Proceedings of the 6th WSEAS International Conference on Multimedia Systems & Signal Processing, Hangzhou, China, April 16-18, 2006 (pp62-66)

A PCA-WAVELET BASED COMPRESSION FOR DISTANCE LEARNING IMAGES

Zong Min WANG[1], Qing Fu FANG[2], Bing ZHOU[1], and Qin LI[1]
[1]Henan Provincial Key Lab on Information Network
[2]Logistics Group Ltd
Zhengzhou University, Zhengzhou, China

Abstract: The large volume of images used in Distance Learning System are required to be compressed in a good ratio to release the storage loading of computer servers. A new image coding Scheme based on Principle Components Analysis (PCA) and Wavelet decomposition is proposed in this paper. Our algorithm includes 1) Principle Components Analysis (PCA)to reduce the information redundancies along temporal dimension; 2) a texture energy (TE) based technique used to optimize the PCA analysis; 3) Wavelet decomposition and optimized LBG algorithm for compression along spatial dimention. The experimental results demonstrate that our proposed coding scheme achieves good performance.

Key–Words: Distance Learning, Image Compression, PCA, Wavelet Decomposition, Image Classification

1 Introduction

To make the distance learning media, the speeches of lecturers are recorded as audio files and the handouts are stored as images to make the distance learning media. According to our experience, the students could be satisfied by low quality audios but they are very sensitive to the quality of images. This causes very high pressure to the servers on VOD services and storage capacity. A new image coding technique is required to compress the handout images.

The most popular image compression standard is JPEG (Joint Photographic Experts Group) [1] that is based on DCT (Discrete Cosine Transform). This standard is open blamed for the blocking effects at high compression ratios [2]. A DWT (Discrete Wavelet Transform) based standard JPEG2000 was established in recent years to overcome that drawback [3]. In stead of operating on 8x8 segments as DCT, wavelet transform applies a global operation on an entire image to avoid the blocking artifacts. By combing wavelet transform and quantization techniques, many wavelet based image coding schemes have been proposed [4] [5] [6] [7]. Because Vector Quantization (VQ) usually produces better quantization than Scalar Quantization (SQ) [8], most of current wavelet based image coding schemes employ VQ for quantization. The codebook of VQ is usually generated through LBG [7] [9] or lattice based methods.

Since there are quantities of temporal redundancies in a sequence of handout images, compression can also be performed in the temporal domain. MPEG (Moving Pictures Experts Group) [10] is the most popular compression standard for video sequence. But it is more suited to a video sequence captured in real-time rather than a sequence of images manually arranged. As a data-driven technique to describe the variance-covariance structure of a data set [11], principle component analysis (PCA) has the ability to reduce large data sets to a smaller number of significant channels. Many PCA based techniques have been proposed for the compression of dynamic images in temporal domain [12] [13].

By analyzing the compression techniques in both temporal and spatial domain, we propose a new compression scheme to improve the compression efficiency of Distance Learning images. Our algorithm Our algorithm includes 1) Principle Components Analysis (PCA)to reduce the information redundancies along temporal dimension; 2) a texture energy (TE) based technique used to optimize the PCA analysis; 3) Wavelet decomposition and optimized LBG algorithm for compression along spatial dimension. Figure 1 gives an overview of the proposed scheme.

2 TE Based Classfication

Since there are always quantities of temporal redundancies in a sequence of handout images, principle component analysis (PCA) is an effective method to do the compress. Figure 2 shows some examples of some temporal redundancies in a sequence of handout images.

