285 Project

CaoHaoyuan

Reimplementation and rethinking of image lossless compression through super-resolution.

Background:

Paper *Lossless Image Compression through Super-Resolution* proposed an efficient while straight forward way to compress the image, by storing the low-resolution(LR) image and utilizing the super-resolution to reconstruct the original image. The overall algorithm is summarized as follow:

- Image is iteratively average pooled(2*2) for three times. Every time, the probability of original pixels conditioned on compressed pixel is captured and original image is entropy coded. Eventually, we have three level entropy-coded image. The lowest one is simply the raw image, while the other two and coded based on conditional probability.
- The probability of pixels conditioned on lower level pixels is the mixture of logistic functions following the the model proposed in PixelCNN++.

Using super-resolution for image/video compression is not a new topic. Before the DNN gets popular, numbers of related papers have been published. Lossless compression using arithmetic encoding in the entropy coding step requires high-quality probabilistic model of the image. Natural image modeling was, however, restricted to very small images until the PixelCNN models were proposed in 2016. PixelCNN is further refined in 2017 by introducing the auxiliary variables including gray-scale images or low resolution images.

What I want to do

The paper *Lossless Image Compression through Super-Resolution* validates that using the current state-of-art probabilistic image modeling in the compression can reach great performance. Conditional probability is captured using PixelCNNs and utilized in the arithmetic coding.

Firstly, I want reimplement this algorithm.

Secondly, I would like to test and find the answers for the following questions:

- Why the author chose the LR images as auxiliary variables instead of the gray-scale images? Can we use LR gray-scale images as auxiliary variables to further boost the compression performance?

- The author proposed hierarchical structure has three levels with 2*2 ave-pooling. Could we use single level hierarchical structure with 4*4 ave-pooling?
- The 2*2 receptive field fundamentally constraints the network to capture large scale high level structure. Will increasing the receptive field improve the compression performance? How much does the throughput decrease if receptive field increase.

Related papers:

1. Sheng Cao, C. Wu, P. Krahenbuhl, *Lossless Image Compression through Super-Resolution.*
2. A. Kolesnikov, C. Lampert, *PixelCNN Models with Auxiliary Variables for Natural Image Modeling.*
3. *Salimans, Tim, Karpathy, Andrej, Chen, Xi, Knigma, Diederik P., and Bulatov, Yaroslav. PixelCNN++: A PixelCNN implementation with discretized logistic mixture likelihood and other modifications. International Conference on Learning Representations (ICLR), 2017.*