Comparison of Image Approximation Methods: Fourier Transform, Cosine Transform, Wavelets Packet and Karhunen-Loeve Transform

Chao-Hsing Hsu, Zhen Guo, Kang Yen
Department of Electrical Engineering
Florida International University
10555 W. Flagler St.
Miami Fl 33174

***Abstract*: -** In this paper we compare the performance of several transform coding methods, Discrete Fourier Transform, Discrete Cosine Transform, Wavelets Packet and Karhunen-Loeve Transform, commonly used in image compression systems through experiments. These methods are compared for the effectiveness as measured by rate-distortion ratio and the complexity of computation.

Key words: Transform coding methods, Nonlinear approximation, Mean square error

1 Introduction

Transform coding methods such as the Discrete Fourier Transform (DFT), the Discrete Cosine Transform (DCT), the Best Wavelets Packet (BWP), and the Karhunen-Loeve Transform (KLT) are basic building blocks of compression systems. All of these methods map an image into a set of N uncorrelated coordinates to reduce information in the image representation. For DFT, DCT, and BWP, the corresponding bases are precise regular signals and independent of the image. But for KLT, the bases selected depend on the image itself. KLT is optimal in the sense that it completely decorrelates a given image in the transformed domain and is also a canonical transformation that minimizes the Mean Square Error (MSE). Unfortunately, the amount of computation ($\sim N^2$ operations) makes this method unpractical, but KLT does provide a benchmark against which other transforms may be judged [1]. M. V. Wickerhauser [2] observed that at high bit-rates (compression ratio 10 or higher) JPEG was better than the orthogonal wavelets and the best-basis of wavelet packets in the MSE sense, and at low bit-rates the wavelet and BWP gave better results. Finding the best wavelets packets requires more than $N2^N$ operations, which is computationally prohibitive [3]. For DCT, since it is closely related to DFT, DCT can be computed by an FFT type algorithm with $N\log N$ operations. Besides, DCT has another very important property that is its asymptotic equivalence to the statistically optimal KLT [1]. So DCT can achieve a good compromise between computational complexity, and coding compression. Therefore, for a fixed computational budget, DCT actually outperforms KLT.

In this paper, we compare the results of image approximation with DCT, KLT, BWP, and FFT. Two different approaches can be used in the selection of the transformed amplitudes. One is linear approximation, which projects the image over N vectors chosen a priori. However, better approximations can be achieved by choosing the N basis vectors based on the image, i.e., nonlinear approximation. The MSE of nonlinear approximation using wavelet bases is superior (with a decay as fast as N^{-2} for $N \rightarrow \infty$) to that of the linear approximation (bounded below N^{-1} for $N \rightarrow \infty$ with the KL bases)[4]. Practically, for a smooth image the nearby pixels are correlated, which means that the value of a pixel conveys information about the likelihood of its neighbors' values. So in the following the nonlinear approximation will be utilized in the comparison of different transformations for their effectiveness as measured by rate-distortion ratio and the complexity of computing the transforms. In addition, it will always be necessary to involve humans in judging acceptable distortion.

2 Decorrelation by Transformations

2.1 Discrete Fourier Transform

The Fourier transform decomposes a signal into frequency components for its analysis. The Discrete Fourier Transform (DFT) estimates the Fourier transform of a signal from a finite number of its sampled points. Since sinusoidal waves last infinite, DFT has very poor local property.

2.2 Discrete Cosine Transform

The Discrete Cosine Transforms (DCTs) cannot be treated simply as a discretized approximation of its continuous Fourier cosine transform. DCT has shown its superiority in bandwidth compression of a wide range of signals such as speech, TV signals, and images.

2.3 Discrete Wavelet Transform

Wavelet transform has emerged as a powerful tool for many applications including data compression and feature detection in sounds, biomedical data and images. The motivation behind its development was the search for fast algorithms to compactly represent of functions and data sets. Compression occurs because pixel values are correlated by the smoothness of the image [5]. The most dissimilarity between DFT and DWT is that individual wavelet functions are compactly supported in space and in frequency. This localization feature makes the representation of image in the transformed domain sparse. This sparseness, in turn, results in a number of useful applications such as dada compression, detecting features in images, and removing noise from signals.