However, a good compression will not be expected by applying PCA directly on a sequence of

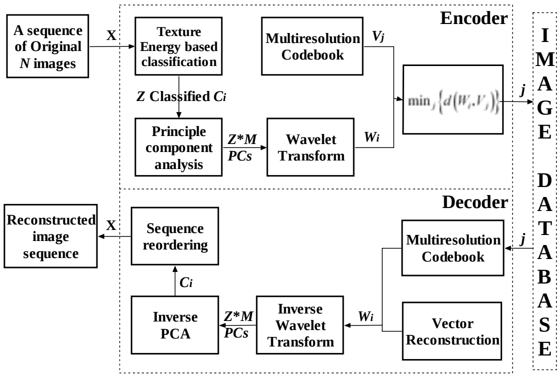


Figure 1. Encoding/decoding scheme

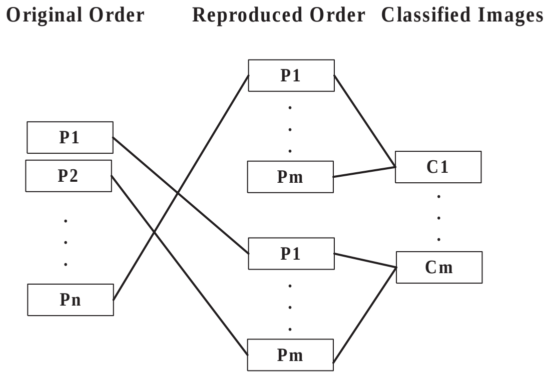


Figure 3. A mapping from original sequence to classified images

handout images because PCA is only well suited to highly correlated data [11] but the images of handout are often obtained from many different sources (the difference of image nature at different rows of figure 2 is obvious). In order to produce better compression, we apply image classification before PCA. Classes comprising relatively highly correlated images are generated by extensively using of a texture based technique proposed in our previous work [14].

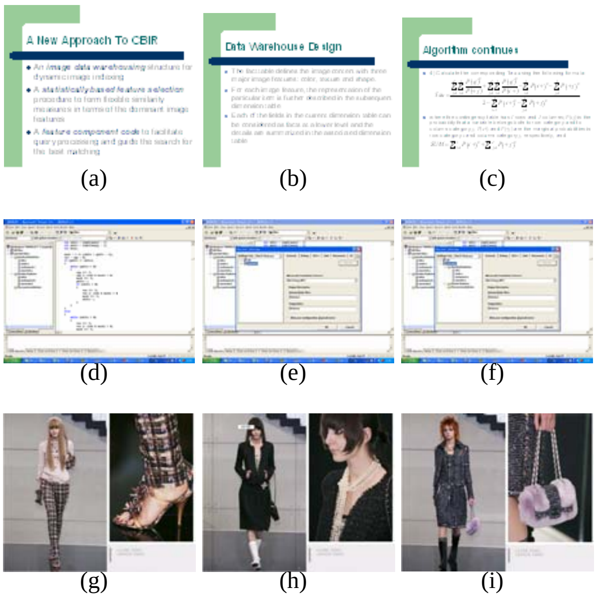


Figure 2. Handout image sequences

There are a finite number of classes $C_i, i = 1, \ldots, Z$. A training collection comprising a number of training images belonging to each class is available. Based on the feature extracted from these sets, a classifier is designed to recognize a given test image of unknown class to one of the Z classes. A page number mapping table is constructed during the classification the record the relation between original order and the classified images. Figure 3 illustrates this mapping.

The classifier is the normalized "texture energy" (TE): the standard deviation of pixel gray scale within a 15×15 window size computed after convolution with a texture tuned mask through task-aimed training. The method consists of first optimizing the mask parameters on the two-dimensional linked training samples to characterize the rotation and scale invariant features of each texture by maximizing the inner-class convergence and inter-class dispersion, then classifying samples from different textures that have been rotated by different angles and magnified with different scale factors. The classification is based on a distance rule which measures the difference between the global texture energy TE of the test image and the reference values in the 2D matrix of the training texture database. The texture is classified to the category for which such a distance is the minimum.

