L Rozaliev[1], Yu A Orlova[1] and A V Alekseev[1]

[1]Volgograd State Technical University, Lenin avenue 28, Volgograd, Russia, 400005

Abstract. This work is devoted to development and approbation of the program for automated sound generation based on image colour spectrum with using the recurrent neural network. The work contains a description of the transition between colour and music characteristics, the rationale for choosing and the description of a recurrent neural network. The choices of the neural network implementation technology as well as the results of the experiment are described.

1. Introduction

Since the music began to be recorded on paper in the form of musical notation, the original "ways" of its composition began to appear. One of the first methods of algorithmic composition was the method of composing music invented by Mozart - "The Musical Game of the Dice". The first computer musical composition - "Illiac Suite for String Quartet" - was created in 1956 by the pioneers of using computers in music - Lejaren Hiller and Leonard Isaacson [1]. In this work, almost all the main methods of algorithmic musical composition are used: probability theory, Markov chains and generative grammars.

The development of computer music, including the sound generation by image, in the last century was severely limited by computing resources - only large universities and laboratories could afford to buy and hold powerful computers, and the first personal computers lacked computing power. However, in the 21st century, almost everyone can study computer music.

Now, computer music can be used in many industries: creating music for computer games, advertising and films. Now, to create background music compositions in computer games and advertising, companies hire professional composers or buy rights to already written musical works. However, in this genre the requirements for musical composition are low, which means that this process can be automated, which will allow companies to reduce the cost of composing songs. Also, the generation of sounds based on image can be applied in the educational process. The development of musical perception in preschool children can be in the form of integrated educational activities, which is based on combinations of similar elements in music and arts (the similarity of their mood, style and genre) [2].

The greatest success of the theory of automation of the process of writing and creating music made up relatively recently (at the end of XX century), but mostly associated with the study and repetition of different musical styles.

Since the process of creating music is difficult to formalize, artificial neural networks are best suited for automated sound generation – they allow identifying connections that people do not see [3]. In addition, to reduce the user role in the generation of music, it was decided to take some of the

musical characteristics from the image. Thus, the purpose of this work is to increase the harmony and melodicity of sound generation based on image colour spectrum through the use of neural networks.

To achieve this purpose the following tasks were identified:

- Determine the correlation scheme between colour and musical characteristics.
- Review the types of neural networks and choose the most suitable type for generating musical compositions.
- Describe the neural network used to generate music compositions by image.
- Choose neural network implementation technology.
- Choose a method for sounds synthesizing.
- Design and develop a program for sound generation using neural networks.
- Make an experiment to assess the harmony and melody of the output musical composition.

2. From colour to musical characteristics

To reduce the user role in the generation of music, some of the musical characteristics are obtained by analysing the colour scale of the image. Thus, the character of the output musical composition will correspond to the input image. This feature makes possible to use this approach for creating background music in computer games, advertising and films.

The key characteristics of a musical work are its tonality and tempo. These parameters are determined by analysing the colour scheme of the image. To begin with, we determine the ratio of colour and musical characteristics [4] (table 1).

Table 1. Correlation between color and musical characteristics.

Colour characteristics	Musical characteristics
Hue (red, blue, yellow...)	Pitch (c, c#, d, d#, …)
Colour group (warm/cold)	Musical mode (major/minor)
Brightness	Octave
Saturation	Duration

Then, it is necessary to define the correlation scheme between the color and pitch name. At the moment, there are a large number of such schemes, but in this work, was chosen the Newton scheme (table 2).

Table 2. Correlation scheme between color and pitch names.

Colour name	Pitch
Red	C
Red – orange	C#
Orange	D
Yellow – orange	D#
Yellow	E
Green	F
Green - blue	F#
Blue	G
Blue-violet	G#
Violet	A
Yellow-green	A#
Pink	H

As can be seen from Table 1, the tonality of a composition is determined by two colour characteristics - a hue and a colour group, the tempo by brightness and saturation. The algorithm for determining the tonality relies on image analysis and table 1. It consists of 3 steps and described below.

Step 1. Converting the input image from RGB to HSV colour space. This step allows transforming the image to a more convenient form, because HSV space already contains the necessary characteristics - the name of the colour (determined by the parameter hue), saturation and brightness (value parameter).

Step 2. Analysing the whole image and determining the predominant colour.

Step 3. Determining the name and colour group of predominant colour.

Step 4. According to table 1 and table 2 define the tone of the musical composition (pitch and the musical mood).

To determine the tempo of composition, it`s necessary to get the brightness and saturation of predominant color, and calculate the tempo, according to these parameters.

In a general form, the approach can be represented as follows (figure 1):

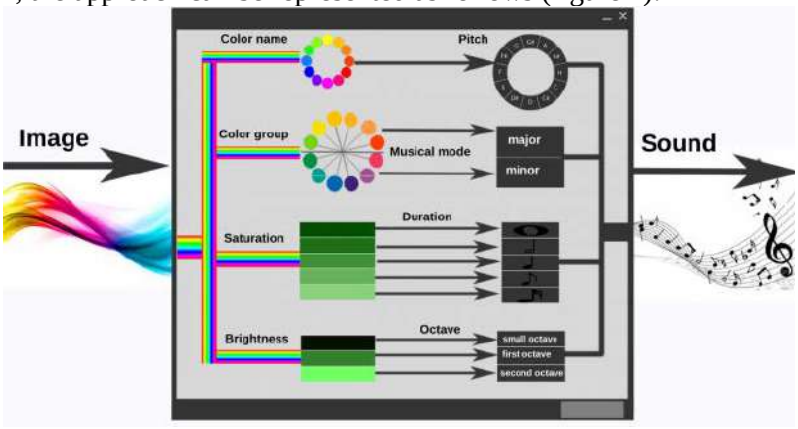


Figure 1. Approach used to generate sounds by image.

3. The choice of a neural network to generate musical compositions

An important feature of feedforward neural networks is that, this neural network has a common limitation: both input and output data have a fixed, pre-designated size, for example, a picture of 100 × 100 pixels or a sequence of 256 bits. A neural network, from a mathematical point of view, behaves like an ordinary function, albeit very complexly arranged: it has a pre-defined number of arguments, as well as a designated format in which it gives the answer.

The above features are not very difficult if we are talking about the pictures or pre-defined sequences of symbols. But for the processing of any conditionally infinite sequence in which it is important not only the content but also the order of information, for example, text or music, neural networks with feedback should be used - recurrent neural networks (RNN). In recurrent neural networks neurons exchange information among themselves: for example, in addition to a new piece of incoming data the neuron also receives some information about the previous state of the network. Thus, the network realizes a "memory" which fundamentally changes the nature of its work and allows to analyse any data sequences in which it is important the order of information [5].

However, the great complexity of RNN networks is the problem of explosive gradient, which consists in the rapid loss of information over time. Of course, this only affects the weights, not the states of the neurons, but it is in them information accumulates. Networks with long-short term memory (LSTM) try to solve this problem through the using filters and an explicitly specified memory cell. Each neuron has a memory cell and three filters: input, output and forgetting. The purpose of these filters is to protect information. The input filter determines how much information from the previous layer will be stored in the cell. The output filter determines how much information the following layers will receive. Such networks are able to learn how to create complex structures, for example, compose texts in the style of a certain author or compose simple music, but they consume a large amount of resources [6].

Thus, to implement the program for automated sound generation based on image colour spectrum, it is necessary to use recurrent neural networks with long short-term memory - RNN LSTM. This kind of neural networks is used to generate musical compositions in various programs such as Magenta. Magenta is an open source music project from Google. Also, RNN LSTM is used in BachBot. This is the program that creates the musical composition in the Bach style. And this kind of neural network is used in DeepJaz - the system that allows to generate jazz compositions based on analysis of midi files.

4. Description of the used neural network

Recurrent neural network (RNN) has recurrent connections which allow the network to store information related to the inputs. These relationships can be considered similar to memory. RNN is especially useful for the study of sequential data, such as music.

In TensorFlow, the repeated connections on the graph are deployed into an equivalent feedforward neural network. Then this network is trained using the technique of gradient descent, called backpropagation through time (BPT).

There are a number of ways in which RNN can connect to itself with cyclic compounds. The most common are networks with long-short term memory (LSTM) and gated recurrent units (GRU). In both cases, networks have multiplicative neurons that protect their internal memory from overwriting, allowing neural networks to process longer sequences. In this work, LSTM is used. All recurrent neural networks have the form of a chain of repeating modules. In standard RNNs, this repeating module will have a very simple structure, for example, one layer of tanh. LSTMs also have this chain, but the repeating module has a more complex structure. Instead of having one layer of the neural network, there are four interacting with each other in a special way [7].

The first step in LSTM is to decide what information we are going to throw out of the cell state. This decision is taken by the sigmoid layer. This layer looks at the value of h_t-1 output and x_t input, calculates a value in the range from 0 to 1 for each C_{t-1} state. If the layer returned 1, this means that this value should be left (remember), if 0 - removed from the state of the cell. For example, in the state of a cell, the characteristics of the current measure can be stored - if the measure is not yet complete, then it is necessary to leave the characteristics in memory, if work is already in progress, then new parameters must be memorized.

The next step is to decide what new information we are going to store in the state of the cell. To do this, firstly the sigmoid layer decides what values we will update. Next, the tanh layer creates a vector of new candidate values, C_t, that can be added to the state.

The next step is to update the old C_{t-1} state in the new C_t state. To do this, it is necessary to multiply the old f_t state, thus deleting the information from the state. Then, it's necessary to add resulting value and $i_t * C_t$. Thus, we get new candidate values, scaled by the update coefficient value of each state value.

At the last step, we need to decide what will output this layer. This output will be based on the state of the cell. First, we pass the input value through the sigmoid layer, which decides which parts of the cell state should be output. Then, we process the state of the cell using tanh (to shift the value between -1 and 1), and multiply it by the output of the sigmoid layer.

The behaviour of a neural network is determined by the set of weights and displacements that each node has. Therefore, for the properly work of neural network we need to configure them to some correct value. First, it is necessary to determine how good or bad any output is according to the input value. This value is called cost. Once the cost is received, we need to use the backpropagation method. In fact, it reduces to calculating the cost gradient relative to the weights (differential of the cost for each weight for each node in each layer), and then it is necessary to use the optimization method to adjust the weights to reduce the cost. In this work, we will use the method of gradient descent.

For the training of a neural network, it is proposed to feed a vector that contains the following parts [8]:

- Note name: MIDI representation of current note. Used to represent the pitch of a note.
- Time when note on.
- Time when note off.
- The velocity of the note playback.

To determine the correct output according to the input, it is suggested to transform the vector as follows: let there be a vector of notes {c, d, e, f, g, a, h}, then the learning vector will be {{c, d}, {d, e}, {e, f}, {f, g}, {g, a}, {a, h}}. This method of learning a neural network is used, for example, to predict time series [9].

5. The choice of technology for the implementation of an artificial neural network

To implement an artificial neural network, the Python programming language was chosen, because the language is cross-platform, it is aimed at improving developer productivity and code readability. In addition, this language is focused on data analysis, and therefore contains a large number of libraries for deep learning.

Theano is an extension of the Python language, which allows to efficiently calculating mathematical expressions containing multidimensional arrays. Since this library is low-level, the process of creating a model and determining its parameters requires writing voluminous and noisy code. However, the advantage of Theano is its flexibility, as well as the availability of the implementation and use of its own components [10].

TensorFlow is an open source library for numerical calculation using stream graphs. This library, as well as Theano, is a low-level library, which means that the development process is complex. However, due to the low level of development of neural networks, a more flexible model can be obtained. Also the advantage of this library is a large community and good documentation [11].

Lasagne is a lightweight wrapper for the Theano library. Programming with Lasagne is quite low-level - it is necessary to declare each level of the neural network by using modular building blocks over Theano. Lasagne acts as a compromise between the flexibility of Theano and the simplicity of Keras [12].

Keras is a high-level API for the development of neural networks written in Python and capable of running based on TensorFlow, CNTK or Theano. This library was developed with an emphasis on the possibility of rapid experimentation. The downside of this library is a small flexibility [13].

MXNet is an open source deep learning system used for training and deploying deep neural networks. Since MXNet is a high-level library, the development of neural networks using MXNet is simpler and faster than using Theano or TensorFlow, but it is inferior to the Keras library due to the large number of supported languages and large scaling possibilities, which makes the program code more cumbersome [14].

To compare the libraries, the following criteria were proposed: flexibility, scalability, parallel computing support, GPU computing support. All considered libraries were estimated according to the above criteria on a five-point scale, where 0 is the minimum value of the criterion, and 5 is the maximum. The results of comparing the libraries are presented in Table 3.

Table 3. Comparison of deep learning libraries.

Parameter	Theano	TensorFlow	Lasagne	Keras	MXNet
Flexibility	5	4	3.5	2	3
Scalability	4	5	4	5	5
Parallel computing support	5	4	5	5	5
GPU computing support	4	5	4	5	5

Thus, we can conclude, that to develop a recurrent neural network for the generation of musical compositions, we should use the Keras library, since this library allows to work based on Theano and TensorFlow, taking advantage of them, while the process of developing neural networks using this library is simple and fast, which allows to create prototypes for rapid experimentation.

6. Sound synthesis

In the process of studying the methods of sound synthesis, four most popular methods of synthesis were considered: additive synthesis, FM synthesis, phase modulation and sampling.

Additive synthesis is very difficult to implement, due to the need for separate control of the volume and height of each harmonic, which even a simple timbre consists of dozens [15].

FM - synthesis is well - suited for synthesizing the sound of percussion instruments, the synthesis of other musical instruments sounds too artificial. The main disadvantage of FM synthesis is the inability to fully simulate acoustic instruments [16].

Phase modulation gives a good enough sound, but is very limited, so it's rarely used in practice.

Sampling is used in most modern synthesizers, since it gives the most realistic sound and is fairly simple to implement [17].

Each of the methods has its advantages and disadvantages, but Sampling was chosen as the most suitable method for sound generating based on image colour spectrum. This method gives the most realistic sound of instruments, which is an important characteristic for the program, and this method is relatively easy to implement. The disadvantage of sampling is its limitation, but for the implementation of the program it is not essential, since for the program needs it is not required possibilities of changing the ready-made presets.

7. Description of program for sound generation

To confirm the effectiveness of the proposed algorithms for generating sounds based on image color spectrum, a Python program was developed. At the input the program receives an image that the user loads manually. This method of loading the image was chosen as the most simple, since at this stage it is only a task to prove the effectiveness of the proposed solutions. Subsequently, this program can be further developed, according to the conditions of applicability to specific areas.

After receiving the path to the image from the user, the program loads the image into memory using the capabilities of the OpenCV library. Converting an image to the HSV color space is also done using this library. Then, the image is analyzed, the tonality and tempo of the work are determined.

After determining the tonality and tempo, the program chooses the most appropriate model for the given situation, based on this model the neural network generates (predicts) a musical composition.

The architecture of the program is shown in Figure 2.

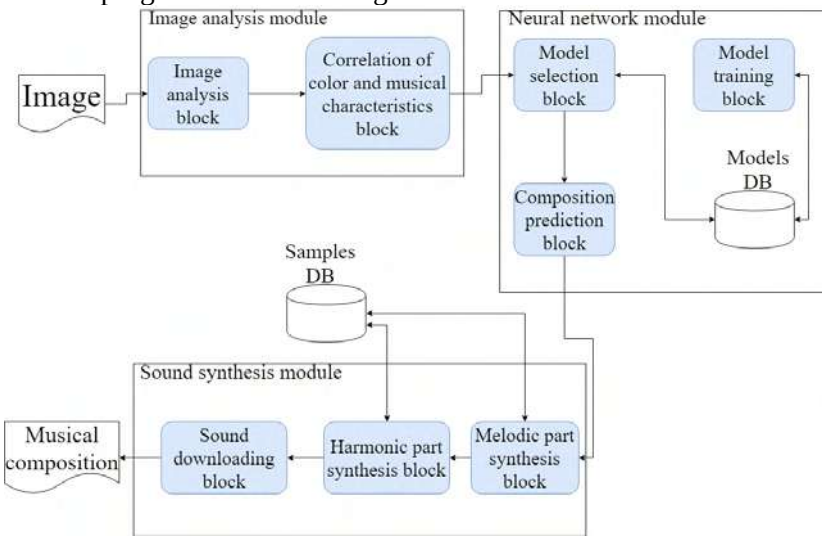


Figure 2. The architecture of the program for sound generation.

8. The experiment

To test the developed program, three experiments were conducted:

- Analysis the dependence of the quality of generated sounds on the size of the training set.
- Evaluation of the quality of generated sounds.
- Analysis the dependence of the quality of generated sounds with using neural networks and without neural networks.

