id58

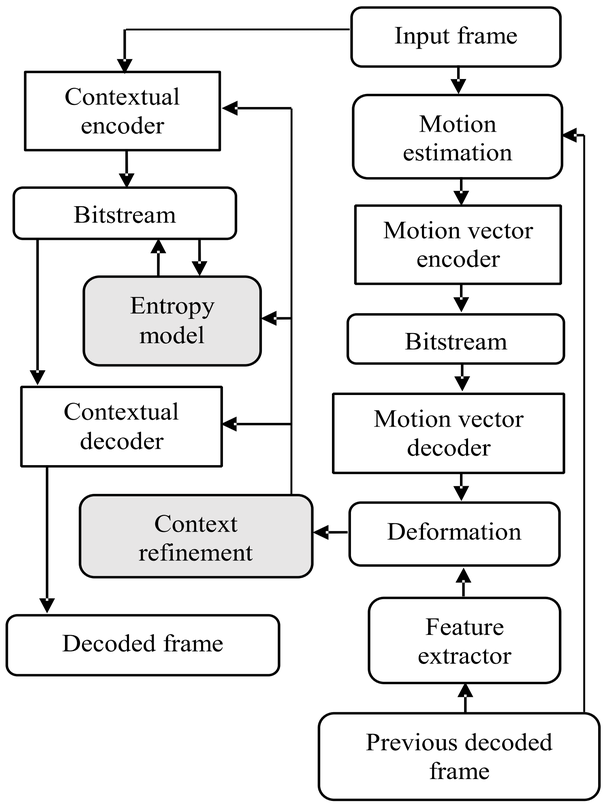
Deep contextual video compression based on machine learning

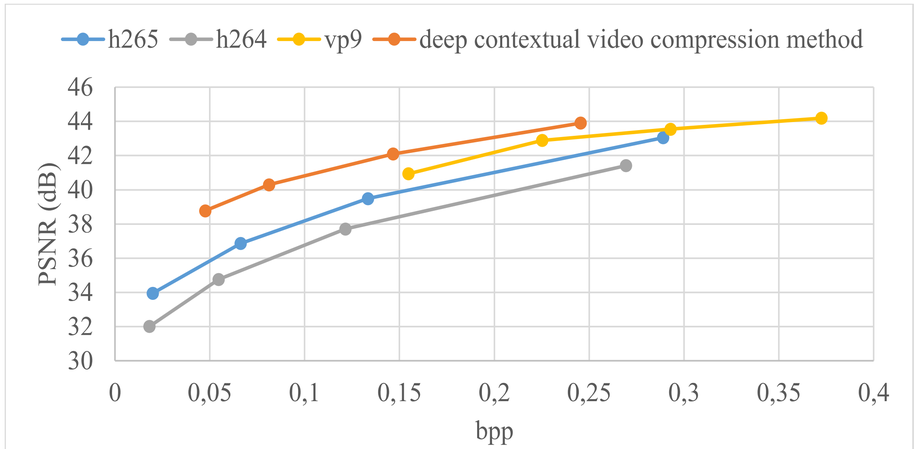
Mariya Chubar, Mikhail Gashnikov
Samara National Research University

General information

The article deals with the conditional coding-based deep contextual video compression method that determines the condition based on learning contextual functions with arbitrary sizes. The flexibility between learning spatial correlation and learning temporal correlation is made possible through conditional coding. The main advantages of the considered architecture of deep contextual compression are the use of conditional encoding when generating a context, its subsequent transmission to the contextual encoder and decoder's input, and the use of an entropy model combining both spatial and temporal priors. We have considered context generation algorithms based on a combination of neural networks instead of directly using the predicted frame. In addition, we also gave a description of the entropy model, including three priors (hierarchical, spatial and temporal). Experimental results prove the advantage of the considered method in PSNR/bpp coordinates on real video sequences in comparison with the efficiency of three standard video codecs — H.264, H.265, VP9.

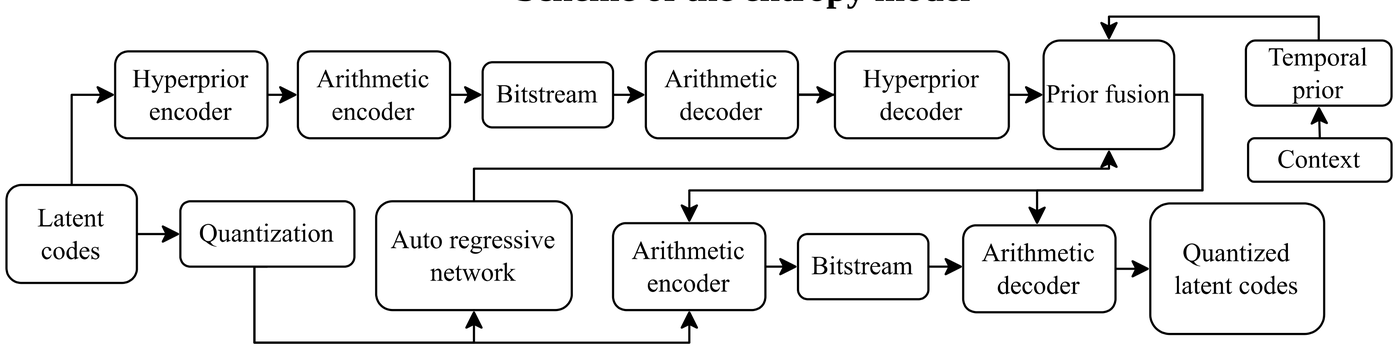
Architecture of the method





The dependence of the efficiency indicator on the degree of compression

Scheme of the entropy model



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