2.4 Wavelet Packets and the best basis

A further degree of freedom can be obtained by choosing the bases adaptively, depending on the signal properties. From families of wavelet packet bases, a fast dynamical programming algorithm is used to select the "best" bases which reflect the signal structures. The advantage of this method over traditional wavelet transform methods is that the bases are chosen automatically to best represent the particular image. In this sense the transform is highly nonlinear [6].

2.5 Karhunen-Loeve Transform

The Karhunen-Loeve Transform (KLT) is a preferred method for approximating a set of vectors by a low dimensional subspace [1]. This subspace is spanned by the eigenvectors of corresponding auto-covariance matrix. This transform is optimal in that it completely decorrelates the signal in the transform domain. Practical implementation of KLT involves the estimation of the auto-covariance matrix of the data sequence, its diagonalization, and the construction of the basis vectors. So the basis vectors are depended on the signal, which cannot be predetermined, and must be completely repeated whenever any new data is added. Although the high computational complexity has made KLT an ideal but impractical tool, it does provide a benchmark against which other discrete transform may be judged [1].

3. Nonlinear Approximations

In linear approximation, we can project a signal f over N vectors chosen a priori.

$$f_N = \sum_{n=1}^{N} < f, \varphi_n > \varphi_n \tag{1}$$

where $\varphi(\varphi_1, \varphi_2, ..., \varphi_N)$ is the basis vector. The truncation error is

$$\varepsilon_N[f] = \| f - f_N \|^2 = \sum_{m=N}^{\infty} |< f, \varphi_m >|^2 \tag{2}$$

The accuracy depends on the characteristics of f relative to the basis vector φ. For instance, if φ is the Fourier basis on [0,1] or a wavelet basis, $\varepsilon_N[f]$ is the L^2-norm. Then the decay of $\varepsilon_N(f)$ depends on the smoothness of f in an L^2-sense. A set of Fourier bases yields efficient linear approximation to smooth signals, which are projected over the N lower frequency sinusoidal waves. For a wavelet basis, the signal is projected over the N largest scale wavelets, which is equivalent to approximating the signal at a coarser resolution. The KL bases consist of N linearly independent eigenvectors and represent the best approximation of a given signal in the MSE sense, for every N.

The performance of the linear approximation can be improved if we choose the N vectors in φ as the most significance N terms (as opposed to the linear approximation using the first N basis functions). The approximation of f with N vectors whose indexes are in I_N is

$$f_N = \sum_{n \in I_N}^{N-1} < f, \varphi_n > \varphi_n \qquad (3)$$

The approximation error is

$$E_N(f) = \| f - f_N \|^2 = \sum_{m \notin I_N} |< f, \varphi_m >|^2 \qquad (4)$$

This error will have a minimum value if the set of N vectors in I_N have the largest inner product amplitude $|< f, \varphi_N >|$. The set I_N depends on the signal f, hence the name non-linear approximation. The difference between a linear and a non-linear approximation has been illustrated in [2] where examples are presented. For the linear transformation, Fourier and wavelet bases give similar results (the MSEs are also roughly equal for the sample images in this paper). In contrast, for the nonlinear approximation, wavelet basis outperform trigonometric bases. Another manifestation of the difference as pointed out in [2] is that for stochastic processes, the KL basis need not be the basis that minimizes the non-linear approximation error.

4. Experimental Comparisons

The transformations described above may be compared for their effectiveness as measured by rate-distortion ratio and humans vision judgment. Nonlinear approximation has been applied to three images: (1) a train image – Capital – with a lot of edges, (2) an artificial image – Ladder – having strong linear relationship between pixels, and (3) a picture of a woman face – Alicia – a typical "piecewise smooth image".

Procedures used in the experiments are giving below.