Since there are no automated ways of assessing the quality of musical works, experts (people with a musical education) were used to evaluate the quality of compositions. For all experiments, the set of ten images were used. For all ten images, output musical compositions were prepared and stored (can be found here: https://github.com/NikitaNikitinVSTU/ImageSoundGeneration). Experts evaluated the compositions according to the following criteria:

- Matching character of image (on a five-point scale).
- Realistic sound of an instrument (piano or guitar).
- Melodiousness of the composition.
- The quality of harmony (accompaniment).
- The pleasantness of the melody for the perception.

- Integrity of the composition.
- Realism/artificiality of the composition.

An example of an abstract image is shown in Figure 3, an example of a landscape is shown in Figure 4 and an example of a city is shown in Figure 5.

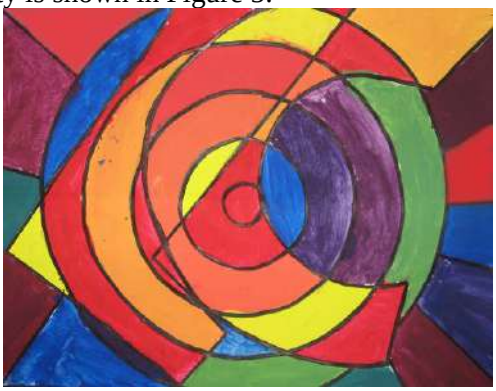


Figure 3. An example of an abstract image.

Figure 4. An example of a landscape.

Figure 5. An example of a city image.

8.1 The experiment of analysis the dependence of the quality of generated sounds on the size of the training set

For this experiment, a training sample consisting of 4295 pieces of music in .midi format was taken. Five models were trained at different sample sizes. For the first model, 120 classical works were taken, for the second model - 500, for the third - 1000 classical works. For the fourth and fifth models, 2000 and 4,295 (the maximum number of music pieces in the .midi format that are available for learning) of works of a different genre were taken. For all models, 3 songs were generated, and sent to experts for analysis. The results of the experiment are shown in Table 4.

Thus, we can conclude that the best model from the point of view of the quality of the generated musical composition is the model, trained for 2000 classical works.

Table 4. Average values for all tests for the first experiment.

Size of the training set	Average value for all tests
120	3.5
500	3.7
1000	4
2000	4.5
4925	4.1

It can be seen from the table that the last model, trained on 4295 tracks, generates sounds worse than the previous model. This is due to the fact that the compositions of a different genre were taken for the training of the last model that at the output gives a work consisting of a mixture of different genres - such works were judged by experts worse than just classical musical works.

8.2 The experiment of evaluating of the quality of generated sounds

For this experiment the set of ten images and output sounds were stored and send to ten experts.

Data from each expert has been processed and analyzed (it presented in table 5).

Table 5. Average values for all test for the second experiment.

Criterion	Average value for all tests
Matching character of image	4.9
Realistic sound of an instrument	3.9
Melodiousness of the composition	4.4
The quality of harmony	4.9
The pleasantness of the melody for the perception	4.6
Integrity of the composition	4.5
Realism/artificiality of the composition	4.3

Thus, after analyzing the assessments of all experts and calculating the average for each criterion [18], we can conclude that the piano is heard by experts to sound more realistic than the guitar. It can also be concluded that the composition generated by abstract images is more pleasant by ear than generation by landscape. In general, the overall impression of the generated sounds from experts is positive. Among the minuses, some experts emphasize the uniformity of harmony, sometimes the laceration and lack of realism of the composition, and not enough realism of the guitar.

Making a conclusion on each criterion, it can be said, that all experts rated matching character of image criteria to a high score, according to the second criterion - the piano instrument sounds quite realistic. The melodiousness of the compositions was divided in half, that is, half of the compositions were rated by experts for the top ball, the other half for 4, a generally good result. The quality of harmony was also evaluated by experts at the top ball. The pleasantness of the melody for the perception received 60% of the highest scores and 40% of the quads, which indicates that some compositions sound not quite realistic. The realism and integrity of the compositions is estimated at an average of 4, which is a natural result for computer sound generation.

8.3 The experiment of analysis the dependence of the quality of generated sounds with using neural networks and without neural networks

In this experiment, 5 musical compositions were created using artificial neural net-works, and 5 without their use. These compositions were sent to experts for analysis. The expert evaluated each pair of works of one of three evaluations:

- 1 - The sounds without the use of neural networks are better than with their use.
- 0 - The sounds without the use of neural networks are like sounds with their use.
- 1 - The sounds with using the neural networks are better than without using them.

The results of expert’s evaluation of pairs of products are presented in Table 6.

Thus, it can be seen from the table that the criteria "Matching character of image" and "Realistic sound of an instrument" did not change when the method of generating sounds was changed (using artificial neural networks and without). This is due to the fact that the parameter "Matching character

of image" is determined by the chosen tonality, which does not depend on the use of neural networks. And the parameter "Realistic sound of an instrument" depends on the synthesizer module of sounds, which was not affected either by changing the approach to generating tracks.

Table 6. Average values for all test for the third experiment.

Criterion	Average value for all tests
Matching character of image	0
Realistic sound of an instrument	0
Melodiousness of the composition	0.8
The quality of harmony	1
The pleasantness of the melody for the perception	1
Integrity of the composition	0.8
Realism/artificiality of the composition	0.8

The most important parameters for assessing the achievement of the research goal - "Melodiousness of the composition", "The quality of harmony" and "The pleasantness of the melody for the perception" were noted by experts as improved. This is due to the fact that without the use of neural networks a "naive" approach to the generation of compositions was used, while neural networks allow us to identify such connections that are not visible to a person, which affects the quality of the generated musical compositions.

The parameters "Integrity of the composition" and "Realism/artificiality of the composition" have improved slightly with the use of neural networks. This is due to the fact that these criteria depend on the previous three criteria, as they have improved, so the use of neural networks has positively affected these criteria.

9. Conclusion

In this work, the scheme of correlation of color and musical characteristics was determined, an overview of the types of neural networks was made and the most suitable type for the generation of musical compositions was selected. Also, the used neural network was described in detail, the technology of implementing the neural network was chosen and the method of synthesis of sounds was chosen. To evaluate the effectiveness of the proposed algorithms, an experiment was conducted to assess the harmony and melodiousness of the output musical compositions.

Analysis of various types and architectures of ANNs concluded that the most suitable network for processing musical information is the recurrent neural networks with long short-term memory (RNN LSTM).

During the description of the used neural network, it was determined that for learning the network it is supposed to input a vector that contains the following parts: MIDI representation of current note, time when note on, time when note off and the velocity of the note playback.

In analyzing the libraries for implementing a neural network in the Python programming language, it was discovered that the Keras library should be used to develop a recurrent neural network, since this library allows to work based on Theano and TensorFlow, taking advantage of them, while the development of neural networks using this library simple and fast, which allows to create prototypes for rapid experimentation.

As a result of the experiment, the model (neural network) on Beethoven's compositions was trained, and compositions of 10 images were generated. These compositions were sent for analysis to experts. As a result of the analysis of expert assessments, it can be concluded that the program generates quite melodic compositions, but it appears that the model was trained on a small number of compositions by only one author.

10. References

[1] Ariza C 2012 Two Pioneering Projects from the Early History of Computer-Aided Algorithmic Composition *Computer Music J*. **4** 40-56

[2] Vygotsky L 2004 Imagination and creativity in childhood *J. of Russian and East European Psychology* **1** 84-96 DOIP: 10.2753RPO1061-0405280184
[3] Mazurowski L 2012 Poland Computer models for algorithmic music composition *Proceedings of the Federated Conference on Computer Science and Information Systems* 733-737
[4] Caivano J L 1994 Colour and sound Physical and Psychophysical Relations *Colour Research and Application* **12(2)** 126-132
[5] Sak H, Senior A and Beaufays F 2014 Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition *ArXiv e-prints*
[6] Doornbusch P 2014 Gerhard Nierhaus Algorithmic Composition Paradigms of Automated Music *Generation Computer Music J.* **34**
[7] Mikolov T 2010 Recurrent neural network based language model *Proceedings of Interspeech International Speech Communication Association* **2010** 1045-1048
[8] Kim H K, Ao S I and Mahyar A 2013 Transactions on Engineering Technologies *Special Issue of the World Congress on Engineering and Computer Science* 796
[9] Fernandez J D and Vico F 2013 AI Methods in Algorithmic Composition A Comprehensive Survey *Journal of Artificial Intelligence Research* 513-582
[10] Bergstra J 2010 Theano a CPU and GPU math expression compiler *Proceedings of the Python for Scientific Computing Conference (SciPy)*
[11] *TensorFlow Large-Scale Machine Learning on Heterogeneous Distributed Systems* (Access mode: https://www.tensorflow.org/)
[12] *Lasagne lightweight library to build and train neural networks in Theano* (Access mode: http://lasagne.readthedocs.org/O)
[13] *Keras The Python Deep Learning library* (Access mode: https://keras.io/)
[14] *MXNet A Flexible and Efficient Library for Deep* (Access mode: https://http://mxnet.io/)
[15] Korvel G, Simonyte V and Slivinskas V A 2015 Modified Additive Synthesis Method Using Source-Filter Model *Journal of the Audio Engineers Society* **63** 443-450
[16] Lazzarini V and Timoney J 2012 Theory and Practice of Modified Frequency Modulation Synthesis *Sound and Music Technology Group* **58** 459-471
[17] Russ M 2012 *Sound Synthesis and Sampling* (London: Taylor & Francis Group) p 568
[18] Goshin Ye V and Kotov A P 2017 Parallel implementation of a multi-view image segmentation algorithm using the Hough transform *Computer Optics* **41(4)** 588-591 DOI: 10.18287/2412-6179-2017-41-4-588-591

Acknowledgments
The work is partially supported by the Russian Foundation for Basic Research (16-47-340320 and 17-07-01601 projects).

Analysis of dependence of the 2FSI subsystem parameters on the labyrinth seal diameter of the centrifugal compressor in the gas pumping unit

L N Butymova[1] and V Y Modorskii[1]

[1]Perm National Research Polytechnic University (PNRPU), Komsomolsky Av. 29, Perm, Russia, 614990

Abstract. The system of centrifugal compressor GTU has several 2FSI subsystems: impeller, labyrinth seals (LS), dummies, etc. The article analyzes dependence of one subsystem - the labyrinth seal - on its diameter. The effect of this parameter on the wave processes occurring in labyrinth seals during vibrations of the rotor is studied. A solution is obtained for a nonstationary aeroelastic three-dimensional gas flow problem in deformable LS. The amplitude-frequency characteristics of the gas dynamic and vibration parameters of the LS CC of GTU are obtained. Time dependences of the gas-dynamical force and the work of gas-dynamic forces in the LS zone at different rotor diameters are obtained.

1. Physical statement

The computational model of the labyrinth seal was considered, which is a set of the LS gas-dynamic gap and the LS steel ring in a 3D statement. The change in the gas dynamic force in time, calculated through the pressure field acting directly on the rotor [1-6], was monitored. The rotor motion of the was described by the harmonic law with given amplitude and frequency [7-11]. When assigning the rotor oscillation frequency, the blade frequency and the rotational frequency were taken into account. The rotor geometric characteristics were varied and their influence on the wave processes occurring in the labyrinth seals during rotor vibrations was estimated. At the LS inlet and outlet, a differential pressure was set. Particular attention was paid to evaluation of wave processes in the circumferential direction [12-18]. As a parameter of the investigated dependences, the rotor diameter in the LS zone is adopted.

2. The solid staging

The solid model (basic variant) is a circular gas-dynamic gap of LS with width 37 mm, height 500 μm and diameter 422 mm. The solid-state LS ring is a complex stepped construction (Figure 1). The construction material is steel, the pressure drop is not taken into account, the working medium is perfect gas.

The plan for carrying out the computational experiment on studying resonances in the 2FSI subsystem “labyrinth seals” (LS) of GTU is shown in Table 1. The parameters vary with respect to the “basic variant”.

Table 1. The plan for the computational experiment.

№ of variant	1 (BV[1])	2	3
D_{rotor}, mm	422	294	550

(BV[1]) - basic variant

The calculation lasts 8.5 hours when examining the perfect gas and absence of pressure drop. Preliminary calculations have shown that the calculation time increases and reaches 48 hours for variants corresponding to real gases and presence of pressure drop in the LS. The amount of computational resources required is limited by the ANSYS capabilities for parallelizing the durability part of 2FSI calculations. Acceleration of calculations of the structure stress-strain state in the 2FSI statement was observed when only up to 8 cores (1 node) were used. With a further increase in the number of calculation cores, the count rate did not increase.

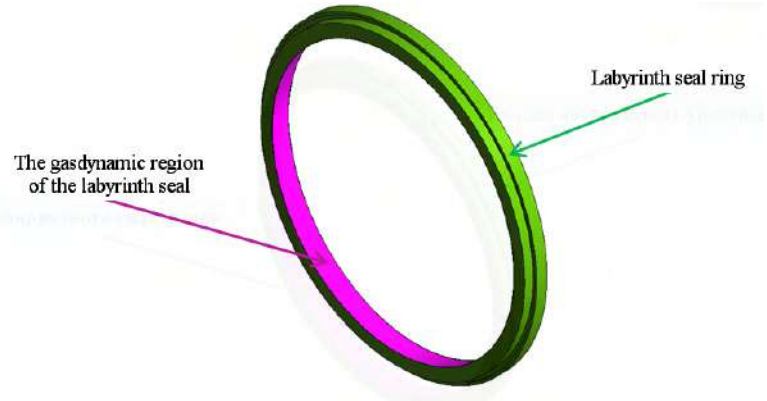
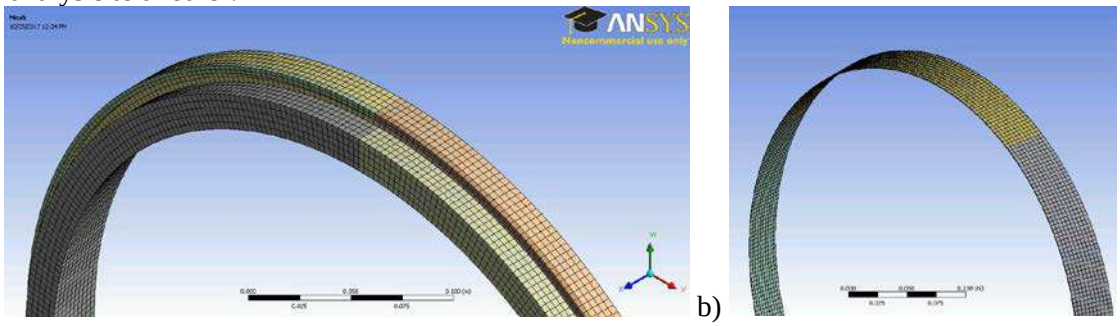


Figure 1. Solid-state model of the 2FSI subsystem of the LS CC GTU.

3. The mesh model

The minimum size of the finite elements of the mesh models of the LS ring and the gas-dynamic gap of the LS is 3.75 mm, the element type is hexahedral (Figure 2). For the feasibility of 2FSI calculations in the construction of mesh models of the gas dynamic region and durability region, it was ensured that the corresponding network elements coincide at the data transfer boundary from one type of analysis to another.



a) b)

Figure 2. The mesh model of the 2FSI subsystem of the LS CC GTU: (a) - mesh model of the labyrinth seal ring; (b) - mesh model of the gas-dynamic region of the labyrinth seal

4. Preparation and carrying out calculations on estimation of influence of LS geometrical characteristics on circumferential oscillations in 2FSI subsystem LS CC GTU

In accordance with the plan, the influence of the LS geometric characteristics on circumferential oscillations in the 2FSI subsystem of the LS GTU is investigated. Time dependences of pressure, gas dynamic force, work of gas dynamic force and total displacements in the LS ring LU on the rotor diameter are obtained (Figure 3).