3 PCA Based Compression along Temporal Dimension

After the classification of a sequence of images, PCA is applied to each C_i. A three bit marker is added in each image to record the page number. If there are N images in C_i, M, $M < N$, PCA channels will be employed. Let $\mathrm{X} = [\mathrm{X}_1 \mathrm{X}_2 \ldots \mathrm{X}_N]$ be the N images in C_i, where X_i is a $K \times 1$ column vector and L is the number of pixels in each image, then the mean vector is defined by

$$\overline{\mathrm{X}} = \mathrm{E}\left\{\mathrm{X}\right\} \tag{1}$$

and the covariance matrix is

$$C = \mathrm{E}\left\{ (X - \overline{\mathrm{X}}) \cdot (X - \overline{\mathrm{X}})^T \right\} \tag{2}$$

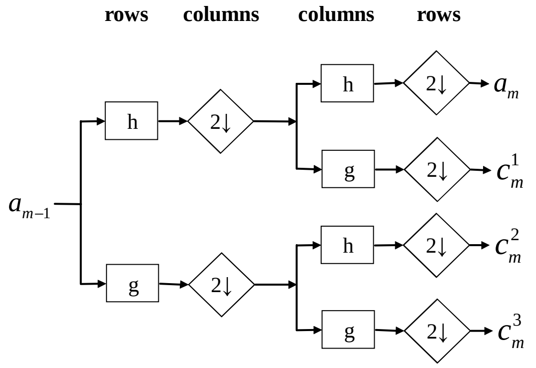


Figure 4. A single-scale filter bank of FWT

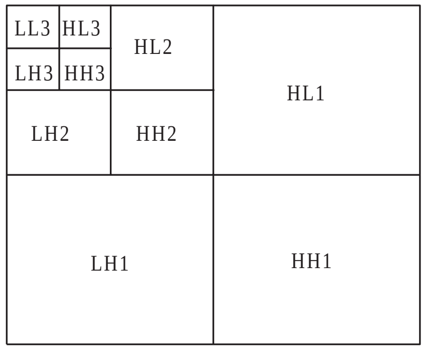


Figure 5. Three level 2-D FWT decomposition

Let $\lambda = [\lambda_1 \lambda_2 \ldots \lambda_L]$ and $A = [\alpha_1 \alpha_2 \ldots \alpha_L]$ be the eigenvalues and eigenvectors of C respectively, then the principal components are defined by

$$P = A \cdot \left(\mathrm{X} - \overline{\mathrm{X}}\right) \tag{3}$$

M principle components (PCs) will be employed as the representation of the image sequence and M is computed by

$$\sum_{i=1}^{M} \lambda_i \bigg/ \sum_{i=1}^{L} \lambda_i \geq T \tag{4}$$

where T is taken as 0.9 in our experiments.

4 Wavelet Based Compression along Spatial Dimension

After the principle component analysis, M, $M < N$, PCA channels containing most of the information of the original N images are generated. Then wavelet transform is performed to each PCA channel for further coding. Employing a pair of filters and downsamplers, wavelet coefficients c and approximation coefficients a can be computed through the Fast Wavelet Transform (FWT)[15]

$$c_{m,n} = \sum_k g_{2n-k} a_{m-1,k} \tag{5}$$

$$a_{m,n} = \sum_k h_{2n-k} a_{m-1,k} \tag{6}$$

where g is a high pass filter and h is a low pass filter; m and n are scaling parameter and shifting parameter respectively; $a_{0,n}$ is the original signal.

Usually biorthogonal wavelet bases are employed in image coding because they have linear phase and perform exact reconstruction [5]. Biorthogonal wavelet bases were introduced in [16]. There are two pairs of filters: the pair h and g introduced above for decomposition and another pair $\widetilde{h}$ and $\widetilde{g}$ for reconstruction. The relations between these filters are

$$g_n = (-1)^n \, \widetilde{h}_{-n+1} \tag{7}$$

$$\widetilde{g}_n = (-1)^n \, h_{-n+1} \tag{8}$$

The so-called 9/7 [17] and 12/4 [18] biorthogonal wavelets are the most popular choice of filters. Both of them are employed in our experiments for a comparative study.

