Approach of Team HHC: A Method for Low Bitrate Image Compression

Chenhao Wu
University of Electronic Science and Technology of China
chwu@std.uestc.edu.cn

Abstract

This paper elaborates on the technical scheme of team HHC in CLIC2021 image track. Considering the target rates in this challenge are relatively low, so a lot of detailed features should be removed, the feature extraction part of the proposed method adopts a more compact space for feature representation. Then a decoder-side enhancement module is designed to help improve the reconstruction quality of the decoded image. Besides, a dual hyperprior entropy model is applied to estimate the probability distribution for encoding/decoding the latent feature, through encoding the mean and scale of Gaussian distributions. This paper introduces the proposed method in 4 parts, including feature extraction, decoder-side enhancement, entropy coding, and training details.

1. Feature Extraction

The proposed method builds on the typical auto-encoder architecture following [1, 3], obtains the compact representation of image feature through stacked convolutional layers with stride 2. Compared to the common 4 layers structure with 16 times downsample rate, a more compact 32 times downsample operation is adopted in the proposed method, based on the consideration that many detail features are omitted in the range of low-bit rates. Feature space with lower resolution can save bits without losing information massively when dealing with relatively smoother features. Specifically, 5 convolutional layers are applied in the proposed approach, with the first 4 of them followed by Generalized Divisive Normalization (GDN) activation layers.

Correspondingly, 5 transpose convolutional layers with Inverse Generalized Divisive Normalization (IGDN) activation layers are applied at the decoder side.

2. Decoder-side Enhancement

The smaller spatial scale of representation space in the feature extraction module will inevitably affect the reconstruction of the decoded image. In order to compensate for lost information and further improve the image quality, a decoder-side enhancement module is adopted to model the refined information omitted by the encoder based on existing feature maps. To be more specific, the enhancement module is utilized to estimate the residual between reconstruction and input image, with both coarse reconstructed image and reconstructed feature as input. Four residual blocks [2] with LeakyReLU layers are utilized in the enhancement module.

3. Entropy Coding

Following the approach of NIPS'18 [3], the proposed method estimates the probability distribution of latent feature map through encoding the mean and scale parameters of Gaussian distributions. To reduce the decoding complexity as far as possible, the context information is not introduced. In practice, two hyperprior networks are utilized to estimate and map these two parameters to hyperprior encoding space respectively and share them between encoder and decoder through entropy coding. When encoding/decoding latent feature map, the encoder/decoder will map the hyperprior information back to feature space, and construct the probability distribution of feature map to carry out entropy coding. Code implementation of entropy coding follows https://github.com/tensorflow/compression.

4. Training Details

Uniform quantization is adopted in the proposed method, which is realized through rounding operation during the actual coding process. While in training, uniform noise is added on the latent feature to simulate the quantize error, through confusing the information within the quantization step, and maintain the gradient at the same time. The training batch size is set to 8, and training images are cropped into patches with the size of 256×256, training process lasts for one million steps.

References

[1] Johannes Ballé, Valero Laparra, and Eero P Simoncelli. End-to-end optimized image compression. In *5th International Conference on Learning Representations, ICLR 2017*, 2017. 1

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 1

[3] David Minnen, Johannes Ballé, and George Toderici. Joint autoregressive and hierarchical priors for learned image compression. In *NeurIPS*, 2018. 1