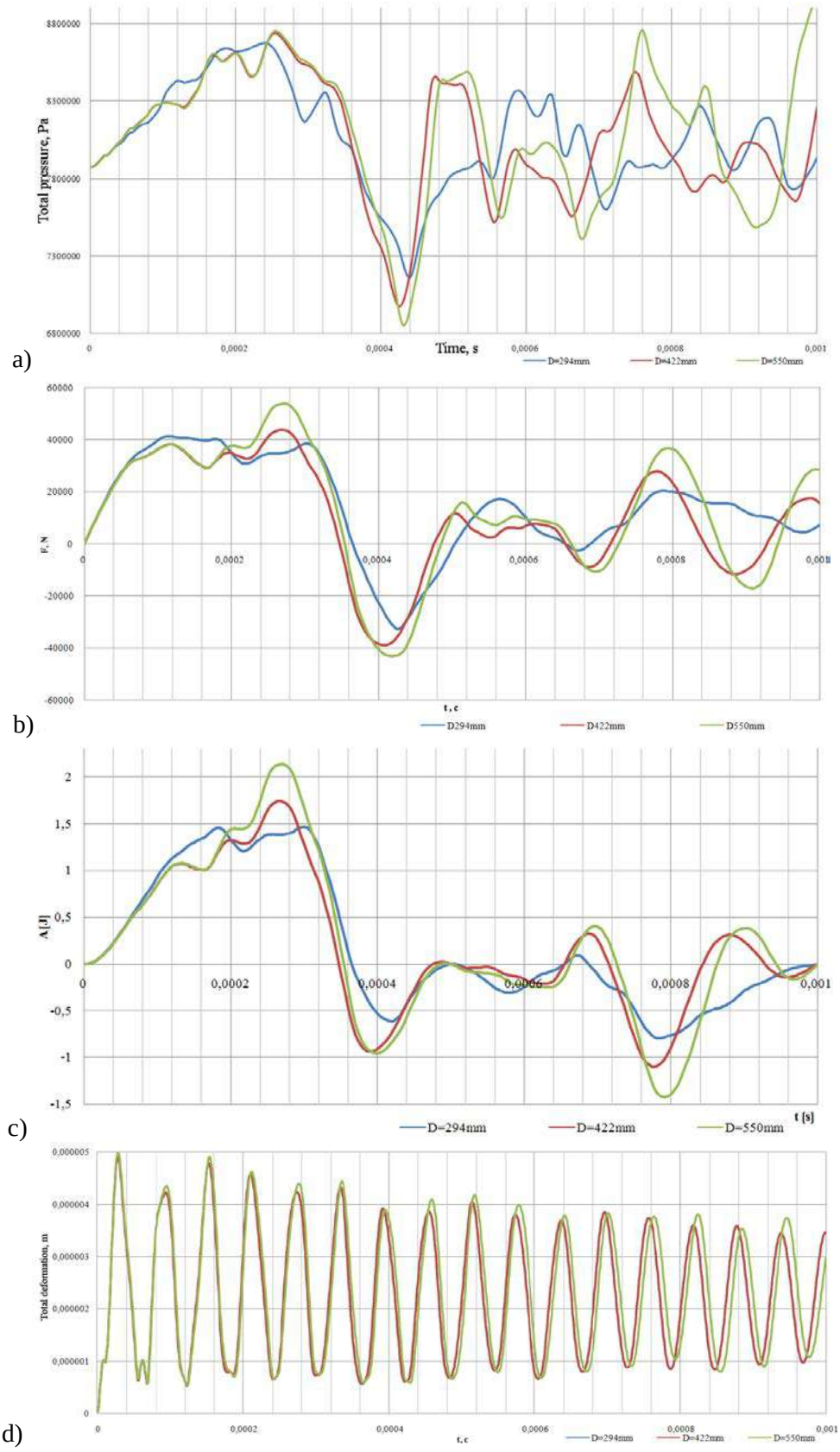


Figure 3. Time dependences of the 2FSI parameters of the LS CC GTU subsystem on the rotor diameter: (a) pressure in the LS gas-dynamic gap; (b) gas-dynamic force; (c) work of the gas dynamic force; (d) total displacements in the LS ring.

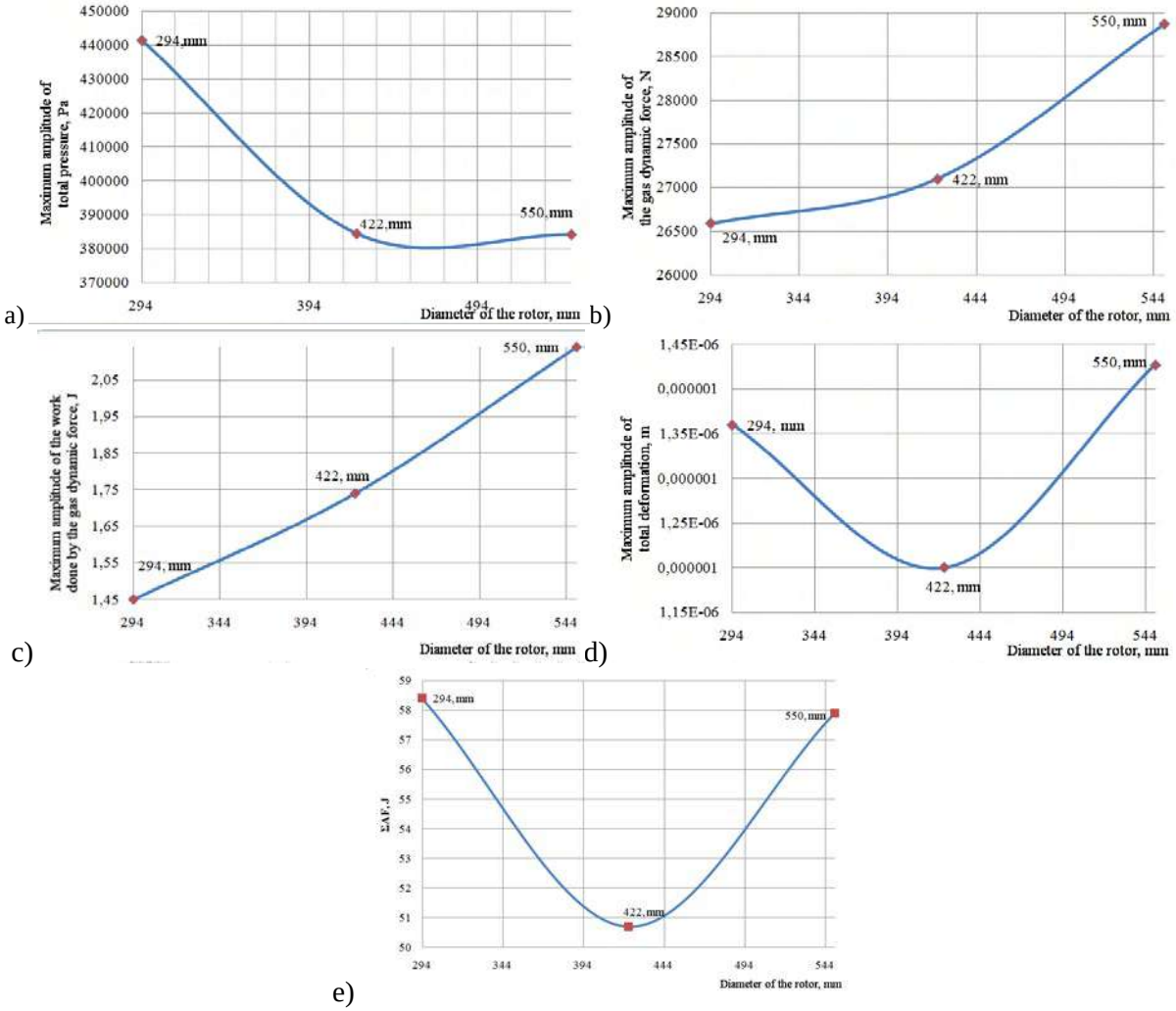


Figure 4. Analysis of the time dependences of the LS subsystem 2FSI parameters CC GTU on the rotor diameter: (a) - maximum amplitude of the braking pressure; (b) - maximum amplitude of the gas dynamic force; (c) - maximum amplitude of the gas dynamic force; (d) - maximum amplitude of the total displacements; (e) - total work of the gas dynamical force.

When analyzing time dependences of the 2FSI subsystem of the LS CC GTU on the LS geometric parameters (Figure 4), it was revealed:

- with increase in the shaft diameter from 294 mm to 550 mm, the maximum amplitude of the gas dynamic force is observed at the diameter of 550 mm and equal to 28.9 kN, the minimum value of the gas dynamic force amplitude is observed at the diameter of 294 mm and equal to 26.6 kN;
- the maximum amplitude frequency of the gas dynamic force is constant and equal to 1.5 kHz for the considered range of diameters and is not equal to the frequency of the driving force (0.95 kHz);
- with an increase in diameter from 294 mm to 550 mm, the maximum amplitude of the oscillation of the braking pressure does not change;
- with increase in diameter from 294 mm to 550 mm, the maximum amplitude of total displacements does not change;
- with increase in diameter from 294 mm to 550 mm (by 1.9 times), increase in the maximum amplitude of the work done by gas dynamic force from 1.45 J to 2.14 J (1.5 times) is observed;
- in figure 4a, it is seen that at the diameter of 294 mm, the pressure oscillation amplitude is maximal. Spectral analysis showed that the diameter equal to 294 mm corresponds to an oscillatory process with a resonant frequency of 1800 Hz, which is multiple of the rotor oscillation frequency (2 times). It should be clarified that the rotor oscillation frequency was set equal to the blade frequency.

The labyrinth seals with diameters equal to 422 mm and 550 mm correspond to oscillatory processes with resonant frequencies of 1260 Hz and 720 Hz, respectively, which are not multiple of the exciting frequency. In this case, the pressure oscillation amplitude at these diameters is reduced (Figure 4a.)

5. Conclusions

1. The influence of the DN diameter on the character of the oscillatory process in the LS is revealed.

2. If the frequencies of oscillations of the rotor and gas coincide (in the circumferential direction), the maximum amplitude of the braking pressure in the LS is observed.

3. The total work of gas-dynamic forces is positive in the considered range of diameters, which indicates the possibility of “build-up” of the system (Figure 4e).

4. To reduce the amplitude of the oscillatory process, it is necessary to "detune" from the coincidence of the oscillation frequencies of the rotor and the gas dynamic volume of the LS by changing the diameter of the LS at the design stage.

6. References

[1] Butymova L N, Modorskii V Ya and Petrov V Yu 2016 Numerical modeling of interaction in the dynamic system “gas-structure” with harmonic motion of the piston in the variable section pipe *ICMAR AIP Conference Proceedings* **1770** 030103

[2] Butymova L N and Modorskii V Ya 2016 One-way FSI simulation of the phase and the geometric parameters of the model of compressor blades on the oscillating gas-dynamic processes pipe *MATEC Web Conf.* **75** 04007

[3] Mekhonoshina E V and Modorskii V Ya 2016 Impact of magnetic suspension stiffness on aeroelastic compressor rotor vibrations of gas pumping units *ICMAR AIP Conference Proceedings* **1770** 030113

[4] Makarov A A and Zaitsev N N 2015 Engineering and theoretical problems of using labyrinth seals in high-speed rotary machines *Bulletin of PNRPU. Aerospace engineering* **42** 61-81 (in Russian)

[5] Butymova L N and Modorskii V Ya 2017 Development and application of a unified algorithm for solving the interdisciplinary problem of modeling aeroelastic processes in the labyrinth seal of centrifugal compressors *AIP Conference Proceedings* **1893** 030067

[6] Modorskii V Y and Shevelev N A 2016 Research of aerohydrodynamic and aeroelastic processes on PNRPU HPC system *ICMAR AIP Conference Proceedings* **1770** 020001

[7] Akzholov M Zh, Alabuzhev P M, Aliev A V, Artistsheva L M and Boyko L G 1995 *Numerical investigation of actual problems of mechanical engineering and mechanics of continuous and granular media by the method of large particles. Investigation of actual problems of mechanics and machine building* (Moscow: National Academy of Applied Sciences, International Association of Developers and Users of the Large-Particle Method) p 1658

[8] Mekhonoshina E V and Modorskii V Y 2015 On a phase-shift of waves at the middle interface *Computer Optics* **39** 385-391

[9] Butymova L N and Modorskii V Ya 2017 Numerical modeling of the labyrinth seal taking into account vibrations of the gas transmittal unit rotor in aeroelastic formulation *CEUR Workshop Proceedings* **1902** 10-17

[10] di Mare L, Imregun M, Green J S and Sayma A I 2010 A numerical study of labyrinth seal flutter *Journal of tribology-transactions of the ASME* **132(2)** 022201

[11] Zhang H, Jia XY, Pan XJ, Jiang B and Zheng Q 2016 Interaction between rotor and annular seals: interlaced and straight-through labyrinth seals *Journal of propulsion and power* **32(6)** 1483-1493

[12] Vannini G, Bertoneri M, Nielsen K K, Iudiciani P and Stronach R 2016 Experimental results and computational fluid dynamics simulations of labyrinth and pocket damper seals for wet gas compression *Journal of engineering for gas turbines and power-transactions of the ASME* **138(5)** 052501

[13] Dairien A, Thouverez F, Blanc L, Helies P and Dehouve J Fluid-structure interaction in a labyrinth gas seal coupled to a flexible stator *Topics in Modal Analysis & Testing* **10** 1-9
[14] Zhang E J, Jiao Y H, Chen Z B, Mo W C and Wang S Parametric influences on the nonlinear dynamic responses of a rotor-bearing-foundation-labyrinth seal system *Proceedings of the ASME international mechanical engineering congress and exposition* **4a** UNSP V04AT05A032
[15] A V Srinivasan, R A Arnoldi and A J Dennis 1984 Aeroelastic instabilities in labyrinth air seal systems *Mechanical Engineering* **106(6)** 91-96
[16] D R Abbott 1981 Advances in labyrinth seal aeroelastic instability prediction and prevention *Journal of Engineering for Power-Transactions of the ASME* **103(2)** 308-312
[17] Phibel R and di Mare L 2012 Comparison between a CFD code and a three control-volume model for labyrinth seal flutter predictions *Proceedings of the ASME Turbo Expo* **6** 597-606
[18] Phibel R, di Mare L, Green J S and Imregun M 2009 Numerical investigation of labyrinth seal aeroelastic stability *Proceedings of the ASME Turbo Expo* **6** 573-583

Acknowledgments
This study was supported by the Russian Science Foundation, project No 14-19-00877.

The Creation of Scalable Tools for Solving Big Data Analysis Problems Based on the MongoDB Database

O I Vasilchuk[1], A A Nechitaylo[2], D L Savenkov[3] and K S Vasilchuk[4]

[1]Volga Region State University of Service, Gagarin st. 4, Togliatti, Russia, 445677
[2]Samara National Research University, Moskovskoye shosse 34, Samara, Russia, 443086
[3]Samara State University of Economics, Sovetskoi Armii st. 141, Samara, Russia, 443090
[4]National Research University of Electronic Technology (MIET), Shokin Square 1, Zelenograd, Moscow, Russia, 124498

Abstract. This article presents analyze of using MongoDB database to storing and effective data mining from open network sources. This paper attempts to use NoSQL instead of traditional SQL database in systems with strongly related information, comparing to relational and non relational approach in the performance and architecture.

1. Introduction

Modern data storage technologies provided a practical opportunity to accumulate huge amounts of information, which allowed a qualitative change in the attitude to the results of analysis of stored information. It became possible to move from a descriptive process of analyzing the results obtained over a certain period of time to predictive data processing technologies that make it possible to offer valid recommendations for the future. The using of relational databases (MySQL, PostgreSQL, Oracle Database and others) to solve large data storage problems becomes problematic. The main advantage of relational databases is the availability of techniques for maintaining data integrity, achieved by storing links between data elements. However, the storage and validation of these links require additional time resources, which, with significant amounts and poor data structure, makes the use of relational databases difficult in some real-time systems.

As an alternative, there is a NoSQL database. One of the advantages of these databases in alternative storage formats and the links between them.

2. Alternative Data Storage for Unified Text Formats

2.1. The Problem Formulation

With the development of metaprogramming, the concept of reflection developed – the ability of the program to use and modify its structure[1]. Reflection, in the context of object-oriented programming, spread the technique where objects created by programs, based on knowledge of the structure of a class, are serialized into representations of given formats[2]. Most often this technique is used in the context of web programming when data is serialized to xml, json and other formats for text data transmission.

This led to the task of filtering such data across different data fields of text formats, including using standard filters (for example, XPath[3]).

A common problem is the problem of data storage, the final form of which is some unified format (hereinafter referred to as UF)[4]. Since a web application user (usually a client application) works only with UF, and internal data views are not available for it, the filters available to the client are reduced to UF fields.

In classical relational databases, storing such formats is associated with the creation of several linked tables and subsequent cross queries [5]. To speed up such queries, indexes are created for the corresponding keys.

Consider an example of a web service that provides access to the UF type JSON for some class "book":

Consider definition of scheme for UF:

```
{
  "definitions": {},
  "$schema": "http://json-schema.org/example/schema#",
  "$id": "http://itnt18.ru/itnt18.json",
  "type": "object",
  "properties": {
    "title": {
      "$id": "/properties/title",
      "type": "string",
      "title": "The Title Schema."},
    "description": {
      "$id": "/properties/description",
      "type": "string",
      "title": "The Description Schema."},
    "comments": {
      "$id": "/properties/comments",
      "type": "array",
      "items": {
        "$id": "/properties/comments/items",
        "type": "object",
        "properties": {
          "user": {
            "$id": "/properties/comments/items/properties/user",
            "type": "string",
            "title": "The User Schema." },
          "comment": {
            "$id": "/properties/comments/items/properties/comment",
            "type": "string",
            "title": "The Comment Schema." }}}}}
}
```

Listing 1: UF definition in JSON format.

We assume that most often the book is filtered by title. And we assume that sometimes we just want to find a book, for example, for a search function on a web site, and sometimes

download a page with a book where besides the book and its description we also want to display all the comments.

