A New Image Compression Approach Based on Epitome Model for a Set of Images

ABSTRACT

This paper proposes a new image compression via epitomic model based on a set of images. To be continued...

Categories and Subject Descriptors

E.4 [**Coding And Information Theory**]: Data Compaction and Compression

General Terms

Algorithm, Experimentation, Performance

Keywords

Image Compression, Image Epitomes, Residual

1. INTRODUCTION

This paper proposes a new image compression approach by epitomic model based on a set of images. Our method is developed based on the epitome model of Jojic et al. [1]. This is a generative model learned through a set of regularly shaped patches over a small range of sizes.

To be continued...

2. APPROACH

This section describes the method that we use to compress a set of images, by learning a concise epitome and transform maps, and by encoding the residuals between the image set and the corresponding reconstructions. Specifically,

- to generate a condensed epitome $\mathbf{E}$, given a set of images $\mathbf{I}$, which represents the most useful information of shape and texture within and among images.
- to learn a transform map ϕ_i for each image, which consists of the row and column indices of all transformed image patches in the epitome, so that each image can be reconstructed via those indices.
- to execute lossless and lossy compression to the residual $\mathbf{R}$, of original images and their corresponding reconstruction, in order to gain better reconstruction quality.

MM '15 26-30 October 2015, Brisbane, Australia

2.1 Epitome Learning and Reconstruction

The epitome model $\mathbf{E}$, is a generative model, learned using a set of regularly shaped patches $\mathbf{P}$, over a small range of sizes.

To learn such a concise epitome model, to be continued ...

2.2 Transform Map Encoding and Decoding

Each image will have a unique transform map, including the row and column indices of the transformed patches in the epitome model. Given those transform maps, the reconstruction can be much quickly achieved, compared with the learning process. But there exists a tradeoff between the transform map and the compression ratio. A large transform map, generating high quality of reconstruction, will have large file size, and thus reduce the compression ratio. The file size of transform maps is determined by the number of image patches and the encoding/decoding method to them.

to be continued ...

2.2.1 *Image patches extraction*

During the epitome learning process, 8×8 size of overlapping patches are chosen, with fixed small spacing, for example, 4 pixels, so that a large number of patches will be extracted for training an informative epitome. However, for the reconstruction, a range of sizes of patch spacing, e.g., 4–8 pixels, is used to generate different transform maps, for seeking an appropriate parameter setting achieving optimal rate-distortion result, in terms of PSNR (peak signal-to-noise ratio) and BBP (bits per pixel).

to be continued ...

2.2.2 *Encoding of the transform maps*

How to effectively encode/decode the indices is another factor, essential to the final compression ration. Given the epitome model with dimension of 256×256, the two translation coefficients of $(row, column)$ can be represented by $2 \times 8 = 16$ bits integral numbers. In addition, the transform map Φ is itself highly redundant both spatially and numerically. So the lossless compression, and even lossy compression including quantization and encoding, can be applied to it, but usually at the cost of some reconstruction error.

to be continued ...

2.3 Residual Processing

As a key part of our method, how to process the residual between the original images and the epitomic reconstruction, plays an important role of achieving large compression ration, and at the same time, retaining the high reconstruction quality. Different methods are used to encode the residual as below.

- **Uniform Quantization + Binary File** The residual is uniformly quantized with different quantization level, and saved as binary files, with bit allocation.
- **Uniform Quantization + Lossless Compression** The residual is uniformly quantized with different quantization, and then be compressed losslessly.
- **Lossy JPEG Compression** The residual is directly compressed via lossy JPEG.
- **Lossy JPEG 2K Compression** The residual is directly compressed via lossy JPEG 2K.

to be continued ...

3. EXPERIMENTS

3.1 Datasets

Experiments are performed on our own dataset, where all the images, in the format of BMP, are captured on the campus at Stevens. It contains **10** categories, i.e., *Edwin, Howe Center, Empire State Building, Flower, Playground, Torch Bearer, World Trade Center, Clock, River View, and Walker Gym.* All the categories can be roughly classified into:

- **Similar images** They consist of similarly regular shape object, like, Edwin, and Howe Center.
- **Variant images** To include variant and irregular shape object, like, Flower, River View.
- **Mixed images** They are the combination of similar images and highly variant ones, like, Playground.

In addition, each category above is further classified into different classes due to the number of images. The reason why we set up so many categories of images, is that we want to demonstrate that our method can achieve reliable compression result within a large range of image categories.

3.2 Experimental setup

Different experiments have been performed. Appropriate setup of parameters for epitome learning, image reconstruction, and indices and residual encoding, is very essential to the final rate-distortion result. With a series of parametric configuration, a number of rate-distortion curves are generated, from which the optimal ones will be picked out. Specifically, **five** parameters and **four** different technical approaches are evaluated in this experimental section.

3.2.1 Parametric Configuration

- **Patch Size s** The graph in Figure XXX-??? explores how the physical file size (for epitome E, transform map Φ, residual R, and total) varies as a function of the patch size s, within the range of [4, 40], with step size of 4.
- **Spacing of Patch Extraction Δs** This parameter together with s, determines the number of training patches, and the size of associated transform map. In our experiments, this parameter ranges from $\frac{s}{2}$ to s.
- **Epitome Size** In our experiments the epitome is square, that means its height and width are identical. Thus for simplicity, we only use its height when referring to its dimension. The height of epitome affects both reconstruction quality and compression ratio. Large epitome, albeit beneficial for good reconstruction, takes up much physical storage, and consequently reduces the final compression ratio. We consider a series of heights varying from 2^4 to 2^{10}.
- **Transformation Map Quantization** On one hand, the transform map Φ compresses well due to its locally spacial coherence and numerically redundance. On the other hand, the map is quantized in different quantization level, for lossy compression.
- **Image Number** Image number varies between 50–555.

3.2.2 Residual Compression

Four different approaches, named respectively as UniQuant-0, UniQuant-1, UniQuant-2, and UniQuant-3, are evaluated to compress the residual images, i.e., the error between the original images and the reconstructions.

- **UniQuant-0** The residual is uniformly quantized with different quantization level, and saved as binary files, with bit allocation, in short, lossy uniform quantization, and lossless compressed residual iamges in the format of binary files.
- **UniQuant-1** The residual is uniformly quantized with different quantization, and then be losslessly compressed with JPEG 2000. That is, lossy uniform quantization, plus lossless JPEG 2000.
- **UniQuant-2** The residual is directly compressed with lossy JPEG with different compression quality. That is, lossless uniform quantization, and lossy JPEG.
- **UniQuant-3** The residual is directly compressed with lossy JPEG 2000, with different compression quality. That is, lossless uniform quantization, and lossy JPEG 2000.

3.3 Results

The results are shown here.
to be continued ...

4. SUMMARY AND FUTURE WORK

We have presented an effective compression approach, based on a generative epitomic model which is capable of learning the information of shape and appearance of repeated regions in a set of images.
to be continued ...

5. ACKNOWLEDGMENTS

Thanks to the assistance from ...

6. REFERENCE

[1] N. Jojic, B. Frey, and A. Kannan. Epitomic analysis of appearance and shape. In ICCV, 2003.