A GAN-based Tunable Image Compression System

Lirong Wu

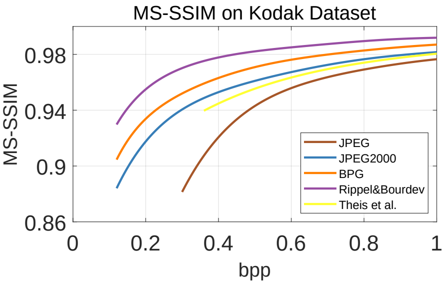
wulirong@zju.edu.cn

College of Information Science and Electronic Engineering
Zhejiang University, Hangzhou, China

March 4, 2020

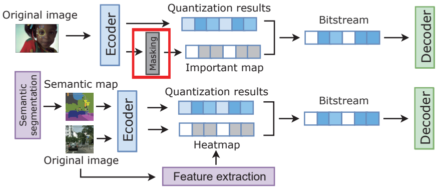
Introduction

- Efficient image compression is significant for the storage, transmission, and processing of image information.
- In recent years, remarkable achievements with DNN-based image compression have been made. However, their compression performance often dramatically drops at low bpp.
- Rethink content-based image compression under the GAN setting.

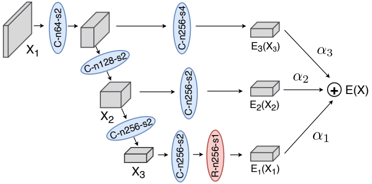


Methods

- We design a simple network (Masking) to identify the important regions of the image and guide the allocation of bits.
- We use the multiscale structure not only in the encoder but also in the discriminator. The symmetrical multiscale structure makes it more adaptable to different sizes of objects.



(a) Top: Ours, Bottom: Others



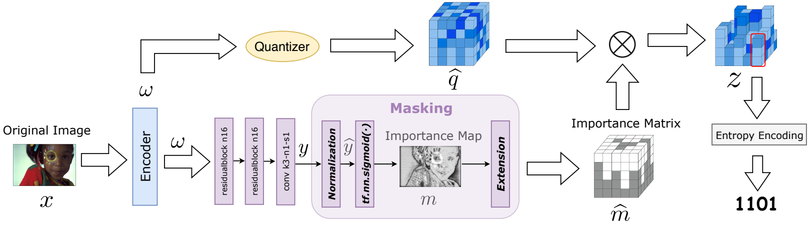
(b) Multiscale design

Methods

- The bit allocation according to the importance of image contents is achieved by constructing a masker.
- The formula of the normalization procedure is specified as

$$\widehat{y}_{i,j} = \frac{y_{i,j} - (\overline{\mu} + n)}{\overline{\sigma}} \tag{1}$$

- Tunable characteristic: We can reassign the user-defined parameter n to achieve different compression ratios without retraining the model.



Methods

- Adversarial Loss

$$\mathcal{L}_A = \sum_{i=1}^{m} \beta_i \Big\{ \mathbb{E}[logD_i(x)] + \mathbb{E}\big[log(1 - D_i(G(x)))\big] \Big\} \tag{2}$$

- Distortion Loss

$$\mathcal{L}_D = E[d(x, \widehat{x})] \tag{3}$$

- Overall Loss

$$\begin{aligned}\mathcal{L}_{G,D,E,B} = \frac{1}{B} \sum_{j=1}^{B} \Big\langle \eta \sum_{i=1}^{m} \beta_i \Big\{ \mathbb{E}\big[log(1 - D_i(G(x^j)))\big] \\ + \mathbb{E}[logD_i(x^j)] \Big\} + \kappa E\big[d(x^j, \widehat{x}^j)\big] \Big\rangle\end{aligned} \tag{4}$$

Results: Qantitative Results

- Our proposed method improves MS-SSIM by more than 10.3% compared to the recently reported GAN-based method [1] to achieve the same low bpp (0.05) on the Kodak dataset.
- Our method preforms much better at low bpp than at high bpp.

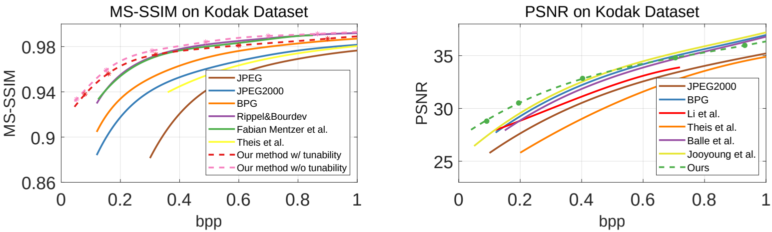


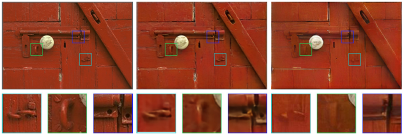
Figure: Qantitative Comparision

[1] Eirikur Agustsson, Michael Tschannen, Fabian Mentzer and Luc Van Gool. Generative adversarial networks for extreme learned image compression. arXiv preprint arXiv:1804.02958, 2018.

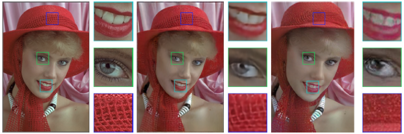
Results: Qualitative Results

- The details of the image, such as the window of the house, the lock on the door, the holes in the women's hat and the fuselage and paddles of the aircraft are well preserved due to the use of important map.

(a) *bpp* / MS-SSIM: Ours 0.039 / 0.927, Agustsson et al. 0.030 / 0.824



(b) *bpp* / MS-SSIM: Ours 0.063 / 0.906, Agustsson et al. 0.069 / 0.795



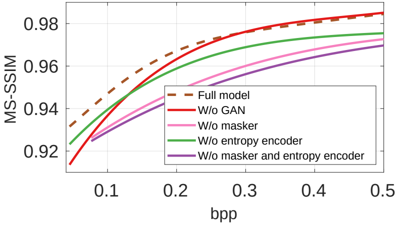
(c) *bpp* / MS-SSIM: Ours 0.058 / 0.921, Agustsson et al. 0.065 / 0.845

(d) *bpp* / MS-SSIM: Ours 0.040 / 0.937, Agustsson et al. 0.034 / 0.844

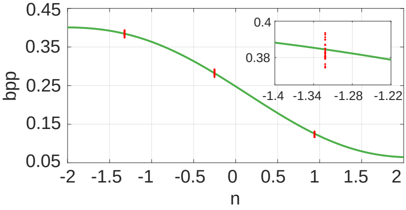
Figure: Visual comparison. From left to right: Original, Ours, Agustsson et al.

Results: Ablation and Tunability Analysis

- The performance of the model with GAN performs better than that of the model without GAN at low bpp.
- GAN's help with image compression is more pronounced at low bpp.
- The compression ratio of the image is determined by the parameter n, which is an intuitive and simple dependency.



(a) Ablation analysis



(b) Tunability analysis

Future work

- Better schemes to capture important regions
  - Attention
- Tunability with better performance
  - Wider scope
  - Stronger generalizability
- Finding a balance between performance and efficiency
  - Feature sharing
  - Pruning and acceleration
- Content-based video compression