It is common practice in the relational database to compile the following schema 1.

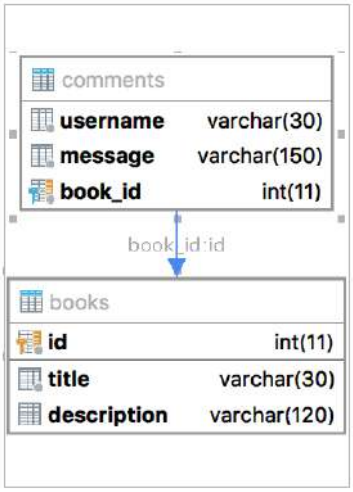


Figure 1. SQL data relations scheme.

MongoDB uses JSON-like format to keep data and can keep data in the output view.

2.2. Object-Relational mapping advantages

ORM or Object-Relational mapping is a programming technique designed to map database relations with object-oriented programming languages entities. It creates virtual objects database inside specific language representation.

The main the goal of technique is get rid of the need to write SQL queries to access database data. The dual way of data representation, relation and object-oriented, usually requires from programmers to write code for getting data from database in relation way, after transforming it to object-oriented, and transform back to relation data to safe changes. Relational databases operate over sets of tables with simple data representations, it leads to use SQL "JOIN" operation to get full object information. Since relational database management systems usually do not implement a relational representation of the physical link layer, the execution of several consecutive queries (referring to one "object-oriented" data structure) can be too expensive.

Relational database management systems work with good performance with global queries, affecting a large area of memory, but object-oriented access is more effective in the work with small amount of data, as it reduces semantic gap between the object and relational data representation[6].

Two way of data representation increases the complexity of object-oriented code to work with relational databases, it becomes more prone to errors.

2.3. Object-Relational mapping disadvantages

The most common problem with ORM as an abstraction over SQL is that it can't fully abstract realization details. Some of the program realizations of ORM works as SQL code generations tools, some of them do not use SQL equivalents at the external level.

The reason why abstraction make sense is simplification of code writing, but if you use ORM framework with knowing SQL as a requirement, it doubles programmers effort, for example the popular ORM framework Hibernate use HQL language SQL for complex requests, which is very

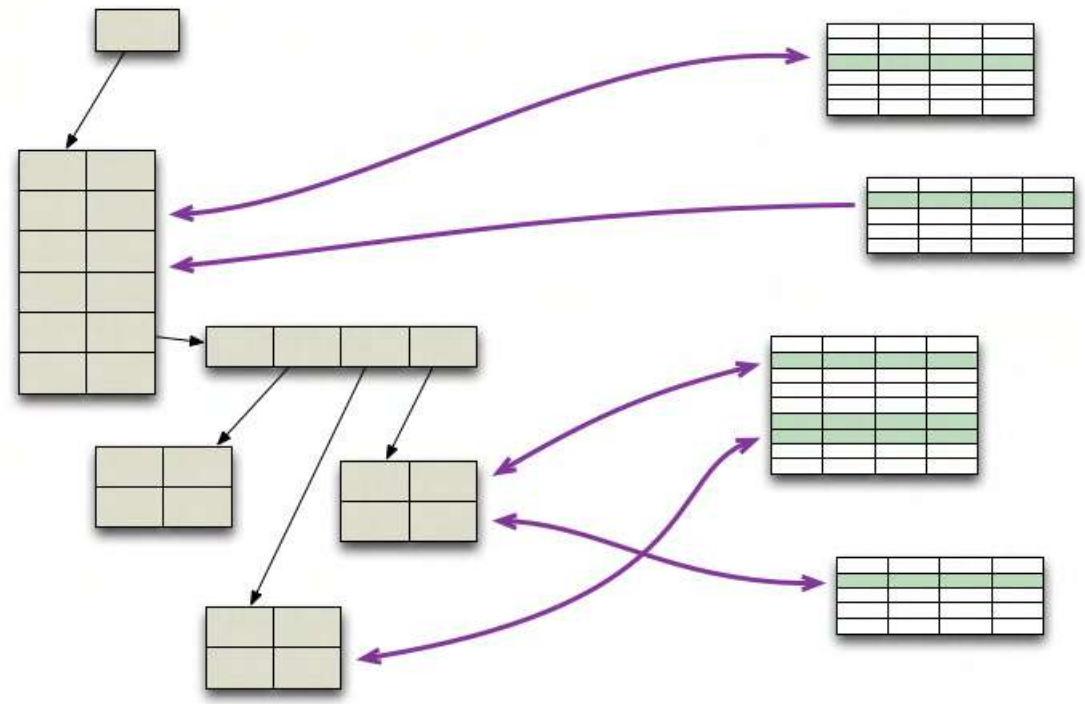


Figure 2. ORM data mapping.

semantically close to SQL. It brakes uniformity of the code abstraction when programmer needs a specific union of data processing.

Inefficiency is another common problem of ORM. If programmer need extract object data from relation database, ORM cannot know which of the object property are going to be used or changed, so it forced to extract all, it cause many requests instead few. The lack of context sensitivity means that ORM can't consolidate requests, which leads to the impossibility of data caching or other compensation mechanisms.

2.4. Object-Document mapping

ODM or Object-Document mapping is an alternative for ORM in document-oriented databases. The basic idea of ODM frameworks is the same, match data to the object, but we have few differences here.

Firstly, in ORM we should complete data for the object, and complete data for backward mapping inside database. In ODM there is no requirement for fully completeness of data. The document can be mapped partly to a database, without multiple table changing.

Secondly, ODM can make data mapping independently from the data source. It makes easier to operate over data in a program.

2.5. NoSQL databases

Relation databases usually based on ACID – Atomicity, Consistency, Isolation, Durability[7]. ACID is common requirement for transaction systems.

NoSQL databases usually based on BASE:

- basic availability – every request will be completed (successfully or not)
- soft state – The system state can be changed without any data changes due to data consistency
- eventual consistency – The data can be inconsistent for some time but will be consistent after a while.

It's obviously that NoSQL databases can not be used in any application. Some application are requires for transaction systems (like banking, ecommerce, etc), but at the same time usual ACID system doesn't suit for systems based on large data storages, like amazon.com and other. So NoSQL databases sacrifice data consistent to make more scalable system, to operate over large amount of data.

Also, NoSQL databases represent following features:

- Application of various types of storage facilities.
- Ability to develop a database without specifying schemas.
- Linear scalability (adding CPUs increases performance).
- Innovation: A lot of opportunities for data storage and processing

2.6. NoSQL databases common types

Unlike relation databases, NoSQL databases have various data schemas, implemented through the use different data structures.

Depending on the data schema and the approaches to distribution and replication, four types of storage can be distinguished: key-value store, document store, column database stores, graph databases.

Key-value storage

The key-value store is the simplest data store that uses the key to access the value. Such repositories are used to store media images, create specific file systems, as caches for objects, as systems well scalable by design. Examples of such storage facilities are Berkeley DB, MemcacheDB, Redis, Riak, Amazon DynamoDB [8].

Bigtable-like databases (column database stores)

In this store, data is stored as a sparse matrix, the rows and columns of which are used as keys. A typical application of this type of database is web indexing, as well as tasks related to large data, with reduced requirements for data consistency. Examples of databases of this type are: Apache HBase, Apache Cassandra, Apache Accumulo, Hypertable, SimpleDB.

Column family stores and document-based repositories have similar usage scenarios: content management systems, blogs, event logging. The use of timestamp allows using this type of storage for the organization of counters, as well as the registration and processing of various data related to time.

The family column stores should not be confused with column stores. The latter are relational databases with separate storage of columns (in contrast to the more traditional line-by-line data storage)[9].

Document-based database management system

Document-oriented databases serve to store hierarchical data structures. They are used in content management systems, publishing, document search, and so on. Examples of this type of database are CouchDB, Couchbase, MarkLogic, MongoDB, eXist, Berkeley DB XML.

Databases based on graphs Graph databases are used for tasks in which data has a large number of links, for example, social networks, fraud detection. Examples: Neo4j, OrientDB, AllegroGraph, Blazegraph, InfiniteGraph, FlockDB, Titan [10].

Since the edges of a graph are materialized, that is, they are stored, traversing the graph does not require additional computation (like JOIN in SQL), but to find the initial vertex of the traversal requires the presence of indices. Graphical databases generally support ACID, and also have different query languages, like Gremlin and Cypher (Neo4j).

2.7. Motivation for MongoDB

As an alternative to the classical approach, there are NoSQL databases. For tasks in which UF is involved, the most common use of document-based one.

It has following advantages:

- It works with unstructured data, which make possible to add new data field with no additional cost.
- It available to compromise finding in performance-reliability.
- Working with ODM frameworks (Object-Document mapping) – alternative for ORM [11]. In cases of optional data it makes possible to map object without cross queries, make the mapping operation faster.
- The JavaScript support on server side.

As a working example, the authors used MongoDB. This choice is due to prevalence and testing in high load projects[12].

Based on benchmarking top NoSQL databases performance tests conducted by "End Point Corporation", the authors systematized MongoDB performance indicators for various hardware configurations, which were summarized in the table 1[13].

Table 1. Query time analysis results (operations per second).

Node numbers	Reading	Reading/Writing	Reading/Writing/Changing
1	2 149	1 278	1 261
2	2 588	1 441	1 480
4	2 752	1 801	1 754
8	2 165	2 195	2 028
16	7 782	1 230	1 114
32	6 983	2 335	2 263

The performance comparison experiment was conducted in the cloud services of "Amazon Web Services EC2", which provides an industrial platform for systems that require a horizontal extension of the architecture, such as distributed non-relational databases. In order to minimize the errors in measurements related to the current load of "Amazon Web Services EC2" services, each set of test scenarios was played three times, at least 24 hours apart, using newly created clusters with hardware configurations described in the table 2 of hardware configurations of cloud services "Amazon Web Services EC2":

Table 2. Hardware configuration of "Amazon Web Services EC2".

Node Class	Configuration	Application
i2.xlarge	30.5 GB RAM, 4 CPU, one SSD with 800GB	database nodes
c3.xlarge	7.5GB RAM, 4 CPU	database client nodes

As an operating system in the nodes used Ubuntu 14.04 LTS AMI in the HVM mode (virtual hardware virtual machine) virtualization, customized with Oracle Java 7. For each test, an empty database was used as the starting point. The client applications were programmed to enter randomly generated information into the database. After the database was finally populated, each of the test scenarios was executed sequentially. All clients performed requests in parallel, and then waited for all operations to be performed and the corresponding results obtained. The client software was supplemented by the installation of the YCSB free software package designed to analyze the performance of NoSQL databases[14]. The study of the company "End Point Corporation" allows you to determine the number of necessary nodes for placing MongoDB in the solution of certain business tasks, based on the anticipated load, when using a SSD drive in each machine-node. Nevertheless, despite its obvious competitive advantages, such as small size and weight, as well as the number of random IOPS that exceed by an order of magnitude the more common HDDs, the SSD is inferior to the latter by cost[15].

2.8. Preconditions

The SQL database with 5 millions entries for book instance and 10 millions for comments (two for each book entry) was used. And 5 millions entries for full (with comments inside view, also two comments on each book) book instance was used in MongoDB collection.

Three main scenarios were examined:

- Find by title
- Find by comment
- Find comments for the book

Since MongoDB stores the data immediately in the final UF view, it will be absolutely identical for searching the book and requesting an extended output.

Averaged values over 10 experiments are used to except the non-deterministic influence of external factors.

2.9. Results

Table 3. MongoDB and MySQL comparison (in seconds).

DB	**MongoDB 3.4**	**MySql 5.4**
Find by title	0.434527	1.646251
Find by comment	5.826707	1.387221
Find comments for the book	0.434527	4.360173

As can be seen from the results(Figure 3 and Table 3) of the research, the search query in the collection of the MongoBD is more effective than the search in the MySQL. The most effective scenario is to search for comments on the book, that is, additional information related to the main entity. The speed of work of MongoDB in this case surpasses MySQL almost 5 times. Nevertheless, it is important to understand that if the search target is additional information(comments here), the velocity of the query for MongoDB will be greatly worse then traditional relational databases.

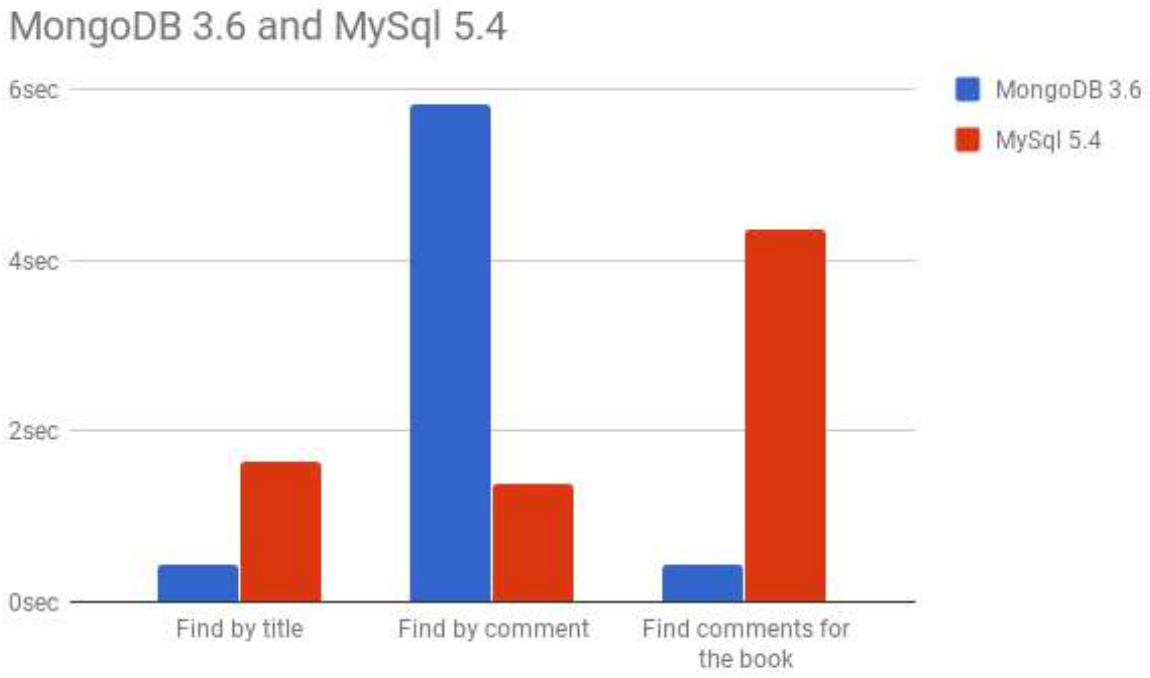


Figure 3. MongoDB and MySQL results comparison.

3. Conclusion

The conducted research allowed to draw a conclusion about the expediency of using document-oriented databases for storing large amounts of data for indexing purposes with a small number of supported links or their absence. Feasibility is confirmed by the fact that the use of document-oriented databases for storing large amounts of data for indexing with a small number of supported links or their absence allows you to select the optimal configurations of computing systems in the framework of current and future business tasks, with the possibility of horizontal and vertical scaling, in conditions better performance than relational equivalents.

4. Reference

[1] Kiczales G J d R, Bobrow D G 1991 *The Art of the Metaobject Protocol* (MIT Press)
[2] *JavaScript Reflect Global Object* (Access mode: https://developer.mozilla.org/en-US/docs/Web/JavaScript/Reference/GlobalObjects/Reflect) (18.10.2017)
[3] *XPath 3.1 Specification* (Access mode: https://www.w3.org/TR/xpath-31/) (18.10.2017)
[4] Bell C 2012 *Expert MySQL* (APress)
[5] Protsenko V I, Kazanskiy N L and Serafimovich P G 2015 *Computer Optics* **39(4)** 582-591 DOI: 10.18287/0134-2452-2015-39-4-582-591
[6] Loureno J R, Cabral B and Carreiro P 2015 Choosing the right nosql database for the job: a quality attribute evaluation *Journal Of Big Data* **2** 18
[7] Lake P and Crowther P 2013 Nosql databases Concise Guide to Databases *Undergraduate Topics in Computer Science* DOI: 10.1007/97814471560175
[8] Kazanskiy N L, Protsenko V I and Serafimovich P G 2014 *Computer Optics* **38(4)** 804-810
[9] Singh M and Kaur K 2015 Sql2neo: Moving health-care data from relational to graph databases *IEEE International Advance Computing Conference* 7154801

[10] Kazanskiy N L, Protsenko V I and Serafimovich P G 2017 *Procedia Engineering* **201** 817 DOI: 10.1016/j.proeng.2017.09.602
[11] Richardson L and Ruby S 2007 RESTful Web Services (Beijing: O'Reilly)
[12] *MongoDB Official site* (Access mode: https://www.mongodb.org) (18.10.2017)
[13] *Benchmarking NoSQL Top Databases* (Access mode: https://www.datastax.com/wp-content/themes/datastax-2014-08/files/NoSQL Benchmarks EndPoint.pdf) (18.10.2017)
[14] *Yahoo! Cloud System Benchmark* (Access mode: https://github.com/joshwilliams/YCSB) (18.10.2017)
[15] *Amazon EC2 Instance Types* (Access mode: https://aws.amazon.com/ru/ec2/instance-types/) (18.10.2017)

Method of application of the genetic algorithm for automatic generation of test data

K E Serdyukov[1] and T V Avdeenko[1]

[1]Novosibirsk State Technical University, K. Marks avenue 20, Novosibirsk, 630073

Abstract. Software testing has always been a time-consuming process, without obvious results but not less important than any other stage of software engineering. One of the major issues of testing is the development of initial test data that allow complete evaluate the program code and find most important parts of the program. Intelligent technologies, in particular, genetic algorithm allow to define initial data sets allowing to evaluate the values of variables for qualitative testing. This article proposes the method of possible integration of genetic algorithm with the system for generating test data.

