2022 IEEE International Conference on Acoustics, Speech and Signal Processing

icassp 2022 Singapore

Universal Efficient Variable-rate Neural Image Compression

Shanzhi Yin, Chao Li, Youneng Bao, Yongsheng Liang, Fanyang Meng, Wei Liu

Harbin Institute of Technology, Shenzhen & PengCheng Laboratory

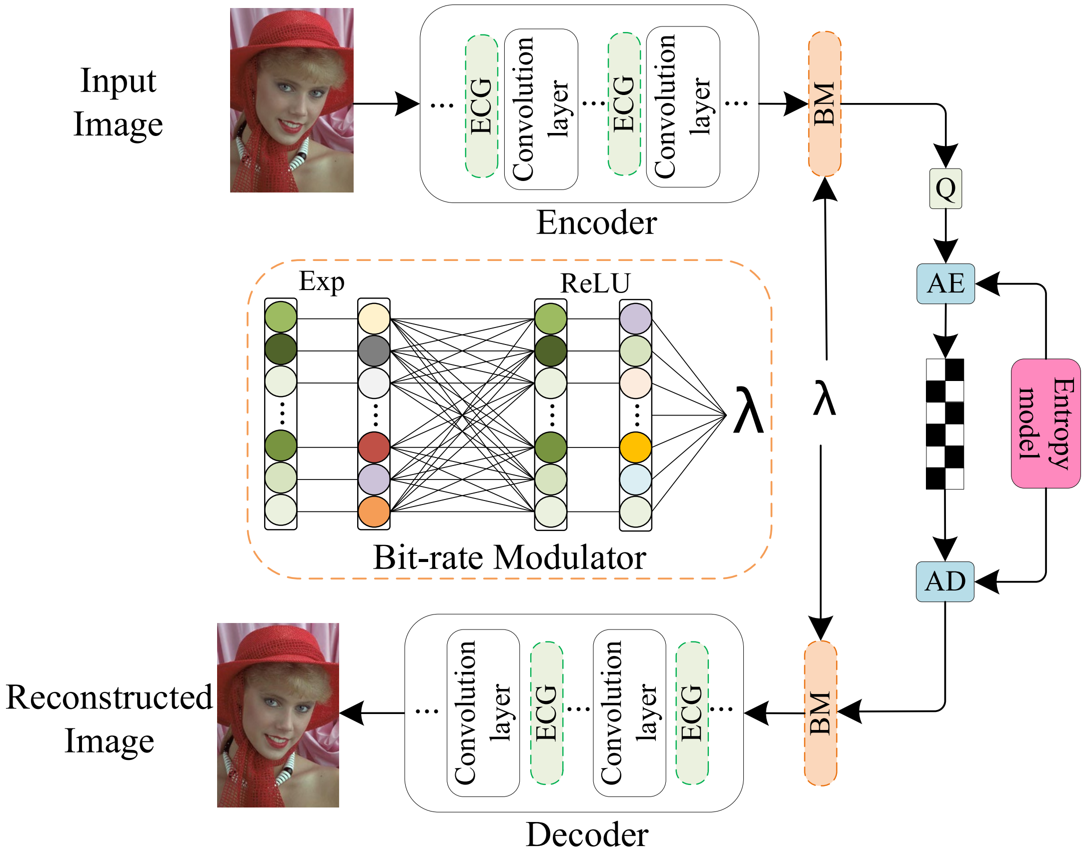
1. Introduction

- **Image compression** is a fundamental technology in signal processing and computer vision.
- In recent years, **many learning-based image compression** methods have achieved **state-of-the-art** performance comparing to traditional image codecs.
- However, there are still some challenges for its **practical deployment**:
  - Bit-rate and reconstruction quality are **fixed** for a single trained model with a predefined trade-off factor.
  - **Computational cost** in learning-based compression models is relatively high due to their complex network architectures.

To deal with such situation and corresponding challenges of learning-based image compression, this paper proposes a **universal variable-rate efficient** method for neural image compression.