It is very easy to extend above FWT to the 2-D case (image). For an image $a_{m,x,y}$, convolving its rows with h and g and downsampling its columns, we obtain two subimages. Then, convolving columns of those two subimages with h and g and downsampling their rows, four quarter-size output subimages a_m, c_m^1, c_m^2, c_m^3, are produced. Figure 4 illustrates this procedure through one single-scale filter bank and this filter bank can be iterated by putting the approximation output to the input of another filter bank. Conventionally, the four subimages are denoted as LL, LH, HL, and HH respectively, where LL is the approximation coefficients and others are the wavelet coefficients. Figure 5 illustrates subimages after three-level 2-D FWT decomposition.

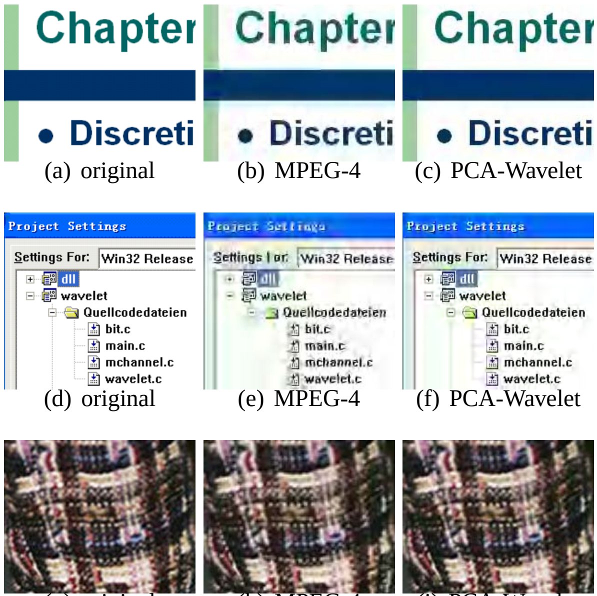
After n level wavelet decomposition, subimage HH_i, HL_i, LH_i and LL_n are then taken as vectors for Vector Quantization. Vector Quantization is performed by mapping an input vector, $x = (x_0, \ldots, x_{k-1})$, to a reproduction vector, $\widehat{x} = q(x)$, belonging to a codebook, $\widehat{\mathrm{A}} = \{y_i; i = 1, \ldots, N\}$, where q is a N-level K-dimensional quantizer that is described by the codebook $\widehat{\mathrm{A}}$ and the partition $\mathrm{P}\left(\widehat{\mathrm{A}}\right) = \{S_i; i = 1, \ldots, N\}$ where $S_i = \{x : q(x) = y_i\}$ is the mapping from input vectors to the codebook [8].

LBG algorithm [9] is a well-known technique to generate the quantizers. To produce the locally optimal quantizer, we optimize the LBG algorithm by employing classified training sequences that are corresponding to the training collections mentioned in section 2. Corresponding to the mutliresolution images generated by FWT decomposition, there are multiresolution codebooks generated from multiresolution training sequences.