1. Introduction

One of the most important stages in the software development is testing. The correspondence of the developed program to the specified requirements, observance of logic in the processing of data and obtaining the correct final results - all this is the testing purposes.

The processes of verification and validation of the program code are improving quite slowly. Development of most types of testing designs and templates is often done manually, without using any automation systems. Consequently, the testing process becomes incredibly complex and consuming, both in time and finance, if take it seriously. To test some programs, it may need up to 50% of the whole time.

One of the main purposes of testing is to create a test array that would ensure sufficient level of quality of the final product, by checking most of the various branches of the code. Nevertheless, the task of finding many branches of code consists of several more local tasks, the solution of which is necessary for the best possible finding of anappropriate testdata. One of the important tasks needed to be solved is the searching of most complex branch of the code. Subsequently, the task is extrapolated to find several important branches. In present paper, we solve the local problem of finding one optimal branch of the program.

Based on the foregoing, we can conclude that automation of testing, or at least the automation of test development, can significantly reduce not only the time costs, but also money. There are other advantages, not so obvious - a greater probability of finding small errors, transparency of test development, testing simultaneously with the development of the program, etc. [1]

Obviously, given the advantages of the process of the automation of testing, the researchers proposed different approaches to this process. For example, in the work [2] it was proposed an approach to automate the programming process of complex production systems in accordance with the standard IEC 61131.

Testing is not a standardized process, it depends on many factors, most of which vary from one program to another. Therefore, scientists came to a conclusion about usage of some research in the field of artificial intelligence, which would allow the development of a hybrid system for automatic finding of tests array [3,4].

The paper is organized as follows. In Section 2, the problem definition is given, and the process of distributing the weights along the code branches is described. Section 3 is devoted to the description of the genetic algorithm, its terminology and method of finding a solution. In Section 4, a detailed description of the operation of the algorithm and the results of the work are presented. Section 5 proposes an additional improvement of the algorithm and shows the results of the specially developed program. In section 6 we give conclusions on the investigation of the algorithm.

2. Genetic algorithms

Formally, the genetic algorithm is not an optimization method, at least in the terms of classical optimization. Its goal is not to find the optimal and best solution, but rather close to it. Therefore, this algorithm is not recommended to use if quick and well-developed optimization methods already exist. But at the same time, the genetic algorithm perfectly shows itself in solving non-standardized tasks, tasks with incomplete data, or those for which it is impossible to use optimization methods because of the complexity of implementation or the duration of implementation [6, 7].

The genetic algorithm is considered complete if a certain number of iterations has been passed (it is desirable to limit the number of iterations, since the genetic algorithm works on the trial-and-error basis, which is a fairly long process), or if a satisfactory fitness function has been obtained. The concept of the fitness function is one of the most important in the whole method. As a rule, the genetic algorithm solves the problems of maximization or minimization and the adequacy of each solution (chromosome) is estimated using the fitness function.

The genetic algorithm works according to the following principles:

- ***Initialization***. The fitness function is defined. An initial population is formed. In the classical theory, the initial population is formed by random filling of each gene in chromosomes. But to increase the rate of convergence of the solution, the initial population can be specified in a certain way, or random values are analyzed in advance to exclude definitely not suitable genes.
- ***Evaluation of population*** [8]. Each of the chromosomes is evaluated by the fitness function. Based on the set requirements, the chromosomes get an accurate value of how well they correspond to the problem being solved.
- ***Selection***. After each chromosome has its own fitness value, the best chromosomes are chosen. Selection can be made by different methods, for example, the first *n* chromosomes from the sorted order, or only the most suitable ones, but not less than *n*.
- ***Crossing*** [9]. The first is significantly different from conventional methods. After selection of the chromosomes that are suitable for solving the problem, they are crossed among themselves. Random chromosomes from all the "chosen" ones in a random order form new chromosome. Crossing occurs on the basis of the choice of a certain position in two chromosomes and replacement of parts of each other. After the required number of chromosomes is filled to create the population, the algorithm proceeds to the next step.
- ***Mutation***. Also a step that is characteristic only for GA. In random order, a random gene can change to random value. The main point in mutation is the same as in biology - to bring genetic diversity to the population. The mutation makes it possible to obtain solutions that could not be obtained with the available genes. Firstly, this helps to avoid getting into local extremes, since the mutation can allow the algorithm to be translated into a completely different branch and, secondly, to "dilute" the population in order to avoid the situation when only the same chromosomes will be in the whole population, which will not generally move towards a right solution.
- After all the steps have been completed, it is estimated whether the population has reached the desired accuracy or has come to limit the number of populations, and if so, the algorithm stops working. Otherwise, the cycle is repeated with the new population until conditions are reached.

3. Problem description

The usage of genetic algorithms in the testing process makes it possible to find the most complex parts of the program, in which the risks due to the assumption of errors are the greatest. Evaluation is due to the use of the fitness function whose parameters are the different weights of each individual operation [5].

By now, many kinds of diagrams have been developed that allow to represent the structure of the program not as a set of actions, but as diagrams with a certain structure. One of such types are diagrams (graphs) of control flows, allowing to represent the whole set of ways of program execution. The main purpose of such diagrams is the describe the program, i.e. to determine the complexity of the code, to check the logic and to directly determine the steps of the program. But if you look at the problem of generating test data, then such a type of diagrams, built on thealready created computer program, allows you to evaluate the quality of the developed code and, within the scope of the task, to assess the importance or complexity of certain branches of the program.

In accordance with the above it was developed a method according to which it would be possible to evaluate the program code and to determine a set of test data that would allow "to walk through" the largest number of operations. The first step is to consider the structural elements of the code. For ease of presentation, you can use the flow diagrams that were mentioned earlierto visualize the structure of the code and to understand in which way the program is executed.

Each action is assigned its own separate block (node), and the directed relationship indicates the direction to the following action. For example, a condition is awarded to one block, but two branches of code will come out of it. Each transition between blocks is assigned a certain weight, depending on which part of the code this action is in, preceded by any complex structural elements, etc.[10].

The weight of each transition is determined by the Pareto rule: 80 percent of the weight on the complex code branch, and 20 percent on the simple one. With the same value of "complexity", the separation occurs on the 50 / 50 basis. For example, if in the code there is an ordinary condition, with one branch with a positive result of the comparison, the action of this condition will account for 80 percent of the weight, and 20 percent for the branch leading to the continuation (without action under this condition). But if both branches are present in the condition, that is, if the condition is fulfilled and not fulfilled, then the weights will be distributed 50% to 50%. An example based on which the algorithm was tested is shown in Figure 1.

The assigned weights can be used to develop test array with the help of genetic algorithms, that is, to estimate how much the calculated weight falls on this or another branch for certain values of the input parameters.

For the sake of convenience, we introduce the following notations:

X – datasets;

F(X) – the fitness function value for each data set depending on the calculated weights.

The task is to maximize the objective function, i.e.

$$F(X) \rightarrow max.$$

4. Results

To test the method, four randomly generated datasets are used: (10,5,12); (3,4,10); (25,30,11); (5,3,17). Table 1 shows the datasets, the calculated fitness function value and the rank determining the best data set.

Table 1. Initial datasets.

№	Dataset X	F(X)	Rank
1	(10,5,12)	896	3
2	**(3,4,10)**	**1196**	**2**
3	**(25,30,11)**	**1308**	**1**
4	(5,3,17)	896	3

In this case, the second and the third variants will be used as the data sets for selection. In order to obtain additional two new variants, their values will be mixed and supplemented by some probability of mutation.

```
0.  public static void Main(int m, int i, int n) {
1.  intsum = 0, avg = 0;
2.  if (m <i)
3.  {
4.  use_m = m;
5.  }
6.  int[] a = new int[m];
7.  Console.WriteLine("Enter the Array Elements
    ");
8.  for (j= 0; j< m; j++)
9.  {
10. a[j] = i;
11. if (a[j] > n)
12. {
13. a[j] = n
14. }
15. }
16. for (k= 0; k < m; k++)
17. {
18. sum += a[k];
19. }
20. avg = sum / m;
21. Console.WriteLine("Average is {0}", avg);
22. Console.ReadLine();
23. }
```

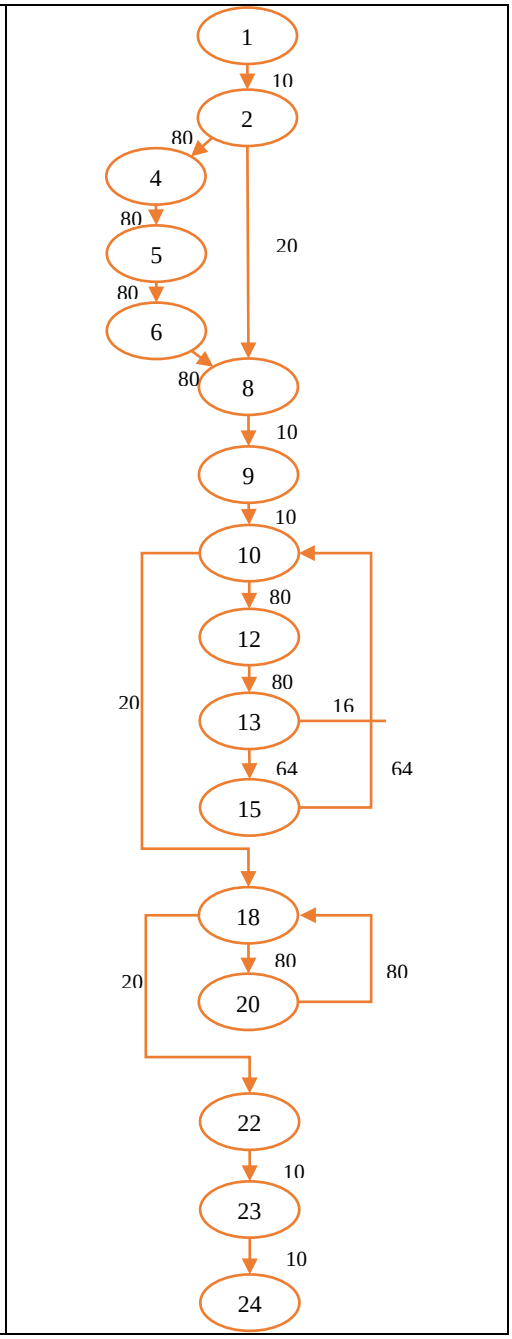


Figure 1. Program code and sequence of blocks of the analyzed program.

Division of the sets will occur at the first and second positions: when crossing (X, X, X) and (Y, Y, Y) the variables (X, X, Y), (X, Y, Y), (Y, X, X) and (Y, Y, X) will be obtained. Parental values remain to maintain the purity of the crossing, i.e. in comparison with the zero generation in the first generation the two new sets will be added. In subsequent generations, six sets of data will be stored.

The mutation will occur with a probability of 10 percent for the chance of changing the value from 1 to a given value in both directions. Under these conditions, the maximum possible additional value for the mutation is 5. As a result of crossing, the data sets presented in Table 2 will be obtained.

As a result, two parents will be added to the additional two sets - (3,4,13) and (25,30,10). Table 3 shows all the new variants of the test data sets.

With equal ranks, priority will be given to the data sets of a newer generation. In the last generation, three sets of data were obtained, which check most of the branches of the program -

(25,30,11), (3,30,11) and (25,30,10). The first set was obtained from the first generation, so it will be excluded and as a result there will be two variants - (3,30,11) and (25,30,10).

Table 2. New data sets obtained as a result of crossing.

№	X	Y	Dataset	**Matation**
1	(3,4,10)	(25,30,11)	(3,4,11)	**(3,4,13)**
2	(3,4,10)	(25,30,11)	(3,30,11)	**(3,30,11)**
3	(3,4,10)	(25,30,11)	(25,4,10)	**(25,4,10)**
4	(3,4,10)	(25,30,11)	(25,30,10)	**(25,30,10)**

Table 3. The first generation of test data.

№	Dataset X	F(X)	Rank	Generation
1	(3,4,10)	1196	2	0
2	(25,30,11)	1308	1	0
3	(3,4,13)	1196	2	1
4	(3,30,11)	1308	1	1
5	(25,4,10)	896	3	1
6	(25,30,10)	1308	1	1

Due to the small initial sample and a small size of code, the datasets quickly find intersecting values - 30 at the second position and 10 at the third one. Therefore, to continue iterating ceases to make sense - in the next generation the sets will consist mainly of repeating sets.

For this code, onecan use test datasets obtained in the last generation. Priority depends on the rank obtained.

Thus, using genetic algorithms, it is possible to find initial test values that would fully test most complex program variants depending on the assigned weights. This is testing through the maximization of the fitness functions, since the most widely describing test variant passes through those branches having the greatest weight.

As the values of chromosomes, the test sample values themselves appear. Many test cases in one iteration become a population. Using the genetic algorithm, it becomes possible to view a set of test samples to find the best initial variants.

5. Improvement of the algorithm

In accordance with the previously considered variant of using, this method is being developed to better correspondence to the real requirements. The weight is considered in accordance with the operability of the program, in other words, the more iterations the program executes, the greater weight will have the original test version. It allows you not only to estimate which branches of the program will be used for such data, but also how they change directly during the program operation and what results will be obtained.

This not only makes it possible to determine the initial test dataset, but also on the basis of the final population to draw a conclusion about changing and modification of data in accordance with the predefined logic. Four additional tests were conducted to present the results.

The first population is formed by random values. Each population contains 100 chromosomes. The total number of the population is also equal to 100. This formed a sufficient number of different variants and the best of them will be selected.

Table 4 shows the results of testing.

In each of the tests, at least two different versions of the data sets were formed, under which the program code will work most of the time and more times will take place inbranches. In addition, you can see certain patterns in the results - the first value is always maximum (random values were limited to 100), the second value is less than the first, but more than the third.

Table 4. Comparison of the results.

Population	Test 1	Test 2	Test 3	Test 4
0	1: 78, 23, 35	1: 97, 3, 6	1: 92, 97, 28	1: 15, 67, 26
	2: 62, 36, 95	2: 82, 77, 64	2: 38, 66, 52	2: 32, 27, 83
	3: 52, 35, 27	3: 24, 47, 57	3: 63, 76, 64	3: 37, 52, 64
	4: 17, 77, 73	4: 90, 13, 82	4: 7, 24, 56	4: 70, 49, 64
	5: 75, 9, 96	5: 81, 69, 24	5: 57, 48, 8	5: 67, 29, 94
20	1: 95, 64, 54	1: 97, 80, 4	1: 99, 13, 10	1: 99, 71, 45
	2: 95, 64, 29	2: 97, 80, 53	2: 99, 13, 11	2: 99, 71, 15
	3: 95, 64, 54	3: 97, 80, 28	3: 99, 13, 11	3: 99, 71, 3
50	1: 95, 64, 54	1: 97, 80, 29	1: 99, 13, 10	1: 99, 71, 60
	2: 95, 64, 29	2: 97, 80, 4	2: 99, 13, 11	2: 99, 71, 3
	3: 95, 64, 54	3: 97, 80, 53	3: 99, 13, 11	3: 99, 71, 3
Final (100)	**1: 95, 64, 54** **2: 95, 64, 29**	**1: 97, 80, 4** **2: 97, 80, 29**	**1: 99, 13, 10** **2: 99, 13, 11**	**1: 99, 71, 60** **2: 99, 71, 45**

6. Conclusion

The task of determining the most suitable test data is very difficult to solve with the help of standard algorithms, since the actual number of the entire array of possible values is incredibly huge, and it is not possible to select really suitable values. At the same time, automation of this process can significantly reduce the analytical burden on users involved in testing.

The use of genetic algorithms allows to compare many different data options for testing a program. The wide possibilities for improvement allow increasing the number of initial test variants, the number of generations and adding new properties, which helps to substantially increase the possibilities of finding more suitable variants. If you follow the passed graphs and reduce the weights of graphs being most often encountered in different variants, you can search for new paths that may not be available at the moment but can be as important as the most frequently encountered ones.

