Learned Image Compression via Sparse Attention and Adaptive Frequency

Supplementary Material

6. Details of the Proposed Module

6.1. Cross-Sparse Window Attention

We illustrate the detailed pipeline for CSWA in Fig. 6. For the specific procedure, please refer to Section 3.2.

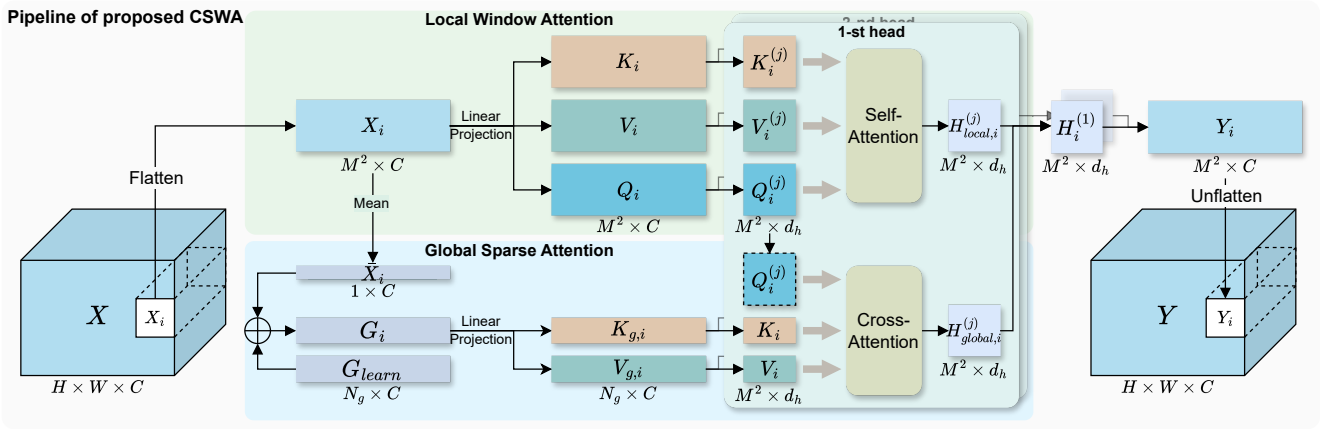

Figure 6. Detailed pipeline of the proposed Cross-Sparse Window Attention.

6.2. Decomposition Weight Generator

The structure of DWG is shown in Fig.7. We use simple convolutions, GELU [20], and Softmax [7] to simulate 4 different frequency bands.

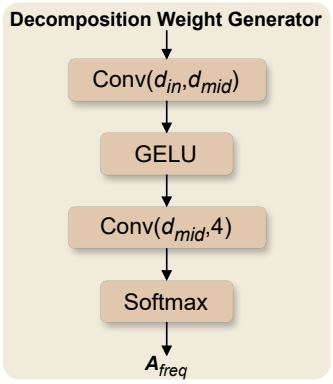

Figure 7. Structure of the proposed Decomposition Weight Generator.

7. Additional experimental results

7.1. Effects of the Proposed Modules

To further investigate the performance impact of the proposed modules, we visualized the training process ($\lambda = 0.05$). Fig. 8 (a) plots the training loss over the first 50 epochs. Fig. 8 (b) illustrates the corresponding RD performance evolution, plotting PSNR against bpp as the loss decreases. As observed in Fig. 8 (a), the models incorporating the proposed modules (e.g., SAB,

AFB, DaR) exhibit a lower overall loss from the beginning compared to the BASE. Notably, the introduction of the adaptive frequency significantly accelerates the convergence speed. This highlights the importance of frequency-domain features and confirms the effectiveness of our proposed modules. Fig. 8 (b) provides a more intuitive view of the RD performance trajectory. As the epochs increase, the models enhanced with the SAB and DaR modules consistently operate in the top-left region relative to the BASE, indicating superior RD performance (achieving higher PSNR at lower bpp). Generally, the RD path trends towards the top-right initially before optimizing steadily towards the top-left. Although the model with AFB starts at a higher bpp, its optimization path moves consistently toward the top-left, further demonstrating that frequency-domain features can accelerate convergence towards a better RD trade-off.