2. Method

Overall Framework



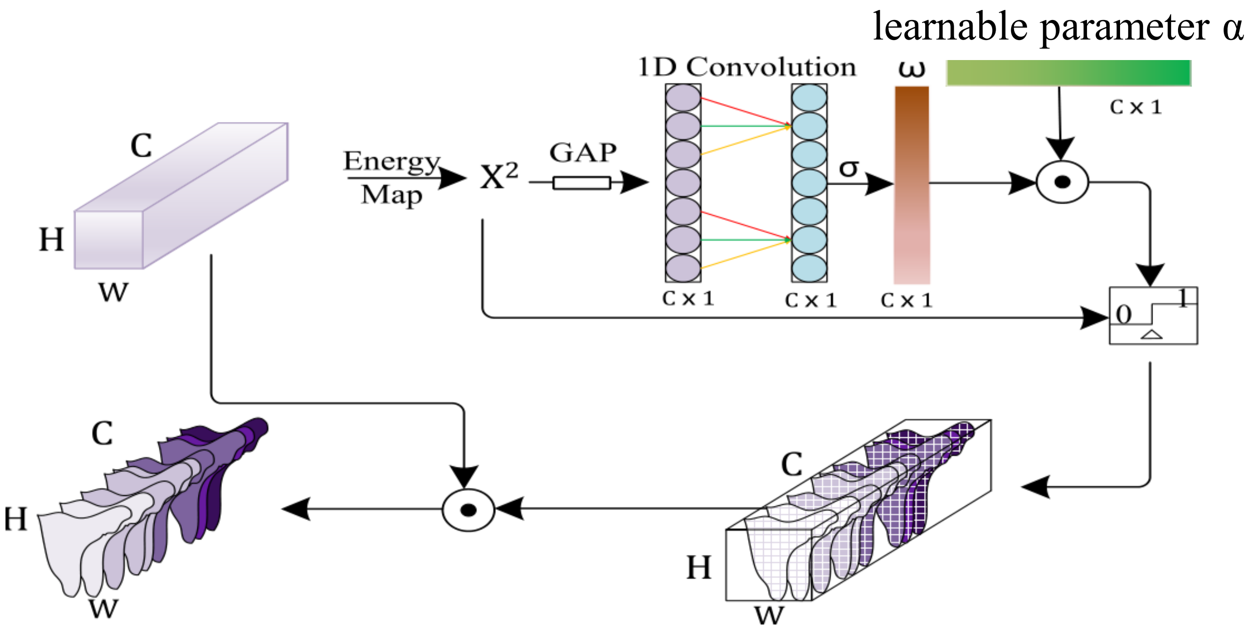
Two novel modules are purposed——**Energy-based channel gating module**(ECG) and **Bit-rate modulator**(BM).

- **ECG** embedded before each convolution layers to get sparse convolutional inputs.
- **BM** inserted outside entropy coding process to modulate the latent representation.
- **Comprehensive** optimization formulation:

$$\underset{\theta,\phi,\xi,\lambda}{\operatorname{argmin}} \sum_{\lambda\in\Lambda}[R+\lambda D+\gamma\sum_{i=1}^{n}(\alpha_n-\alpha_t)^2]$$

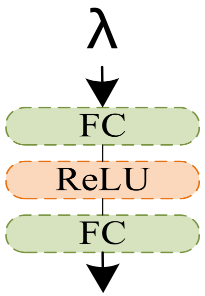
2. Method

Energy-based channel gating



- Inputs of **different intensities** cause different influence on the results → learnable dynamic feature map pruning with **channel-wise threshold**.
- Global pooling for **intra-channel** information
- 1D-convolution for **inter-channel** information