The generated tests can serve not only for testing the algorithm, but also for analyzing its performance. The data in its pure form already allow us to determine the laws of the program code, and with a deeper study they can allow us to draw conclusions about a possible subsequent improvement. Such an analysis can be used as a topic for further research in the field of data analysis.

In addition, it is necessary to solve several ambiguous tasks, because of which a simple assignment of weights cannot provide a qualitative search for the branches of the program:

- Transitions between different parts of programs. For example, in object-oriented programming it can be the launch of other subclasses, each of which has its own paths, the launch of other functions, and the calculation of changes in values. A lot of recommendations for programmers are related to the fact that the most difficult to understand parts are taken out in separate procedures, which makes it very difficult to find really important parts of the program.
- Loops that significantly complicate an adequate evaluation of the code, because some may occupy a huge share in the resource costs, but at the same time have a fairly simple logic. As a result, it becomes difficult to assess whether any particular cycle is an important part of the program and in fact how many iterations can be run under certain conditions.
- Insufficient code development. Often an input restriction is introduced outside the code, because of which, when using a genetic algorithm, a data set can potentially be caught, under which the program will never be executed. On the one hand, it helps to find such problem places, on the other hand, it puts a restriction directly on the algorithm.

These are not the only problematic tasks, because of which there are difficulties in correctly assessing the operability of the program. Nevertheless, their solution will allow not only to fulfill the current given goal of the proposed method, but also to obtain additional benefits from its use, such as determining non-optimal parts of the code, unnecessary operations, "littering" the code, and so on. As a result, further research will allow us to obtain new, previously undefined results.

7. References

[1] Zanetti M C, Tessone C J, Scholtes I and Schweitzer F 2014 Automated Software Remodularization Based on Move Refactoring. A Complex Systems Approach *3th international conference on Modularity* 73-83
[2] John K and Tiegelkamp M 2010 *Programming Industrial Automation Systems: Concepts and Programming Languages, Requirements for Programming Systems, Decision-Making Aids* (Springer, Berlin) p 388
[3] Luger F 2009 *Artificial Intelligence Structures and Strategies for Complex Problem Solving* (University of New Mexico) p 679
[4] Berndt D J, Fisher J, Johnson L, Pinglikar J and Watkins A 2003 Breeding Software Test Cases with Genetic Algorithms *Proceedings of the Thirty-Sixth Hawaii International Conference on System Sciences* 36
[5] Praveen R S and Tai-hoon K 2009 Application of Genetic Algorithm in Software Testing *International Journal of Software Engineering and Its Applications* **3(4)** 87-96
[6] Serdyukov K and Avdeenko T 2017 Investigation of the genetic algorithm possibilities for retrieving relevant cases from big data in the decision support systems *CEUR Workshop Proceedings* **1903** 36-41
[7] Yang H L and Wang C S 2008 Two stages of case-based reasoning - Integrating genetic algorithm with data mining mechanism *Expert Systems with Applications* **35** 262-272
[8] Mühlenbein H 1992 How genetic algorithms really work: Mutation and hillclimbing *Parallel Problem Solving from Nature* **2**
[9] Spears W M 1993 Crossover or mutation? *Foundations of Genetic Algorithms* **2**
[10] Coyle L and Cunningham P 2004 Improving recommendation ranking by learning personal feature weights *Proc. 7th European Conference on Case-Based Reasoning* 560-572

Acknowledgments
The work is supported by a grant from the Ministry of Education and Science of the Russian Federation within of the project № 2.2327.2017/4.6.

Analysis of the influence of citizens' altruism on the effectiveness of the socially-optimal actions stimulation system

M I Geraskin[1]

[1]Samara National Research University, Moscowskoe shosse 34, Samara, Russia, 443086

Abstract. The problem of the development of the state information stimulation system of Russian citizens' socio-optimal actions is considered according to the optimum of collective utility function as criterion. The algorithm of the information system is formed taking into account the conditions of individual rationality, Pareto efficiency and non-manipulability. The algorithm for analyzing the effect of various functions of the probability distribution density of the altruism degree of Russian citizens on the effectiveness of the system for stimulating socially-optimal actions is developed. The simulation of social groups behavior, covering more than ten percent of Russia's economically active population, is confirmed the stability of the stimulation system for citizens' opportunistic behavior.

1. Introduction

In a transitional economy [1,2] trends of individual rationality are growing in the society. The state develops the moral improving programs [3,4] to overcome this trends. In this case, the purpose of the state is the social effect of citizens' acts, performing on the basis of maximization of collective but not individual utility function, hereinafter referred to as socio-optimal actions. Achieving this purpose requires the involvement in socially useful activity of large group of the population and personified registration of socio-optimal actions. It needs to organize the state information system, based on the information resources of currently working in Russia programs [5-7].

The concept of socio-optimal actions' stimulation provides for the establishment of information system of personified registration of the actions of citizens (hereinafter, agents). The system also includes the distribution of state stimulation fund in the form of incentives between agents according to certain mechanisms. The dynamics of the system is a two-period. In the first period (registration period) performed socio-optimal actions are recorded, and in the end of this period the stimulation fund is distributed. In the next period (period of stimulation) the earlier distributed incentives are used.

In addition to the utilitarian stimulating function information system also solves the problem of the formation of the agent's status in the hierarchy of the citizens, used for non-material motivation. On a longer time horizon, the state's social priorities could changed by varying the attributes of socio-optimal actions and their monetary valuation can be varied as a result of inflation. Therefore, to comparability of agents' statuses the system accumulates not only incentives as the current cash equivalent of social activity, but also agents' rating in comparable dimension.

The object of stimulation is socio-optimal actions of citizens, that is, actions that correspond to certain attributes. The actions should maximize collective utility function without increasing the individual utility function. Therefore, the attributes correspond to the terms of gratuitousness, public utility and unconnectedness with professional activities of citizens. Consequently, socio-optimal

actions do not require special qualification, whereby the stimulation object's dimension is duration of action excluding the content of the action. The subject of stimulation is citizen, performing a socio-optimal action in certain period. The apparatus of stimulation is the state represented by certain ministries (departments).

2. Methods

The investigations of stimulation systems and distribution mechanisms produce the following mechanisms corresponding the individual rationality. Competitive mechanism is developed with noncooperative [8] and cooperative [9] behavior of agents, its Pareto efficiency and optimality according to additive utility function criterion are proved. The step-by-step resource distribution mechanism (SRDM) is obtained [10], for which proved [11] that nonmanipulability and Pareto efficiency simultaneously only for SRDM; also SRDM, as shown in [12], is equivalent to mechanisms of direct and reverse priorities. It was shown [13] that unique SRDM exists, in which the incentive is distributed [14] as minimum of agent's information and the average undistributed rest of incentives. The approach to the distribution based on the penalty and incentive functions [15] showed the Pareto efficiency and optimality according to additive utility function criterion for compensatory mechanisms; according to a compensatory mechanism incentives are equal to agents' costs. Thus, only SRDM satisfies [16] all above conditions. Since SRDM implies consistent registration of agents' actions and further distribution of the incentives, it is impossible to use in the system, where actions perform independently and record simultaneously with asymmetrical agents' behavior [17,18]. The interaction distribution algorithm in a strongly coupled system with a transferable utility is developed [19,20].

On the base of this ideas an algorithm of the information system (Fig. 1) was developed [21]. The algorithm satisfied the conditions of individual rationality, Pareto efficiency, non-manipulability and optimality by the collective utility criterion.

In [19] and later in this paper the following notation is used:

$Z = \{z_i, i = 1,..., I\}$ is a set of attributes of socially-optimal actions;

$K(t) = \{1,..., n(t)\}$ is a set of agents, including citizens who perform actions that correspond to the attributes of *Z*; index *n(t)* denotes the number of agents in a certain period *t*;

$A(Z,t) = \{a_k(Z,t), k \in K\}$ is a vector of socially-optimal actions; the vector includes quantitative estimations of the actions of the *k*-th agent corresponding to the attributes *Z* in the *t*-th period; estimation *a* is expressed in the time spent by the agent for performing these actions; the vector belongs to the allowable set $\overline{A} = \left\{a_k \in \left[0, a^{\max}\right], a^{\max} > 0, k \in K\right\}$, where the symbol $a^{\max}$ denotes the upper limit of the agents' available time;

$F(t) \in \left(0, F^{\max}\right], F^{\max} > 0$ is a stimulation fund in the *t*-th period $t \in (0,T]$; further, the index *t* is omitted, all parameters of the model correspond to a certain period of time;

$u_k = \psi(a_k), k \in K$, $\psi(a_k) = \alpha a_k^{\beta}, \alpha \in \left(0, \alpha^{\max}\right], \beta \in \left(0, \beta^{\max}\right], \beta^{\max} \in (0,1], k \in K$ is dimensionless function of registration of socially-optimal actions; where *u* is a score, α, β are constant coefficients;

$E = \sum_{k \in K} a_k$ is a system effectiveness indicator;

$x_k = \varphi(u_k) = \varphi^{\min} + \left(b_1 - b_2 \sum_{k \in K} u_k\right) a_k, k \in K$ is monetary incentive function; where $\varphi^{\min}$ is guaranteed incentive of non-zero agent's action, $b_1, b_2 > 0$ are coefficients of the incentive function, calculated from formulas

$$b_1 = \frac{F - n\varphi^{\min}}{\sum_{k \in K} u_k} \frac{2\overline{u} + \sum_{k \in K} u_k}{2\overline{u}}, b_2 = \frac{F - n\varphi^{\min}}{2\overline{u} \sum_{k \in K} u_k}, \overline{u} = \frac{1}{n} \sum_{k \in K} u_k .$$

The algorithm of a one-period cycle of the information system is shown in Fig. 1.

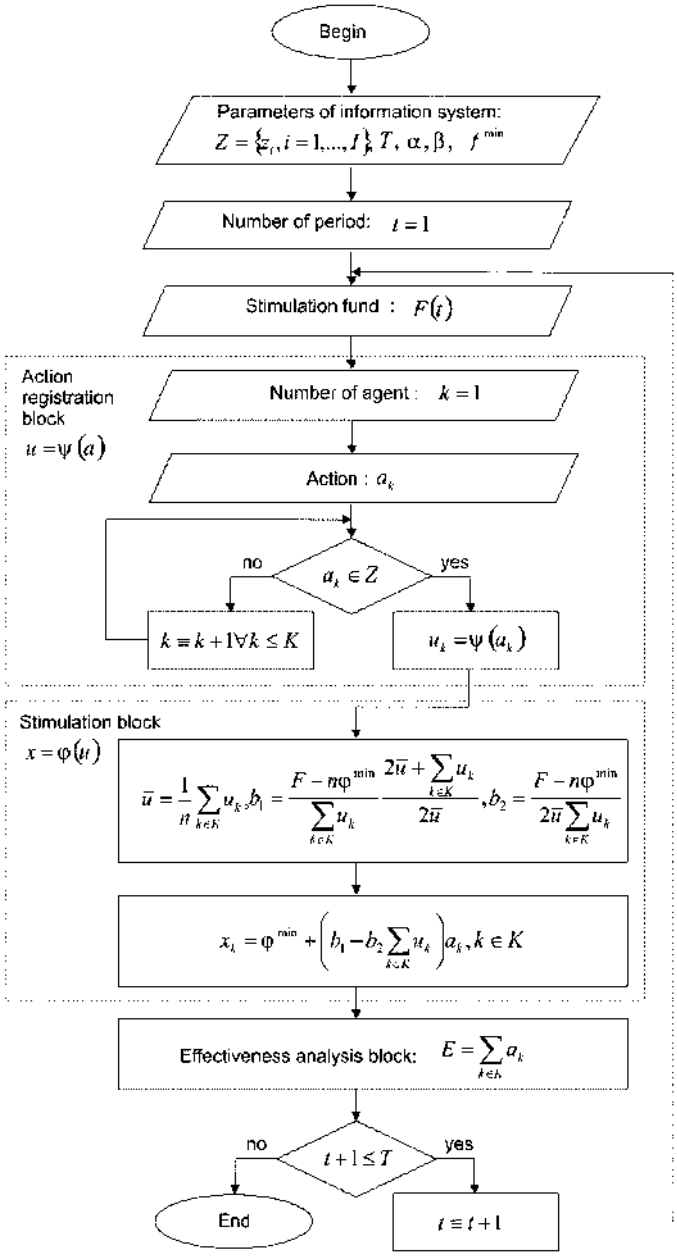


Figure 1. Algorithm of information system cycle.

The action registration block identifies agents' actions in the scoring. The stimulation block is intended for distribution of the stimulation fund depending on the vector of scoring actions. The effectiveness analysis block controls the dynamics of the change in the social efficiency criterion of the system in the selected time interval until the maximum number of work periods is reached.

Let us consider the problem of the information system algorithm modeling for various probability density functions of the existence of social groups with a greater or lesser inclination for altruism.

3. Results and discussion

Let the following continuous functions are defined for the *k*-th agent

$$a_k(D_k) = D_k^{\delta_{ak}}, d_k(D_k) = D_k^{\delta_{dk}}, \delta_{ak}, \delta_{dk} \in [0,1], \delta_{ak} + \delta_{dk} = 1, D_k >> 1, k \in K, \quad (1)$$

where d_k is the working time interval; D_k is an available time fund equal to the physical time fund except for the rest time; δ_{ak}, δ_{dk} are elasticity coefficients of «charitable» time and working time on D_k.

The functions (1) express the connection between the time interval of charitable actions, the time interval of working time and the available time fund. The derivatives of the functions (1) decrease with increase of D_k, expressing the propensity of individuals [22] to increase the rest time with growth D_k.

Definition: altruism (propensity to charity) of the *k*-th agent is a type of function (1), under which $\delta_{ak} > \delta_{dk}$.

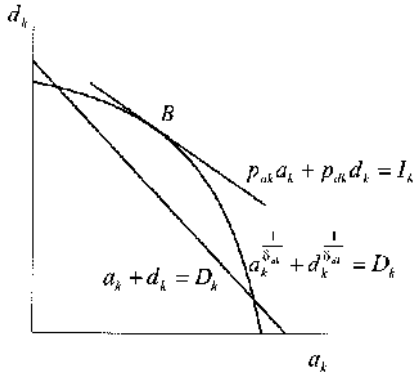


Figure 2. Graphical interpretation of the problem (2), (3).

The introducing of the socially-optimal actions stimulation system by the algorithm (Fig. 1) leads to the fact that altruistic actions bring a non-zero income.

Therefore, to select the value of the action, the *k*-th agent, taking into account the individual rationality, solves the following problem: to maximize the total income from the working and «charitable» time with a restriction[1] on the available time fund:

$$\max_{a_k \in A} I_k = \max_{a_k \in A} \left(p_{ak} a_k + p_{dk} d_k \right), p_{ak}, p_{dk} > 0, p_{ak} = \frac{x_k}{a_k}, \tag{2}$$

$$a_k^{\frac{1}{\delta_{ak}}} + d_k^{\frac{1}{\delta_{dk}}} = D_k, k \in K, \tag{3}$$

where p_{ak} is the price of the «charitable» time; p_{dk} is the price (tariff rate) of working time.

Note that the restriction (3) does not coincide with the identity $a_k + d_k = D_k$, because of the «charitable» time is not an absolute substitute in relation to working time, as shown in Fig. 2.

The solution of the problem (2), (3) has the form[2]:

$$\frac{\delta_{dk}}{\delta_{ak}} \left(D_k - a_k^{*\frac{1}{\delta_{ak}}} \right)^{\delta_{dk}-1} a_k^{*\frac{1}{\delta_{ak}}-1} = \frac{p_{ak}}{p_{dk}}, \tag{4}$$

$$d_k^* = \left(D_k - a_k^{*\frac{1}{\delta_{ak}}} \right)^{\delta_{dk}}, k \in K, \tag{5}$$

[1]The restriction (3) follows from (1): $D_{ak} = a_k^{\frac{1}{\delta_{ak}}}, D_{dk} = d_k^{\frac{1}{\delta_{dk}}}, D_{ak} + D_{dk} = D_k$.

[2]The equation $d_k = \left(D_k - a_k^{\frac{1}{\delta_{ak}}} \right)^{\delta_{dk}}$ follows from (2); therefore $d'_{ka_k} = -\frac{\delta_{dk}}{\delta_{ak}} \left(D_k - a_k^{\frac{1}{\delta_{ak}}} \right)^{\delta_{dk}-1} a_k^{\frac{1}{\delta_{ak}}-1}$; as at the tangency point *B* the angular coefficients of the functions (2) and (3) are equal, the optimum condition has the form (4).

where the symbol «*» denotes the optimal values.

