Convolutional Autoencoders and JPEG2000 for Image Compression

David Torpey

December 13, 2017

Abstract

In this paper a novel technique for image compression is introduced. The method is multi-stage, and involves using a Conlutional Autoencoder to learn the sptial relationship in the input image, and hopefully decorrelate the data sufficiently well. The JPEG2000 compression standard is then applied to the encoded images that result from the autoencoder. The JPEG2000 is able to take advantage of the decorrelation achieved by the autoencoder, to achieve even better compression of the data.

1 Introductin

To do...

2 Related Work

To do...

3 Proposed Method

The autoencoder will be described as two separate networks - the *encoder* and the *decoder*. The encoder consists of an input layer, followed by 2 pairs of convolutional and pooling layers, and a final convolutional layer. This final convolutional layer of the encoder network is the encoded form the input into the autoencoder. Without loss of generality, we will assume all images are grayscale. We can define the encoder network as the mapping $f : \mathbb{R}^{n_i \times m_i} \rightarrow \mathbb{R}^{n_e \times m_e \times c_e}$, where $n_i \times m_i$ is the dimensionality of the input image; $n_e \times m_e$ is the dimensionality of the encoded images, and c_e is the number of these encoded images. The decoder network can essentially be seen as a mirror image of the encoder network. It consisting of 2 pairs of up-sampling and deconvolutional layers, followed by a final deconvolutional layer that is the reconstructed version of the input image. This decoder network can be defined as the mapping $g : \mathbb{R}^{n_e \times m_e \times c_e} \rightarrow \mathbb{R}^{n_i \times m_i}$.

For all convolutional and deconvolutional layers, the ReLU activation function is used. Further, the Adam optimization algorithm is chosen to train the network. The loss function used for training is the mean squared error (MSE) between the input image and the reconstructed image. This is perhaps a weakness in the proposed method, and is a potential area of investigation for further research. When using MSE as a loss function for autoencoders, the reconstructions are often blurry. More realistic loss metrics are needed, such as features from a pre-trained CNN, or the SSIM.