Deep Multiple Description Coding by Learning Scalar Quantization

Lijun Zhao*, Huihui Bai*, Anhong Wang†, and Yao Zhao*

*Beijing Jiaotong University
Institute of Information Science
Beijing, 100044, P. R. China
hhbai@bjtu.edu.cn

†Taiyuan University of Science and Technology
Institute of Digital Media & Communication
Taiyuan, 030024, P. R. China
wah_ty@163.com

Unlike multiple description quantizer and correlating transform-based multiple description coding (MDC) framework, sampling-based multiple description coding is more flexible to be compatible with standard coders. But most of sampling-based MDC methods always manually design a specific sampling method or extend the existing sampling methods. Recently, a convolutional neural network based MDC method [1] tries to adaptively sample the input image to create multiple descriptions. But, the coding efficiency of these approaches cannot satisfy the practical applications. In summary, we should study the topic of multiple description coding, especially at low bit-rates.

In this paper, we propose a deep multiple description coding framework, whose quantizers are adaptively learned via the minimization of multiple description compressive loss. Firstly, our framework is built upon auto-encoder networks, which have multiple description multi-scale dilated encoder network and multiple description decoder networks. Secondly, two entropy estimation networks are learned to estimate the informative amounts of the quantized tensors, which can further supervise the learning of multiple description encoder network to represent the input image delicately. Thirdly, a pair of scalar quantizers accompanied by two importance-indicator maps is automatically learned in an end-to-end self-supervised way. The hard scalar quantizers is used in the forward propagation, but they are replaced by soft quantizations during gradient back-propagation. Finally, multiple description structural dissimilarity distance loss is imposed on multiple description decoded images in pixel domain for diversified multiple description generations rather than on feature tensors in feature domain, in addition to multiple description reconstruction loss. Through testing on Set4 from [1] and the commonly used Kodak PhotoCD database, it has been verified that our method has better coding efficiency than "tcsvt18" [1] and "tip14" [2] regarding image structural similarity.

[1] L. Zhao and H. Bai, etc, Multiple description convolutional neural networks for image compression, IEEE Transactions on Circuits and Systems for Video Technology, 2018.

[2] L. Meng and J. Liang, etc., Multiple description coding with randomly and uniformly offset quantizers, IEEE Transactions on Image Processing, 23(2), 582-595, 2014.

This work was supported in part by the Fundamental Research Funds for the Central Universities (No.2017YJS053, 2018JBZ001) and Key Innovation Team of Shanxi 1331 Project (KITSX1331).