- (A) The LBG Algorithm:
  - (1) Initialization:Given N = Number of levels of the quantizer, a distortion threshold $0 \leq \varepsilon \leq$, an initial N-level codebook , $\widehat{\mathrm{A}}_0 = \{y_i; i = 1, \ldots, n\}$and a training sequence $X = \{x_j; j = 1, \ldots, n\}$. Set $m = 0$ and $D_{-1} = \infty$.
  - (2) Given $\widehat{\mathrm{A}}_{\mathrm{m}} = \{y_i; i = 1, \ldots, n\}$, find its minimum distortion partition $\mathrm{P}\left(\widehat{\mathrm{A}}_{\mathrm{m}}\right) = \{S_i; i = 1, \ldots, N\}$ of the training sequence: $x_j \in S_i$ if $d\left(x_j, y_i\right) \leq d\left(x_j, y_l\right)$ for all l. d is the Euclidean distance. Compute the average distortion by
  $$\begin{aligned} D_m &= D\left(\left\{\widehat{\mathrm{A}}_m, \mathrm{P}\left(\widehat{\mathrm{A}}_m\right)\right\}\right) \\ &= \frac{\sum\limits_{j=1}^{n} \min_{y \in \widehat{\mathrm{A}}_m} d(x_j, y)}{n} \end{aligned}$$
  - (3) If $(D_{m-1} - D_m)/D_m \leq \varepsilon$, halt with $\widehat{\mathrm{A}}_m$ and $\mathrm{P}\left(\widehat{\mathrm{A}}_m\right)$ to describe the final quantizer. Otherwise, continue.
  - (4) Find the optimal codebook $\widehat{x}\left(\mathrm{P}\left(\widehat{\mathrm{A}}_m\right)\right) = \{\widehat{x}\left(S_i\right); i = 1, \ldots, N\}$, which is the centroid of each S_i. Set $\widehat{\mathrm{A}}_{m+1} \triangleq \widehat{x}\left(\mathrm{P}\left(\widehat{\mathrm{A}}_m\right)\right)$. Replace m by $m+1$, and go to (2).
- (B) Choice of $\widehat{\mathrm{A}}_0$:Choosing $\widehat{\mathrm{A}}_0$ by splitting gives good performance [7], which is briefly described here:
  - (1) Initialization: Set $M = 1$, and define $\widehat{\mathrm{A}}_0 = \widehat{x}\left(\mathrm{A}\right)$, which is the centroid of the training sequence.
  - (2) Given the codebook $\widehat{\mathrm{A}}_0\left(M\right) = \{y_i, i = 1, \ldots, M\}$, split each y_i into two close vectors $y_i + \sigma$ and $y_i - \sigma$ where σ is a fixed perturbation vector. The collection $\widetilde{\mathrm{A}}\left(M\right) = \{y_i + \sigma, y_i - \sigma; i = 1, \ldots, M\}$ has $2M$ vectors. Replace M by $2M$.
  - (3) If $M = N$, halt with $\widehat{\mathrm{A}}_0 = \widetilde{\mathrm{A}}\left(M\right)$. If not, run LBG algorithm for a M-level quantizer on $\widetilde{\mathrm{A}}\left(M\right)$ and then go to (2).

5 Experimental Results

Experiments are implemented using the images from handouts in the Distance Learning database belonging to Distance Learning School of Zhengzhou University. 37 handout sequences of different subjects are employed. Each of those sequences contains 32-83 images. The performance of the proposed algorithm is evaluated in terms of $PSNR$ and the perceptual qualities.



(a) original (b) MPEG-4 (c) PCA-Wavelet
(d) original (e) MPEG-4 (f) PCA-Wavelet
(g) original (h) MPEG-4 (i) PCA-Wavelet

Figure 6. Reconstructed images at 28 kbit/s

$PSNR$ is defined by

$$PSNR = 10 \lg \left(\frac{255^2}{MSE}\right) \tag{9}$$

where MSE is the mean squared error of the reconstruction.

$$MSE = \frac{1}{N} \sum_{i=1}^{N} \left(f_i - \widehat{f}_i\right)^2 \tag{10}$$

Figure 6 displays some reconstructed images from different test sequenecs at 28 kbit/s. The average $PSNR$ of the 37 test sequences is illustrated in figure 7 with comparisons to MPEG-4. The performance of our PCA-wavelet based scheme achieves better performance than MPEG-4 at low and medium bit-rate. Both 9/7 and 12/4 biorthogonal wavelets are tested in this experiment. The performances of different wavelets are generally similar. The 9/7 wavelet achieves a little better $PSNR$ at low bit-rate and the 12/4 wavelet a little better at medium bit-rate.