The elasticity coefficients δ_{ak}, δ_{dk} are known, and price vectors are given

$$\boldsymbol{P}_a = \{p_{ak}, k \in K\}, \boldsymbol{P}_d = \{p_{dk}, k \in K\}, \quad (6)$$

the first vector is calculated on the basis of the results of algorithm (Fig. 1) by formula (2), and the second vector is given based on the statistics of the national labor market. Thus, the solution of the equation (4) gives an individually rational vector of socially-optimal actions $\boldsymbol{A}$.

Let the coefficient δ_{ak} in the general population of agents (population) is a random variable [23] with a normal distribution law [21]:

$$f(\delta_a) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{w(\delta_a - \bar{\delta}^l)^2}{2\sigma^2}}. \quad (7)$$

where $\bar{\delta}, \sigma$ are the mathematical expectation and the standard deviation of the initial distribution of the random variable; l is a coefficient taking into account the asymmetry ($l > 1$ - left asymmetry, $l < 1$ - right asymmetry at $0 < \bar{\delta} < 1$) in comparison with the normal law ($l = 1$); w is a coefficient that takes into account kurtosis ($w < 1$ is a more uniform distribution, $w > 1$ is a less uniform distribution) compared with the normal law ($w = 1$).

In this case, the expected values of the agents' number having these values of the elasticity coefficients $\delta_{ak}, \delta_{dk} = 1 - \delta_{ak}$, are calculated by the formulas

$$\hat{n}(\delta_a) = f(\delta_a) n, \hat{n}(\delta_d) = n - \hat{n}(\delta_a),$$

where $\hat{n}(\delta_a), \hat{n}(\delta_d)$ are the expected values of the agents' number at given values of the elasticity coefficients.

The algorithm of the effectiveness analyzing block of the information system for various parameters of the probability density function of the social groups presence with a greater or lesser inclination for altruism is shown in Fig. 3.

The simulation of the stimulation effect on the population's behavior is carried out by changing the asymmetry and kurtosis of the probability density function (7) of a normal distribution. The characteristics of the population's distribution variants by social groups with greater or lesser inclination for altruism are given in Table. 1. The variant θ = 1 corresponds to the normal distribution law; variant θ = 2 simulates the case of exceeding the mathematical expectation of the population's inclination to altruism over the median value due to right asymmetry; variant θ = 3 simulates the case of a more uniform distribution of the population's propensity for altruism in comparison with the normal distribution law due to the decrease in kurtosis.

Table 1. Variants of population's distribution.

Model parameter	Variant		
	θ=1	θ=2	θ=3
Characteristics	normal distribution	normal distribution with right asymmetry	normal distribution with reduced excess
l	1	0,8	1
w	1	1	0,5

The following initial data are considered: the value of the stimulation fund is set $F = 1000000$ thousand rubles, the values of the registration function coefficients (α = 67, β = 0,8) are selected from the condition $\bar{u} = 100$ at the initial value of the time price and θ = 1, the minimum stimulus is assumed equal to zero, the number of the population n = 9727 thousand corresponds to the group of Russian Federation population with income below the subsistence minimum, the price of working time is P_d = 0.21 thousand rubles per hour based on the average wage of 35.369 thousand rubles per month, the range of the elasticity coefficient of the «charitable» time is chosen taking into account the possibility

of equation (4) solution equal to $\delta_a \in [0{,}1;0{,}7]$. Since $\varphi^{\min} = 0$, the price of «charitable» time is the same for all agents and it is equal to $P_a = \frac{x_k}{a_k} = b_1 - b_2 \sum_{k \in K} u_k$.

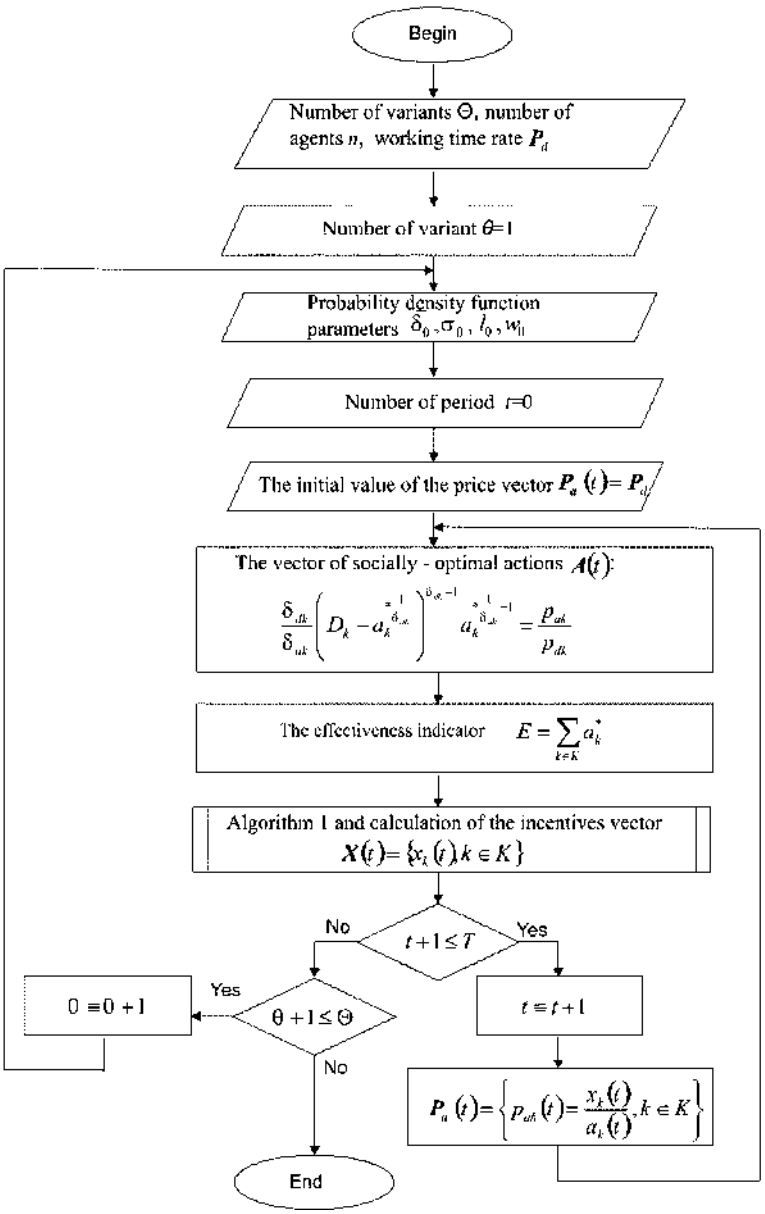


Figure 3. Algorithm of the effectiveness analyzing block of the information system.

In Fig. 4 the distribution functions of the population depict depending on the elasticity coefficient of the «charitable» time for various simulation variants.

In Fig. 5 the functions of average per capita effectiveness indicator $E_{av.} = E/n$ for different simulation variants are given. The sharp increase in the effectiveness indicator at t = 1 is due to the increase in the price of the «charitable» time (Fig. 8) in the first period in comparison with the initial date adopted by P_a=P_d to adjust the stimulation system in the zero period. From Fig. 6 it follows that the reason for this growth is a significant increase in socially-optimal actions in groups with relatively low of «charitable» time elasticity coefficient in the range (0.2, 0.5).

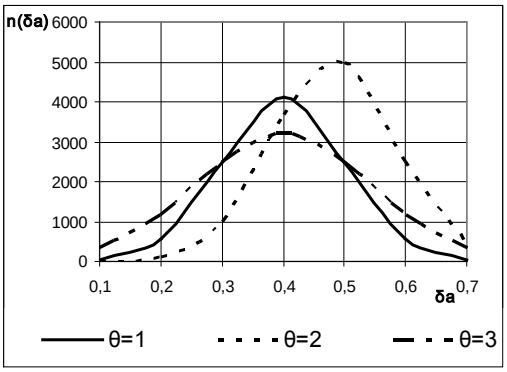


Figure 4. Distribution functions of the population (thousand) depending on the value of the elasticity coefficient of the «charitable» time for various simulation variants.

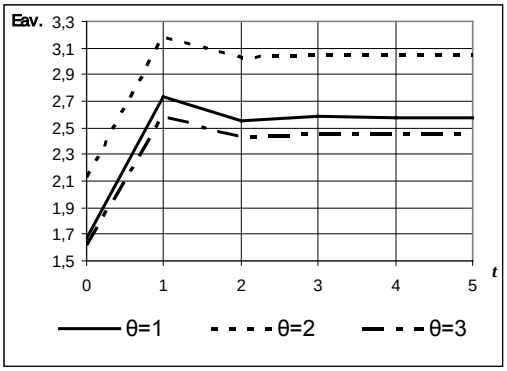


Figure 5. Dynamics of the average per capita effectiveness indicator for various simulation variants.

Population groups with very low values (0,1 ... 0,2) and very high (0,5, 0,7) values of the «charitable» time elasticity coefficient are less susceptible to the stimulating effect of the high price P_a: the first because of the prevalence of work in their disposable time fund, the latter due to high altruism. Since at t = 2, as the price P_a decreases (Fig. 8), the effectiveness indicator also decreases taking into account (Fig. 6) sharp reduction of such actions in the range (0.2, 0.5). Consequently it can be concluded that population groups with relatively low «charitable» time elasticity coefficient are prone to opportunistic behavior, that is, the motivation for socially-optimal actions for them is individual rationality. As a result, in all variants θ = 1,2,3 the value of the average effectiveness indicator stabilizes even at t = 3.

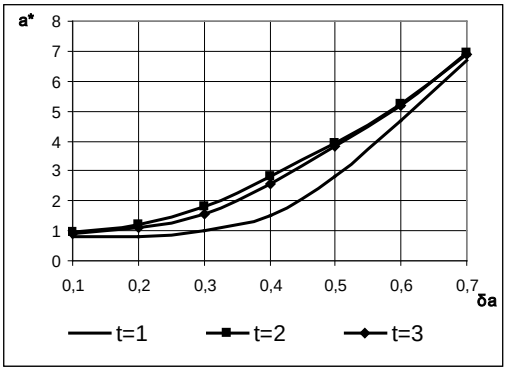


Figure 6. Dependence of socially-optimal actions on the elasticity coefficient of the «charitable» time for θ=1.

Stimulation is most effective at θ = 2, the least effective at θ = 3. Consequently, the results of the simulation are consistent with the following provisions: 1) an increase in the prevalence in a society of citizens with a higher propensity for altruism leads to an increase in the social effect from charity stimulating; 2) in case of stimulation a more equal prevalence in a society of individuals with high and low inclination to altruism leads to a reduction in the number of socially optimal actions in comparison with the normal distribution.

In Fig. 7 the average registration score per capita $u_{av.} = \overline{u}$ dependences on the period for different simulation variants are given. In all variants, the value $u_{av.}$ stabilizes at t = 3. The highest values of

average score are recorded at θ = 2, the lowest values are recorded at θ = 3, which is due to the dependence of the average score through the registration function on the social effect $E_{av.}$.

In Fig. 8 the «charitable» time price dependences on the period for different simulation variants are given, which also stabilize at t = 3. The lowest price *Pa* corresponds to θ = 2, the highest price is reached in the case θ = 2. Since the price *Pa* is a state incentive for a single socially-optimal action, the stimulation system is most economically effective at θ = 2, the least effective at θ = 3.

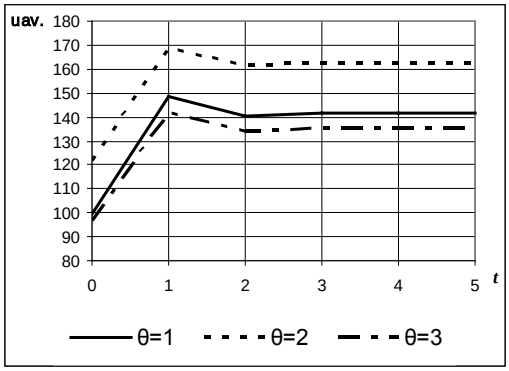


Figure 7. Dynamics of the average score per capita of the registration function for various simulation variants.

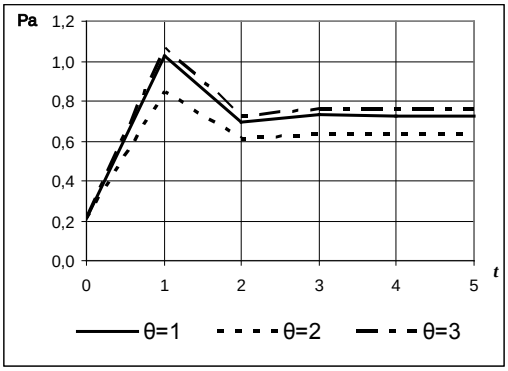


Figure 8. Dynamics of the «charitable» time price (thousand rubles / hour) for various simulation variants.

The stabilization of the effectiveness indicator of the simulation system (Fig. 5) and the price of time (Fig. 8) at t> 2 leads to the conclusion about the stability of the simulation system for opportunistic behavior of citizens: persons with relatively low values of «charitable» time elasticity, performing socially-optimal actions based on individual rationality, demonstrate the greatest negative sensitivity to the change in the price of time as a stimulus. Due to this, the positive deviation of the stimulus in one period from a certain equilibrium value is damped by the opportunistic behavior of citizens, leading in the next period to a decrease in the number of socially-optimal actions.

4. Conclusion

The problem of informational support of the state strengthening morality strategy by stimulating citizens' actions performed on the basis of maximizing the collective utility function is considered. The following main results are obtained in the article.

The algorithm for analyzing the effectiveness of the information simulation system for various parameters of the probability distribution function of social groups with greater or lesser inclination to altruism in society is developed.

The influence of stimulation on the population's behavior is studied by changing the asymmetry and kurtosis of the probability density function of the normal distribution. The case of exceeding the mathematical expectation of the population's inclination to altruism over the median value is simulated by introducing asymmetry. The case of a more uniform distribution of the population's inclination to altruism in comparison with the normal distribution law is simulated by the introduction of kurtosis.

The analysis shows that the stimulation system is resistant to opportunistic citizens' behavior due to the fact that persons performing socially-optimal actions based on the individual rationality demonstrate the greatest negative sensitivity to the change in the stimulus.

5. References

[1] Roland G 2000 Transition and Economics. Politics, Markets, and Firms (Cambridge: MIT Press) p 840

[2] Braguinsky S, Yavlinsky G 2000 Incentives and Institutions. Transition to a Market Economy in Russia (NJ.: Princeton University Press) p 420

[3] RF Government Decree of 30.12.2015 N 1493 *"On State program" Patriotic Education of Citizens of the Russian Federation for 2016 - 2020"*

[4] RF Government Decree of December 27, 2012 N 2567-r *"On the state program of the Russian Federation" Development of Culture and Tourism "2013 - 2020"*
[5] RF Government Decree of 15.04.2014 N 313*"On approval of the Russian Federation, the state program" Information Society (2011 - 2020)"*
[6] RF Government Decree of 27.12.2012 N 1406 *"On the federal target program" Development of the Russian judicial system for 2013 - 2020"*
[7] RF Government Decree of 15.04.2014 N 320 *"On approval of the state program of the Russian Federation" Public Financial Management and regulation of financial markets"*
[8] Burkov V N, Danev B, Enaleev A K, Nanev T B, Podvalny L D and Yusupov B S 1988 Competitive mechanisms in problems of distribution of scarce resources *Automation and Remote Control* **11** 142-153
[9] Burkov V N, Enaleev A K and Kalenchuk V F 1989 Coalition with the competitive mechanism of resource distribution *Automation and Remote Control* **12** 81-90
[10] Burkov V N, Enaleev A K and Lavrov Y G 1992 Synthesis of optimal planning and incentive mechanisms in the active system *Automation and Remote Control* **10** 113-120
[11] Burkov V N, Iskakov M B and Korgin N A 2010 Application of generalized median schemes for the construction of non-manipulable mechanism multicriterion active expertise *Automation and Remote Control* **71(8)** 1681-1694
[12] Korgin N A 2009 Equivalence of non-manipulable and non-anonymous priority resource distribution mechanisms *Managing large systems* **26.1** 319-347
[13] Burkov V N, Gorgidze I I, Novikov D A and Yusupov B S 1997 Models and cost and revenue distribution mechanisms in the market economy (Moscow: Institut problem upravleniya) p 356
[14] Korgin N A 2010 Use of intersection property for analysis of feasibility of multicriteria expertise results *Automation and Remote Control* **71(6)** 1169-1183
[15] Chumak V G, Ramzaev V M and Khaimovich I N 2015 Challenges of Data Access in Economic Research based on Big Data Technology *CEUR Workshop Proceedings* **1490** 327-337
[16] Burkov V N, Korgin N A and Novikov D A 2016 Problems of aggregation and decomposition mechanisms of management of organizational and technical systems *Management issues* **5** 14-23
[17] Lyubimov V V and Lashin V S 2017 External Stability of a Resonance during the Descent of a Spacecraft with a Small Variable Asymmetry in the Martian atmosphere *Advances in Space Research* **59(6)** 1607-1613
[18] Lyubimov V V 2015 Numerical Simulation of the Resonance Effect at Re-entry of a Rigid Body with Low Inertial and Aerodynamic Asymmetries into the Atmosphere *CEUR Workshop Proceedings* **1490** 198-210
[19] Dodonova N L and Kuznetsova O A 2017 About scarce resources allocation in conditions of incomplete information *CEUR Workshop Proceedings* **1904** 130-134
[20] Geraskin M I and Kuznetsova O A 2017 Agents' Interaction algorithm in a strongly coupled system with a transferable utility *CEUR Workshop Proceedings* **2018** 32-42
[21] Geraskin M I 2017 Algorithms of the information stimulation system of Russian citizens' socio-optimal actions *CEUR Workshop Proceedings* **1903** 92-99
[22] Novikov D 2013 *Theory of Control in Organizations* (New York: Nova Science Publishers) p 341
[23] Kulikovskikh I M 2017 Anomaly detection in an ecological feature space to improve the accuracy of human activity identification in buildings *Computer Optics* **41(1)** 126-133 DOI: 10.18287/2412-6179-2017-41-1-126-133

Personal data segmentation based on conjugation index usage

P V Hripunov[1] and D A Zherdev[2]

[1]Pension Fund of the Russian Federation, Shabolovka str. 4, Moscow, Russia, 119991
[2]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086

Abstract. The paper proposes a method for processing personal data that allows them to be divided into many segments or classes. The customer database is used as the source data. We use the indicator of conjugacy that has already proved the effectiveness in both recognition and clustering of data problems.

