Abstract

Very recently, deep generative models have been used to optimize or replace some image coding, with very promising results. However, so far no systematic and independent study of the coding performance of these algorithms has been carried out. In this paper, we conduct a subjective evaluation of two recent deep-learning-based image compression algorithms, comparing them to JPEG 2000 and to the recent BPG image codec based on HEVC Intra. We found that compression approaches based on deep auto-encoders can achieve coding performance higher than JPEG 2000, and sometimes as good as BPG. We also show experimentally that the PSNR metric is to be avoided when evaluating the visual quality of deep-learning-based methods, as their artifacts have different characteristics from those of DCT or wavelet-based coders. In particular, images compressed at low bitrate appear more natural than JPEG 2000 coded pictures, according to a no-reference naturalness measure.

Context and motivation

- Traditional image coding methods mainly based on linear signal models (DCT, wavelet transform, linear prediction)
- Recently proposed image compression algorithms based on deep generative models such as (variational) auto-encoders or generative adversarial networks
- Goal: assess visual quality of images compressed with these methods

Subjective Experiment

- Compared methods:
  - Ballé et al. (2016) [1]: End-to-end image compression using a variational auto-encoder
  - Toderici et al. (2016) [2]: Variable bitrate image compression using recurrent auto-encoders
  - JPEG 2000
  - BPG (HEVC Intra)
- For Ballé et al., bitrate estimated by implementing a simple run-length + entropy encoding
- Test material: 6 uncompressed images of 736x960 pixels, total of 113 compressed stimuli
- 23 participants, no outliers detected
- Method: DSIS with continuous scale

Objective Evaluation

Fidelity metrics

- For some metrics (especially PSNR), drop in accuracy with DL methods

Qualitative results

[MOS vs PSNR]( attachment:mos_vs_psnr.png)

Naturalness

MOS = 12.4
PSNR = 20.85 dB
MOS = 8.1
PSNR = 21.35 dB

Qualitative results

[ballé2016]( attachment:ballé2016.png)
[ballé2016]( attachment:ballé2016.png)
[toderici2016]( attachment:toderici2016.png)


MMSP 2018, Vancouver, Canada