DSSLIC: Deep Semantic Segmentation-based Layered Image Compression

Mohammad Akbari, Jingning Han
Google Inc.
jingning@google.com
akbari@sfu.ca, jiel@sfu.ca

Experimental Results

Introduction

o Deep learning in Computer Vision
  • Image classification, face recognition
  • Semantic Segmentation (image to map) [1]

Proposed Approach

o Generative Adversarial Networks (GANs)
  • Competition of Discriminator and Generator
  \[ V(D, G) = \sum_{i=1}^{n} \log(D_i(x_i)) + \sum_{i=1}^{n} \log(1 - D_i(G(z_i))) \]

o GAN-based high-quality map to image [1]

o Learning-based Image Compression (IC) [2-5]
  • Exploit image features using neural networks
  • Learn analysis and synthesis transforms
  • Better results than standard codecs with hand-crafted components (e.g., JPEG & BPG)

o Previous works:
  • GAN-based autoencoder for IC [2]
  • Soft-to-hard vector quantization approach [3]
  • GDN-based analysis/synthesis transform [4]
  • Semantic map-based IC for low bit rates [5]

Conclusion

o Deep learning-based image compression method proposed using semantic segmentation maps.

o Objective Functions:
  • Pixel-wise losses: \[ L_1 = 2 \| x - x' \|_1, \quad L_{SISSE} = -1(x, \hat{x}) \cdot C(x, \hat{x}) \cdot S(x, \hat{x}) \]
  • Perceptual losses: \[ L_{DSS} = \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \beta_j \| D_i(x_i) - D_j(x_j) \|_1 \]
  • Adversarial training:
    \[ L_G = -\sum_{i=1}^{N} \log(D_i(G(z_i))) + \log(1 - D_i(x_i)), \quad L_D = \sum_{i=1}^{N} \log(1 - D_i(x_i)) + \sum_{i=1}^{N} \log(D_i(G(z_i))) \]

Future works:
  • YUV-based image coding
  • Object-based adaptive image compression
  • Deep scalable image compression

References