Orthogonalization and parameterization of convolutional kernels in machine learning for image and video compression

id21

Ruslan Yuzkiv, Mikhail Gashnikov
Samara National Research University

General information

We study orthogonalization and parametrization of convolutional filters within the framework of the image and video compression method based on machine learning. We use the convolutional filters to interpolate less sparse video frame meshes based on sparser video frame meshes. We consider superresolution neural networks and decision trees as machine learning algorithms at the interpolation stage. Decision trees adaptively select an interpolating function from a predefined set of convolutional filters with parameterized orthogonal weights. We use orthogonalization and parametrization of convolution filter weights to increase the efficiency of the machine learning interpolation algorithm, which in turn leads to an increase in the efficiency of the image and video compression method in general. Computational experiments demonstrate the advantage of the proposed algorithm in natural videos.

The proposed method

The decision tree node chooses an interpolating value through specialized feature P thresholding:

$$I(\tau,m,n) = I_L(m,n)\langle P(m,n) < \tau\rangle + I_R(m,n)\langle P(m,n) \geq \tau\rangle$$

Simplest convolutional filters when interpolating a video frame over a half-decimated video frame:

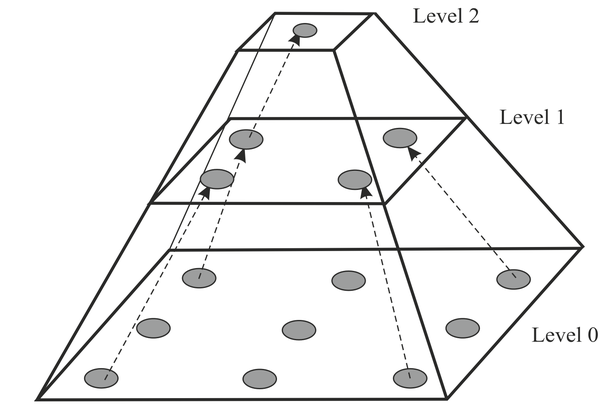
$$\begin{bmatrix} 1/4 & 0 & 1/4 \\ 0 & 0 & 0 \\ 1/4 & 0 & 1/4 \end{bmatrix}$$

An example of a system of two parameterized orthogonal convolutional filters :

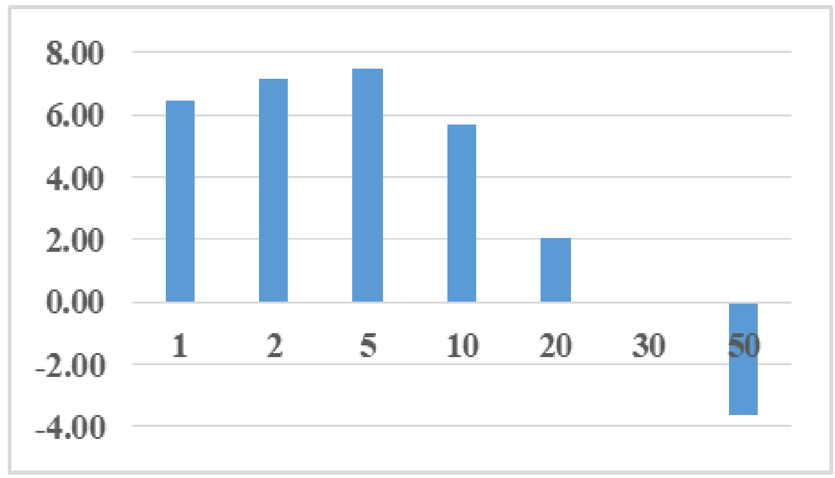
$$\begin{bmatrix} a/2 & 0 & (1-a)/2 \\ 0 & 0 & 0 \\ (1-a)/2 & 0 & a/2 \end{bmatrix}$$

$$\begin{bmatrix} (a-1)/(4a-2) & 0 & a/(4a-2) \\ 0 & 0 & 0 \\ a/(4a-2) & 0 & (a-1)/(4a-2) \end{bmatrix}$$

Video frame as a set of sparse meshes



The experiment



The percentage of gain in archive size due to the use of parameterized orthogonalized convolutional filters depending on the absolute error

ACKNOWLEDGMENT: This work was supported by Russian Science Foundation project 22-21-00662