(i) Apply DCT to an image. We keep the N largest coefficients in the transformed domain and set the others to zero. Then we apply the inverse DCT to reconstructing the image.

(ii) For an image, calculate the eigenvalues and the eigenvectors. We sort the eigenvectors according to the values of the eigenvalues. Keep only the first N eigenvectors as the basis. Then we apply the KLT to restoring the image.

(iii) Use Daub6 (Daubechies 6 wavelets) as the basis functions and search for the "best-basis". We choose entropy as the cost function to extract the N largest coefficients. Then we apply BWT to approximate the original image.

(iv) Perform the forward FFT and keep the N most important terms. Then we carry out the inverse FFT to reproduce the original image.

Distortion measures used in the calculation are the Mean Square Error (MSE) and the Peak Signal-to-Noise Ratio (PSNR) between the original and the approximated images [2].

$$\text{PSNR} = 10\ \text{LOG}_{10}\,(255^2/\text{MSE}) \qquad (5)$$

Figure 1 shows the original image and its approximations by different transforms with 7000 coefficients. BWP and DCT produce very similar approximations from the human vision point of view; however, DCT works better in the PSNR sense. KLT performs a little bit worse than BWP and DCT. The approximation using DFT is unrecognizable by human eyes. The original image in Figure 2 is a very simple and linear one in two dimensions. KLT produces almost the same image as the original one with only 350 coefficients. This implies that KLT provides the best basis for linear decorrelation. BWP (using Daub6) does not perform well at the edge. If using Daub 1, the Haar transform, we can get a much better "edge approximation". This is because Daub 6 basis functions have much larger space support than Daub1. This also tells us wavelets have infinite basis functions, but DCT has only one. From the Figure 2(e), we can see some sinusoidal waves. KLT is the best in this case. In Figure3, we use 6000 coefficients to approximate the original image. DCT is somehow better than BWP. In both cases, the details are recognizable. FFT only keeps the lower frequency components and omits the details (higher frequencies).

The above mentioned experiments were performed by a PC (HP pavilion 6635 with 533MHz). The processing time for FFT, DCT, KLT, and BWP are 0.55s, 1.32s, 4.06s and 107.71s, respectively.

5. Discussion and Conclusions

From the experiments performed, we obtain the following observations:

(i) If pixels in an image have a strong linear relationship, KLT will provide the most efficient approximation (See Figure 2).

(ii) For nonlinear approximation, Discrete Cosine Transform provides the best results in most our experiments.

(iii) For nonlinear approximation, Best Wavelet Packets Transform (using Daubechies 6 wavelets) is not as efficient as DCT in most our experiments.

(iv) For nonlinear approximation, Fourier bases are not efficient at all.

(v) Discrete Cosine Transform and Wavelet Packets basis always yield "visually pleasant" images.

References:

[1] Rao, K. R.; Yip, P., Discrete Cosine transform: Algorithms, Advantages And Applications, Academic Press, 1990.

[2] Wickerhauser M.V., "High-Resolution Still Picture Compression,"Digital Signal Processing: a Review Journal, Vol.2, No.4, 1992, pp.204-226

[3] Mallat, S., A Wavelet Tour of Signal Processing, 2nd Edition, Academic Press, 1999.

[4] Cohen, A., Daubechies, I., Guleryuz, O.G., Orchard, M.T, "On the Importance of Combining Wavelet-based Non-Linear Approximation with Coding Strategies," Preprint.

[5] Mallat, S., "A Theory for Multi-resolution Signal Decomposition: The Wavelet Representation," IEEE Trans. on Pattern and Machine Intelligence, Vol.11, No.7, 1989, pp.674-693.

[6] Cohen, A., D'Ales, J., "Non-linear Approximation of Random Functions," SIAM J. of Appl. Math., Vol.57, No.2, 1997, pp.518-540.