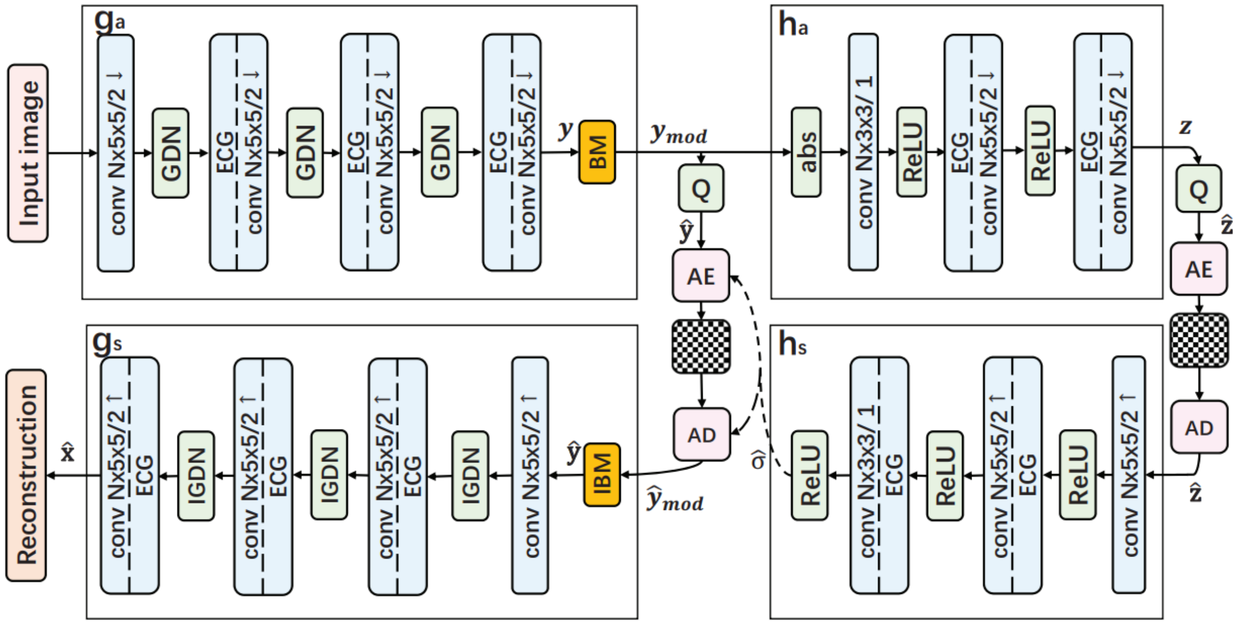
Bit-rate modulator



- Mapping a trade-off factor λ into a vector
- Simple: Two full-connected layer
- Effective: Plug-in manner

3. Experiments

We use ScaleHyperprior model as an example to show the implementation details and optimization strategies of our method.



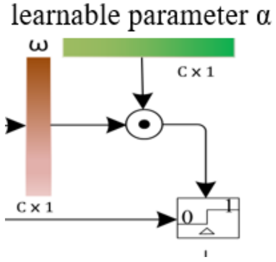
Distortion: mean square error measured on the test set

$$D(x,\hat{x};\theta,\phi,\xi,\lambda) = \mathbb{E}_{x\sim p_x}[||x-\hat{x}||^2]$$

Rate: cross entropy of the estimated distribution of y and the its actual distribution

$$R(\hat{y};\theta,\phi,\xi,\lambda) = \mathbb{E}_{\hat{y}\sim p_y}\{log_2 q_y[Q(y\odot bm(\lambda))]\}$$

In ECG, the **learnable adjustment vector α** affect the final gating threshold *th*, **larger** the α is, **higher** the final threshold on each channels will be, and the output feature map of ECG will be **sparser**.