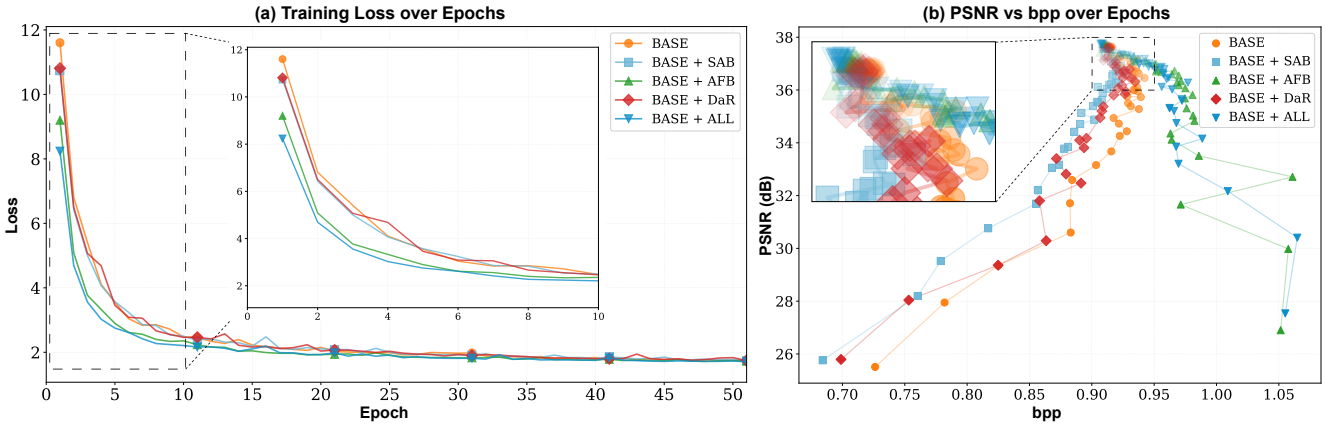

Figure 8. Evolution of loss and RD performance during the training process.

7.2. Additional Details in Reconstructed Images

We additionally compare the reconstruction details of our algorithm and the baselines on Kodak, as shown in Fig. 9. SAAF demonstrates superior overall reconstruction quality and achieves the optimal PSNR/bpp metrics.

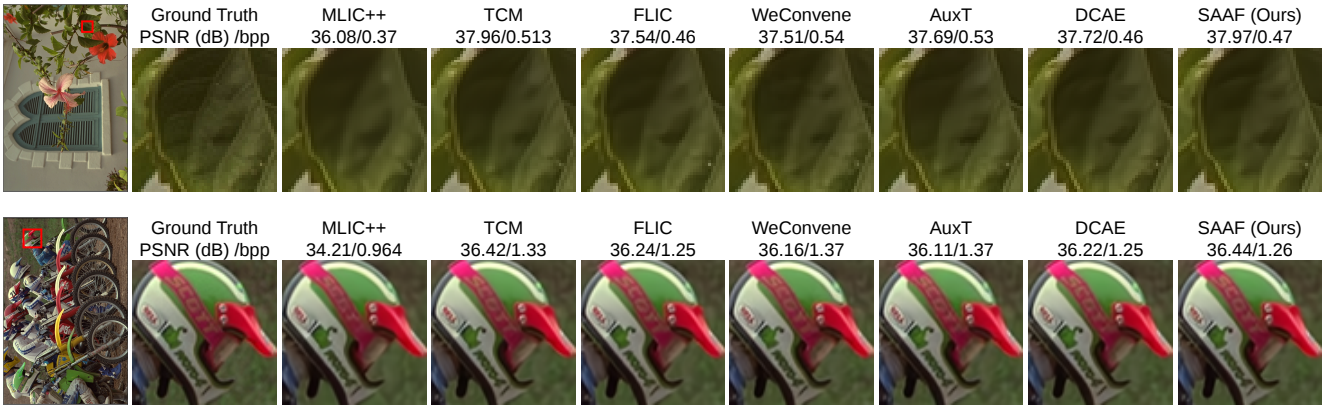

Figure 9. Visual comparison of reconstruction quality.