Sound Signal Invariant DAE Neural Network-Based Quantizer Architecture of Audio/Speech Coder Using the Matching Pursuit Algorithm

Vladislav Avramov [1],

Vadzim Herasimovich [2],

Alexander Petrovsky [3]

2018

1, 2, 3 Department of EMU, Belarusian State University of Informatics and Radioelectronics, P. Browki 6, Minsk 220013, Belarus

Keywords: Neural network, Deep autoencoder, Audio/speech coding, Quantization, Stepwise activation function.

Abstract.

The paper is devoted to the quantization algorithm development based on the neural networks framework. This research is considered in the context of the scalable real-time audio/speech coder based on the perceptually adaptive matching pursuit algorithm. The encoder parameterizes the input sound signal frame with some amount of real numbers that are need to be compactly represented in binary form, i.e. quantized. The neural network quantization approach gives great opportunity for such a goal because the data quantized in whole vector but not in separate form and it can effectively use correlations between each element of the input coded vector. Deep autoencoder (DAE) neural network-based architecture for the quantization part of the encoding

algorithm is shown. Its structure and learning features are described. Conducted experiments points out the big compression ratio with high reconstructed signal quality of the developed audio/speech coder quantization scheme.

Article published in: Advances in Neural Networks – ISNN 2018. Lecture Notes in Computer Science. – Springer, Cham, 2018. – Vol. 10878. – P. 511-520. – DOI: https://doi.org/10.1007/978-3-319-92537-0_59.

Internet link to the article:

https://link.springer.com/chapter/10.1007/978-3-319-92537-0_59.