1. Introduction

Data mining problem is a primary problem inprocessing of huge amount data. Different methods of pattern recognition, classification, images clustering and others have foundthe implementation in the data mining. Many works [1-4] study clusterization processes of big data. In this study, we research the recognition ability of some personal data that was presented by a digit vector. For classification within some probability, it is necessary to find out is a similar element consist in the database. To achieve this, we use conjugation index. The index effectiveness was shown in study of face recognition problems [5], as well as objects recognition in radar images [6], [7].

2. Segmentation and classification of personal data

In this section the clustering approach based on conjugation index usage is described. We can decide that a vector belongs to a class by calculation of conjugation index value. The higher probability of conjugation index shows that a vector has similarity to the vectors that form a class:

$$\mathbf{X}_k = \left[\mathbf{x}_1(k), \mathbf{x}_2(k), ..., \mathbf{x}_j(k), ..., \mathbf{x}_M(k)\right], k = \overline{1,K},$$

where $\mathbf{x}_j = \left[x_1, x_2, ..., x_i, ..., x_N\right]$ is a $N\times 1$ feature vector.

The conjugation index can be presented as:

$$R_k\left(\mathbf{x}_j\right) = \frac{\mathbf{x}_j^T \mathbf{Q}_k \mathbf{x}_j}{\mathbf{x}_j^T \mathbf{x}_j}, \quad k = \overline{1,K},$$

where K is a class count,

$$\mathbf{Q}_k = \mathbf{X}_k \left[\mathbf{X}_k^T \mathbf{X}_k\right]^{-1} \mathbf{X}_k^T, \quad k = \overline{1,K},$$

is a $N \times N$ matrix of k -class.

Each letter in a single categorical data is coded by some index thus digital vector of a string can be formed. Each letter of Russian alphabet "А-Я" coded by numbers1-33. In the result the new database of vectors can be formed.

There are three fields in the database: first name, second name and middle name. For convenience, the maximum number of possible symbols in the database was chosen to be 100. As the result, all

vectors in the new dataset contain 300 features. All vectors in dataset must be the same size. For example, if the first field size equals 12 symbols then the first 12 features of a vector contain the field value and other 88 features are filled by zero value.When we processed the database, all personal data was encrypted by summing up with some digital key.

We use the similar procedure for clustering which was used in work [6] for clustering of radar images. At first step of the whole set we choose the two most "distanced" vectors. These vectors have the minimal value of the correlation ratio and we can be labeled them as $\mathbf{x}_1$, $\mathbf{x}_M$.

Then the algorithm from the remain set of vectors adds two new vectors ($\mathbf{x}_2$, $\mathbf{x}_{M-1}$). Each one of these vectors must have the maximum of the correlation ratio:

$$R_{1,2} = \frac{(\mathbf{x}_1^T \mathbf{x}_2)^2}{\|\mathbf{x}_1\|\|\mathbf{x}_2\|},$$

$$R_{M-1,M} = \frac{(\mathbf{x}_{M-1}^T \mathbf{x}_M)^2}{\|\mathbf{x}_{M-1}\|\|\mathbf{x}_M\|},$$

with one vector was obtained at the first step. In the result received pairs of vectors $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_{M-1}$, $\mathbf{x}_M$ form the subspaces that were formed by matrices $\mathbf{X}_{1,2}$ and $\mathbf{X}_{M-1,M}$ correspondingly. Then using the remaining set next two vectors $\mathbf{x}_3$, $\mathbf{x}_{M-2}$ that are closest to the subspaces are joined to previously formed subspaces using computation ofconjugation index with a maximum value.

Since the database contains a large number of vectors, process continues due finding of specific number of vectors in both subspaces. For example, for such dataset the resulted subspace contains 15 vectors in both matrices $\mathbf{X}_k$, $\mathbf{X}_l$, which correspond to two subclasses.

The procedure described above is repeated iteratively with all unlabeled vectors. Clustering is continued until all the vectors will be specified to any of the subspaces.At the recognition stage with a certain decision rule, the vector closest to one of the subclasses formed in thedescribed manner is considered to belong to the class.

3. Results and discussion

In this paper, the problem of the determining possibility whether there is some given record in the database is examined. After clustering process we can figure out the belonging of a vector to some class. The subclass stores a small number of vectors in comparisonof the initial database.Therefore, after the classification of the current vector, it will be easy to analyze data in a subclass and determine isit possible to add a new value into the database.

Thus, to verify the above assumption, we performed the experiment. From the database of 1041100 records there was performed the random selection of 1040 records five times. Each selection was divided onto 80 subclasses, a subclass consists of 13 vectors. After the clustering procedure, the generated vectors were classified.

The testing vectors were formed using the existed in the dataset records with some modifications. For example, there was simulated situation of incorrect handwritten letters conversion when the personal data was filled in some document. As it was shown in the work [8] the problem of text recognition is a difficult and can have many solutions. Figure 1 a, b presents the images of two handwritten words. The word in Figure 1 a was correct converted by some letters recognition software into “ADAM” as opposed to word in Figure 1 b that led to incorrect result: “ADRM”.

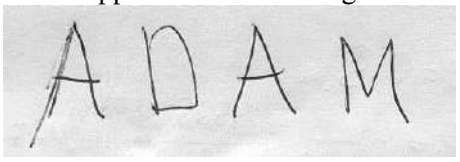
a)  b) 

Figure 1. Examples of a) correct and b) incorrect handwritten letters conversion.

In the classification experiment, 20 vectors of the type described above were tested. All vectors were successfully classified based on the many-to-many approach [9] extending the possibilities of the

binary classification of the support subspaces algorithm [7]. Moreover the average value of conjugation index was 0.95 for true defined class. This fact undoubtedly indicates the reliability of using the conjugation index in problems of this kind. This is an advantage for following research of such kind both with databases of a more complex type, with a larger field number, and for classification using a whole database of one million or more records.

4. References

[1] Yang Y and Guan J 2002 CLOPE: a fast and effective clustering algorithm for transactional data *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining* 682-687
[2] Zhang T, Ramakrishnan R and Livny M 1996 BIRCH: an efficient data clustering method for very large databases *ACM Sigmod Record* **25(2)** 103-114
[3] He Z, Xu X and Deng S 2005 A cluster ensemble method for clustering categorical data *Information Fusion* **6(2)** 143-151
[4] Huang Z 1997 A fast clustering algorithm to cluster very large categorical data sets in data mining *DMKD* **3(8)** 34-39
[5] Fursov V and Kozin N 2007 Recognition through constructing the eigenface classifiers using conjugation indices *IEEE Conference on Advanced Video and Signal Based Surveillance* 465-469
[6] Minaev E and Fursov V 2016 Support subspaces method for fractal images recognition *CEUR Workshop Proceedings* **1638** 379-385
[7] Zherdev D A, Kazanskiy N L and Fursov V A 2015 Object recognition in radar images using conjugation indices and support subspaces *Computer Optics* **39(2)** 255-264 DOI: 10.18287/0134-2452-2015-39-2-255-264
[8] Bolotova Y A, Spitsyn V G and Osina P M 2017 A review of algorithms for text detection in images and videos *Computer Optics* **41(3)** 441-452 DOI: 10.18287/2412-6179-2017-41-3-441-452
[9] Bishop Ch M 2006 *Pattern Recognition and Machine Learning* (New York: Springer) p 738

A fractal statistical fraud detection analysis in databases

P V Hripunov[1]

[1]Pension Fund of the Russian Federation, Shabolovka str. 4, Moscow, Russia, 119991

Abstract. The paper proposes a method for fraud detection analysis in databases. The main idea of the method is to use fractal analysis of numerical information in databases to identify anomalies caused by substitution or distortion of initial data.In this paper, a criterion based on fractal analysis is used. The results of experiments on simulated data are presented.

1. Introduction

The task of identifying fraud in corporate and government databases is one and the most important in data mining. In recent years, a large number of methods for solving this problem have been developed [1]. The main approaches to fraud detection today are: neural networks [2,3], logistic model [4], support vector machine [4], decision trees [5], genetic algorithm [5], text mining [6], Bayesian belief network [7], featureless approach [12] and others. In this paper, a criterion based on fractal analysis is used. This is a fairly new approach for the problem of searching for fraudulent operations in databases. The effectiveness and prospects of this approach is shown in [8,9]. At the same time, the use of fractal methods in other areas, for example, for pattern recognition on images [10], to detect intentional distortions [11], is well developed. In this paper, the methods of fractal analysis for image recognition are adapted to work with large databases.

2. Iterated function systems for database analysis

Classiciterated function systems (IFS) algorithm for images searches the best affine transformation from domain to range block for every range block[10]. As a result, an input image is coded by several affine transformations:

$$\begin{aligned} \mathbf{I}^* &= F\left(\mathbf{I}\right) = \mathbf{C}_{1,4}\mathbf{I} + \mathbf{c}_{5,6}, \\ u^*_{i,j} &= c_7 \cdot u_{i,j} + c_8, \end{aligned} \tag{1}$$

where $\mathbf{I}^* = (i^*, j^*)^T$, $\mathbf{I} = (i, j)^T -$) is the coordinates of pixel from domain and range block accordingly, $\mathbf{C}_{1,4} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}$, $\mathbf{c}_{5,6} = \begin{bmatrix} c_5 \\ c_6 \end{bmatrix}$ – transformation coefficients, $u^*_{i,j}$, $u_{i,j}$ – is the pixel brightness from range and domain area, a c_7, c_8 – contrast and brightness shift parameter.

The transformation is conducted in a class of contraction mapping to obtain a unique and stable fractal image. Parameters of transformations $c_1 - c_8$ are computed by IFS fractal compression algorithm: $c_1 - c_4$ are selected from the possible sets, c_5, c_6 are calculated in the process of searching the best affine transformation from domain to range block, c_7, c_8 – are calculated on the average brightness of domain and range blocks. Specificity analysis of numerical and textual data requires

adaptation of this approach. The main element for the analysis was selected rows of database tables. First, the source data in the database tables often contain heterogeneous information: text, numbers, images, binary data, etc. In our experiments, only text and numeric information was used, and in the preprocessing phase, the text was converted into numbers in accordance with the character encoding table. As a result of preprocessing, each row of the table was represented as a one-dimensional array of numbers. After that, each line of the database is divided into one-dimensional range and domain areas, and by analogy with the formula (1), self-similar data sections are searched. As a result, at the training stage, we select from the database knowingly genuine and correct rows, and form a set of corresponding transformations. At the stage of recognition of fraudulent records in the database, using sets of received transformations, we find the distance:

$$D_i = \frac{d(F_i^* I^*, I^*)}{I_s^*}, \quad (2)$$

where I^* – initial database row, I_s^* – size of initial database row, F_i^* – set of transformations for correct rows, d - Euclidean norm. A distance value significantly greater than the average value for a particular table in the database means that the current row can be fraudulent.

3. Results and discussion

To test the approach described above, we used test database tables describing pension contributions in the corporate enterprise database.An example of the initial data is shown in Figure 1.

	Регистр...	Но...	Акти...	Период	Физлицо	Организация	Период ...
−	Списан...	1	✔	04.06.2010...	Краснов Васи...	ИнвестЗаказ...	01.01.20...
−	Списан...	2	✔	04.06.2010...	Женилов Серг...	ИнвестЗаказ...	01.01.20...
−	Списан...	3	✔	04.06.2010...	Иванов Иван ...	ИнвестЗаказ...	01.01.20...
−	Списан...	4	✔	04.06.2010...	Сидоров Петр ...	ИнвестЗаказ...	01.01.20...
+	Начисл...	1	✔	30.06.2010...	Краснов Васи...	ИнвестЗаказ...	01.06.20...
+	Начисл...	2	✔	30.06.2010...	Женилов Серг...	ИнвестЗаказ...	01.06.20...
+	Начисл...	1	✔	30.06.2010...	Петров Петр И...	ИнвестЗаказ...	01.06.20...
+	Начисл...	1	✔	30.06.2010...	Иванов Иван ...	ИнвестЗаказ...	01.06.20...
+	Начисл...	1	✔	30.06.2010...	Сидоров Петр ...	ИнвестЗаказ...	01.06.20...

Figure 1. Fragment of the database table.

To train the method, 10,000 rows were used from the database.Then, we changed the database, in the first version we added 100 new lines with the correct information, in the second version we added 100 lines with false and fraud information.And counted the distance by formula (2) for both versions.The results of the distribution of distances are shown in Figure 2.

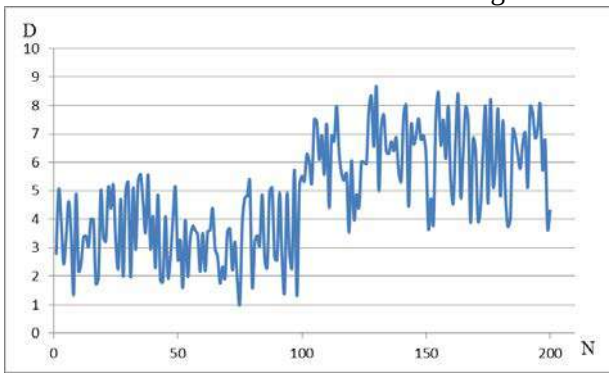


Figure 2. Distribution of distances, N=1…100 – correct rows, N=101…200 – false rows.

As a result of the experiment, it was found that for 73% of false rows, the distance was significantly larger than the correct ones.In future work planned to find the area of applicability of this approach and to conduct experiments on a large sample of data.

4. References

[1] West J and Bhattacharya M 2016 Intelligent financial fraud detection: a comprehensive review *Computers and security* **57** 47-66
[2] Ngai E, Hu Y, Wong Y, Chen Y and Sun X 2011 The application of data mining techniques in financial fraud detection: a classification framework and an academic review of literature *Decision Support Systems* **50** 559-569
[3] Zhang D and Zhou L 2004 Discovering golden nuggets: data mining in financial application *Systems, Man, and Cybernetics, Part C: Applications and Reviews IEEE Transactions* **34** 513-522
[4] Bhattacharyya S, Jha S, Tharakunnel K and Westland J C 2011 Data mining for credit card fraud: a comparative study *Decision Support Systems* **50** 602-613
[5] Ravisankar P, Ravi V, Raghava Rao G and Bose I 2011 Detection of financial statement fraud and feature selection using data mining techniques *Decision Support Systems* **50** 491-500
[6] Humpherys S L, Moffitt K C, Burns M B, Burgoon J K and Felix W F 2011 Identification of fraudulent financial statements using linguistic credibility analysis *Decision Support Systems* **50** 585-594
[7] Kirkos E, Spathis C and Manolopoulos Y 2007 Data mining techniques for the detection of fraudulent financial statements *Expert Systems with Applications* **32** 995-1003
[8] Padua R N and Borres M S 2017 From Fractal Geometry to Statistical Fractal *Recoletos Multidisciplinary Research Journal* **1(1)**
[9] Uy K J D and Zanoria M L E 2017 A Fractal Statistical Analysis of Enron Stock Prices *Recoletos Multidisciplinary Research Journal* **2(2)**
[10] Minaev E Y and Nikonorov A V 2012 Object detection and recognition in the driver assistance system based on the fractal analysis *Computer Optics* **36(1)** 124-130
[11] Ozawa K 2008 Dual fractals *Image and Vision Computing* **26(5)** 622-631
[12] Yumaganov A S and Myasnikov V V 2017 A method of searching for similar code sequences in executable binary files using a featureless approach *Computer Optics* **41(5)** 756-764