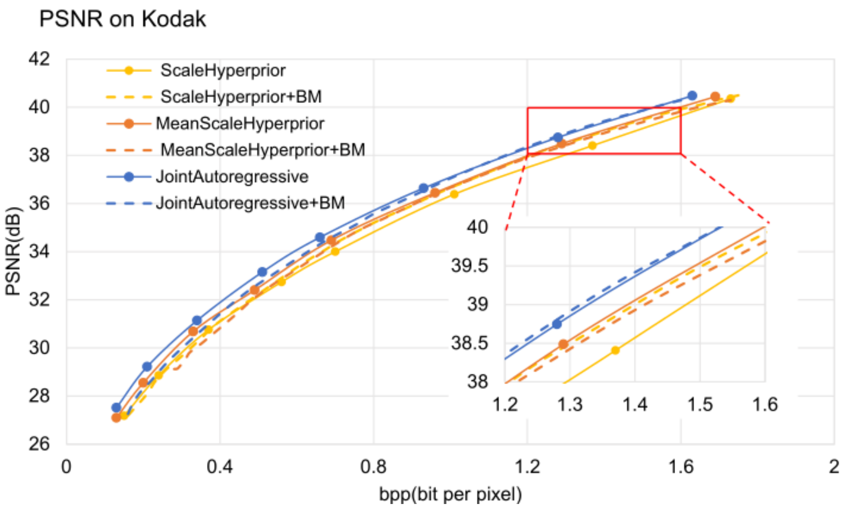
Final optimization formulation:

$$\underset{\theta,\phi,\xi,\lambda}{\text{argmin}} \sum_{\lambda\in\Lambda}[R+\lambda D+\gamma\sum_{i=1}^{n}(\alpha_n-\alpha_t)^2]$$

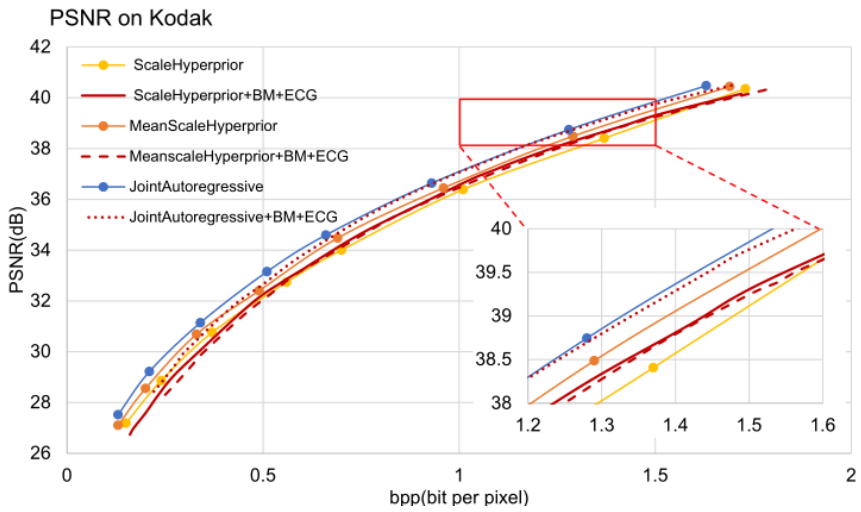
3. Experiments

Model	Performance	Quality							
		1	2	3	4	5	6	7	8
ScaleHyperprior	PSNR drop(%)	0	0.346	0.228	0.269	0.336	0	0	0.216
	FLOP reduction	**2.54**×	**2.86**×	**2.60**×	**2.54**×	**2.54**×	**2.07**×	**2.14**×	**2.03**×
MeanscaleHyperprior	PSNR drop(%)	0.39	0.22	0.77	0.69	0.37	0.61	0.71	0.78
	FLOP reduction	**2.34**×	**2.50**×	**2.56** ×	**2.68**×	**2.33**×	**2.12**×	**2.12**×	**2.24**×
JointAutoregressive	PSNR drop(%)	0.207	0.335	0.437	0.807	0.465	0.150	0.354	0.553
	FLOP reduction	**2.43**×	**2.67**×	**2.48**×	**2.23**×	**2.29**×	**2.02**×	**2.06**×	**2.02**×

For model with ECG: We can see that the FLOP reduction of more than 2× can be achieved in three neural image compression models with very slight PSNR degradation around 0.5% and no more than 1%



For models with BM: continuous rate flexibility can be achieved.



For efficient models: Comparable performance to original models.

Sparsity around 0.5 in convolution operations

Storage saving of 80.42%,82.04% and 83.07% respectively.

2022 IEEE International Conference on Acoustics, Speech and Signal Processing

icassp 2022 Singapore

Thank you!

Harbin Institute of Technology, Shenzhen & PengCheng Laboratory