# Asymmetric Gained Deep Image Compression With Continuous Rate Adaptation

## **1. Additional Implemental Details**

## 1.1. Complexity And Memory Consumption

Parameter and computational quantity are important indexes of whether the learned image compression methods can be popularized and applied. The proposed variable-rate HCVR method and CVR method avoid the multiplication of model memory and only introduce the rarest complexity and memory consumption. Compared with the corresponding fixed-rate models, the additional parameters and computation brought by the CVR method are described as:

$$Para_{CVR} = c \times n \times 2,$$
  
FLOPs<sub>CVR</sub> = c × h × w × 2, (1)

where  $Para_{CVR}$  and  $FLOPs_{CVR}$  represent the additional parameters and computation with gain units of the CVR method respectively. c, h, w represent the number of channels, height and width of the latent representation respectively, and n is the number of the gain vectors used in the gain units of the CVR method (n is usually set to 6 in our AG-VAE framework). Compared with the corresponding fixed-rate models, the additional parameters and computation amount brought by the HCVR method are depicted as:

$$Para_{HCVR} = c \times n \times 2 + c_{hp} \times n_{hp} \times 2,$$
  
FLOPs<sub>HCVR</sub> =  $c \times h \times w \times 2 + c_{hp} \times h_{hp} \times w_{hp} \times 2.$  (2)

where  $Para_{HCVR}$  and  $FLOPs_{HCVR}$  represent the additional parameters and computation with gain units of the HCVR method respectively.  $c_{hp}, h_{hp}, w_{hp}$  represent the number of channels, height and width of the hyperprior latent representation respectively, and  $n_{hp}$  is the number of the gain vectors used in the gain units of the HCVR method  $(n_{hp}$  is usually set to 6 in our AG-VAE framework). Based on the single fix-rate VAE-based image compression model, the proposed CVR method or the HCVR method can utilize the trivial additional parameters and computation to endow the fixed-rate model with continuous rate adaptation while avoiding performance degradation.

#### **1.2.** Parameter Selection

In training, the channel importance guides the gain units to obtain proper gain values for different channels, which



Figure 1. The gain vectors for different rate-distortion tradeoff obey similar distribution alongside the channel.

achieves reasonable rate allocation. As shown Figure 1, different gain vectors obey similar distribution alongside the channel because of the same channel importance variation. The number of the gain vectors can be selected flexibly according to your requirements. In our experience, n and  $n_{hp}$ should be set 6 for the long bit rate range (such as  $0 \sim 1.2$ bpp) and 3 for the short bit rate range (such as  $0 \sim 0.2$  bpp) respectively to ensure the rate-distortion performance of the CVR method or the HCVR method.

#### **1.3.** Asymmetry Gaussian Entropy

Compared with the symmetry Gaussian entropy model, the asymmetry Gaussian entropy model obtains higher freedom of parameters, achieving more flexible entropy estimation for natural images. As shown in Figure 2, the asymmetry Gaussian entropy needs fewer bits to encode the selected channel. And a marked difference of  $\sigma_l^2$  and  $\sigma_r^2$  distribution demonstrates that the entropy estimation model for the image should not degrade to the symmetric Gaussian model.

## 2. Visualization

In this section, we compare more images from the Kodak dataset reconstructed by the proposed AG-VAE (optimized for the MS-SSIM or MSE) and the classical image compression codecs, including BPG and VTM. To guarantee the fairness of visual comparison, we adapt the rate of images reconstructed by the AG-VAE to be consistent with the counterparts reconstructed by the SOTA classical image compression codecs VTM. From Figure 5 and Figure 3, we can conclude that the AG-VAE optimized for MS-SSIM or MSE can outperform the classical image compression codecs VTM and BPG in MS-SSIM at the same BPP. Besides, the proposed AG-VAE optimized by MSE or MS-



Figure 2. Visualization of asymmetry and symmetry Gaussian entropy models for the channel with the highest entropy, which utilizes the *kodim*21 from Kodak dataset with approximately 0.5bpp.

SSIM can recover more details and provide better visual quality than the classical image compression codecs for different images. To highlight the visible differences, we box and cut the same areas of these reconstructed images, which are displayed in Figure 6 and Figure 4.



**Original Image** 

AG-VAE [MS-SSIM] 0.230bpp, 25.23dB, 0.964





VTM 0.230bpp, 27.57dB, 0.932

BPG 0.248bpp, 26.90dB, 0.922

Figure 3. PSNR, MSSSIM, and visualization comparison of the proposed AG-VAE (optimized for the MS-SSIM or MSE) and classical image compression codecs(BPG, VTM) for the kodim24 from Kodak dataset with approximately 0.230 bpp. The proposed AG-VAE optimized for MS-SSIM or MSE outperform the BPG and VTM in MS-SSIM. We box and cut the same areas of these reconstructed images to highlight the visual comparison.



Figure 4. Visualization comparison of the proposed AG-VAE (optimized for the MS-SSIM or MSE) and classical image compression codecs(BPG, VTM) for the kodim24 from Kodak dataset in the boxed areas of Figure 3.



**Original Image** 

AG-VAE [MS-SSIM] 0.449bpp, 22.80dB, 0.965



VTM 0.449bpp, 25.82dB, 0.923

BPG 0.480bpp, 25.35dB, 0.914

Figure 5. PSNR, MSSSIM, and visualization comparison of the proposed AG-VAE (optimized for the MS-SSIM or MSE) and classical image compression codecs(BPG, VTM) for the kodim13 from Kodak dataset with approximately 0.449 bpp. The proposed AG-VAE optimized for MS-SSIM or MSE outperform the BPG and VTM in MS-SSIM. We box and cut the same areas of these reconstructed images to highlight the visual comparison.



Figure 6. Visualization comparison of the proposed AG-VAE (optimized for the MS-SSIM or MSE) and classical image compression codecs(BPG, VTM) for the kodim13 from Kodak dataset in the boxed areas of Figure 5.