Figure 1 (a) Capital --- Original Image 256X256

Figure 1. (b) BWP with 7000 coefficients
PSNR = 26.8

Figure 1. (c) KLT with 7000 coefficients
PSNR = 28.8

Figure 1. (d) DCT with 7000 coefficients
PSNR = 27.3

Figure 1. (e) FFT with 7000 coefficients
PSNR = 20.8

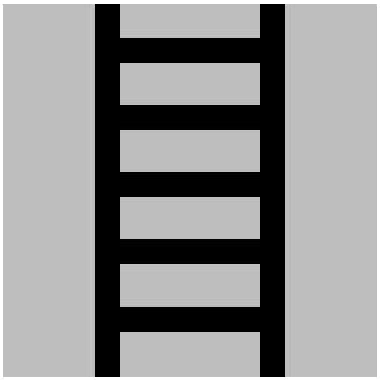

Figure 2. (a) Ladder Original Image 256X256

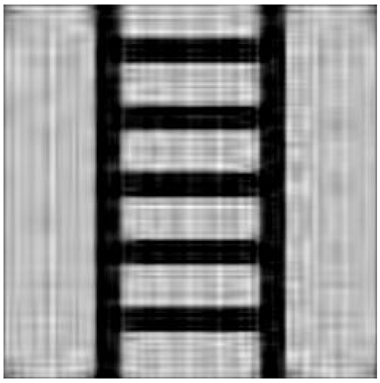

Figure 2. (b) BWP with 1000 coefficients
PSNR = 20.7

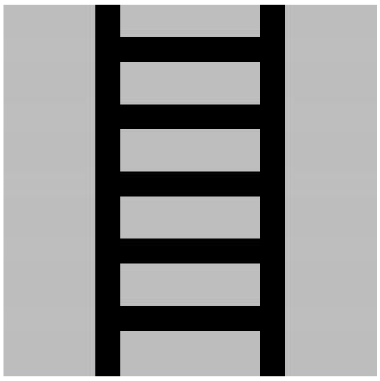

Figure 2. (c) KLT with 350 coefficients
PSNR = 74.32

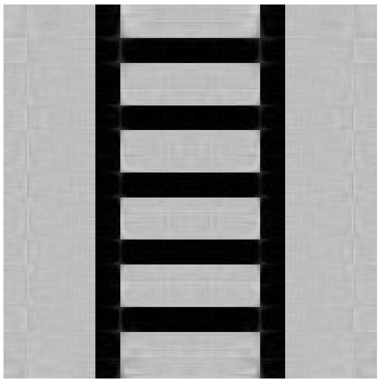

Figure 2. (d) DCT with 1000 coefficients
PSNR = 33.3

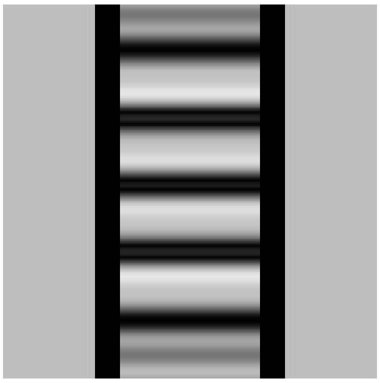

Figure 2. (e) FFT with 1000 coefficients
PSNR = 19.8

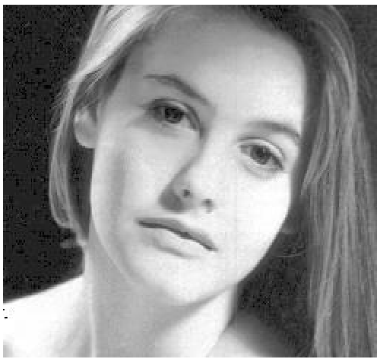

Figure 3. (a) Alicia Original Image
256X256

Figure 3. (b) BWP with 6000 coefficients
PSNR = 31.9 db6

Figure 3. (c) KLT with 6000 coefficients
PSNR = 32.7

Figure 3. (d) DCT with 6000 coefficients
PSNR = 32.7

Figure 3. (e) FFT with 6000 coefficients
PSNR = 29.5