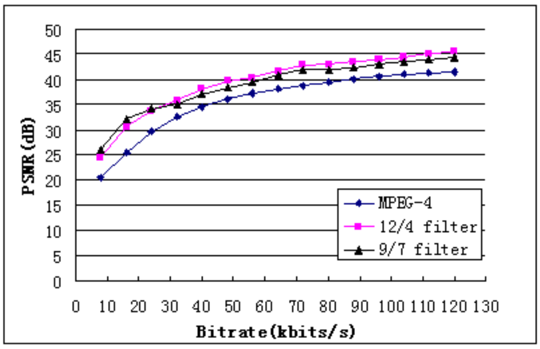


Figure 7. PSNR vs. Bit-rate

6 Conclusion

We proposed a new PCA-Wavelet based compression scheme for Distance Learning images. Our coding scheme includes the Principle Components Analysis (PCA) in temporal domain and wavelet based compression in spatial domain. A texture energy (TE) based technique is used to classify the image sequence so that the principle component analysis (PCA) is optimized to produce better representative PCA channels for our handout images. Further, employing classified training sequences generated based on the TE of the handout images, the LBG algorithm is optimized to produce locally optimal codebook. The experiments results demonstrate that our algorithm achieves good performance at low bit-rate.

Acknowledgements: The authors would like to thank the support from Distance Learning School and Henan Provincial Key Lab on Information Network of Zhengzhou University. The Research Grants from The Science & Technology Department of Henan Province and The Education Department of Henan Province are greatly appreciated.

References:

[1] G. K. Wallace, "The jpeg still picture compression standard," *Communication of the ACM*, vol. 34, no. 4, pp. 30–49, April 1991.

[2] A. Zakhor, "Iterative procedures for reduction of blocking effects in transform image coding," *IEEE Transactions on Circuits and System for Video Technology*, vol. 2, no. 1, pp. 91–95, March 1992.

[3] "Jpeg 2000 image coding system," ISO/IEC, Tech. Rep. 15444-1-2000, 2000.

[4] J. M. Shopiro, "Embeded image coding using zerotree of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3445–3462, 1993.

[5] M. Barlaud, P. Sole, T. Gaidon, M. Antonini, and P. Mathieu, "Pyramidal lattice vector quantization for multiscale image coding," *IEEE Transactions on Image Processing*, vol. 3, no. 4, pp. 367–380, 1994.

[6] M. Antonini, M. Barlaud, and P. Mathieu, "Image coding using lattice vector quantization of wavelet coefficients," in *IEEE ICASSP*, vol. 4, April 1991, pp. 2273 – 2276.

[7] A. Averbuch, D. Lazar, and M. Israeli, "Image compression using wavelet transform and multiresolution decomposition," *IEEE Transactions on Image Processing*, vol. 5, no. 1, pp. 4 –15, January 1996.

[8] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Kluwer Academic, 1992.

[9] Y. Linde, A. Buzo, and R. Gray, "An algorithm for vector quantizer design," *IEEE Transactions on Communications*, vol. 28, no. 1, pp. 84–95, January 1980.

[10] "Mpeg-4 visual," ISO/IEC, Tech. Rep. 14496-2-1999, 1999.

[11] B. S. Everitt and G. Dunn, *Applied Multivariate Data Analysis*. London: Arnold, 1991.

[12] V. Chameroy and R. D. Paola, "High compression of nuclear medicine dynamic studies," *Int. J. Card. Imaging*, vol. 5, pp. 261–269, 1990.

[13] Z. Chen, B. J. Parker, D. D. Feng, and R. Fulton, "Temporal processing of dynamic positron emission tomography via principal component analysis in the sinogram domain," *IEEE Transactions on Nuclear Science,*, vol. 51, no. 5, pp. 2612 – 2619, October 2004.

[14] J. You, W. Li, and D. Zhang, "Hierarchical palmprint identification via multiple feature extraction," *Pattern Recognition*, vol. 35, pp. 847–859, 2002.

[15] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Recognition and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[16] A. Cohen, I. Daubechies, and J. C. Feauveau, "Biothogonal bases of compactly supported wavelets," *Comm. Pure. Appl. Math.*, vol. 45, pp. 485–560, 1992.

[17] J. D. Villasenor, B. Belzer, and J. Liao, "Wavelet filter evaluation for image compression," *IEEE Transactions on Image Processing*, vol. 4, no. 8, pp. 1053 – 1060, August 1995.

[18] O. Egger and W. Li, "Subband coding of images using asymmetrical filter banks," *IEEE Transactions on Image Processing*, vol. 4, no. 4, pp. 478–487, April 1995.