4th WSEAS International Conference on ELECTRONICS, CONTROL and SIGNAL PROCESSING, Miami, Florida, USA, 17-19 November, 2005 (pp.257-262)

A New Coding System for Monochromatic Images Based on Wavelet Transform and Singular Value Decomposition (HDWTSVD)

Humberto Ochoa[a,b], K.R. Rao[a]

[a]Universidad Autónoma de Cd. Juárez
Henry Dunant 4016, Zona Pronaf
Cd. Juárez, Chihuahua. México

[b]University of Texas at Arlington, Box 19016
416 Yates Street, Nedderman Hall, Rm 530
Arlington, TX 76019

Abstract: In this paper the HDWTSVD algorithm to encode monochromatic images is proposed. The algorithm combines DWT and SVD techniques. The input image is divided into tiles of 64x64 pixels. A criterion based on the average standard deviation of 8x8 subblocks is used to choose DWT or SVD. If the tile exhibits a high average standard deviation, it is compressed by using SVD otherwise by DWT. Eigenvalues and eigenvectors are scalar and vector quantized respectively. If the MSE of the highest three eigenvectors, with respect to the vectors in the codebooks, is greater than a threshold then the eigenvectors are scalar quantized and sent, otherwise the indices of de codevectors are sent. The remaining eigenvectors are quantized using small codebooks. The maximum number of levels of subband decomposition for each tile is 3. Each subband is coded using the HC-RIOT. Each subband is coded using HC-RIOT. The proposed algorithm is applied to image coding and its performance is discussed.

Keywords: coding of monochrome images using wavelets, singular value decomposition, HC-RIOT, discrete wavelets, vector quantization.

1. Introduction

The phenomenal increase in the generation, transmission, and use of digital images in many applications is placing enormous demands on the storage space and communication bandwidth. Data compression algorithms are a viable approach to alleviate the storage and bandwidth demands. They are key enabling components in a wide variety of information technology applications that require handling a large amount of information. From text and image representation in digital libraries to video streaming over the Internet, current information transmission and storage capabilities are made possible by recent advances in data compression.

In recent years, many compression techniques have been developed in different fields, specially in the subband coding (SBC) field - namely wavelets in applied mathematics, subband coding in digital signal processing, and multiresolution in computer vision that have converged to a unified theory. SBC is a powerful technique, which is efficiently implemented using filter banks to split and merge the image without distortion. In SBC the filtered images are downsampled to their respective Nyquist rates. The downsampling operation, performed by integer decimation for practical reasons, introduces distortions due to aliasing and filtering. Reconstruction theory for filter banks demonstrates that alias-free and distortion-free solutions exist [1], [2], [3].

There exist various techniques to construct wavelet bases, or to factor the filters into basic building blocks. One of these is lifting, which is known as the second generation wavelets. A construction using lifting, is completely spatial and is used when Fourier techniques are no longer available.

The basic idea of compression using the DWT is to exploit the local correlation that exists in most of the images for building an aproximation. In the first generation wavelets, the Fourier transform is used to build the space-frequency localization. However, in the second generation wavelets, this can be done in the spatial domain and can reduce the computational complexity of the wavelet transform by a factor of two [4]. Orthogonal transformations provide a good performance for signals with high correlation. However, they achieve a poor performance for signals with low correlations.

Together with the DWT, new algorithms to encode the resulting subbands have emerged. For example, the Embeded Zerotree Wavelet (EZW), introduced by Shapiro [5], the Set Partitioning in Hierarchical Trees (SPIHT) proposed by Said and Pearlman [6], and the HC-RIOT developed by Syed [7], belong to this category. The former is a quantization and coding strategy that incorporates some characteristics of the wavelet decomposition. It takes advantage of the fact that there are wavelet coefficients in different subbands that represent the same spatial location in the image. The second uses a partitioning of the trees in a manner that tends to cluster insignifficant coefficients together in a larger subset. The last algorithm combines techniques of the zerotree entropy (ZTE) algorithm introduced by Martucci [8] and a modified SPIHT to improve the quality of the images at low bit rates.

The Singular Value Decomposition (SVD) technique provides optimal energy-packing efficiency for any given image, but its application is very limited due to the computational complexity associated with the computation of eigenvalues and eigenvectors [9]. The best results for SVD image compression have been obtained by combining SVD and vector quantization (VQ) of the eigenvectors [10], [11].

2. SVD Coding

The SVD is a transform suitable for image compression because it provides optimal energy compaction for any given image [9]. A good representation of the image can be achieved by taking only a few largest eigenvalues and corresponding eigenvectors. A (NxN) matrix $\boldsymbol{A}$ is decomposed to form two orthogonal matrices, $\boldsymbol{U}$ and $\boldsymbol{V}^{\boldsymbol{T}}$, representing the eigenvectors, and a diagonal matrix $\boldsymbol{\Sigma}$ representing the eigenvalues.

$$\boldsymbol{A} = \boldsymbol{U\Sigma V}^{\boldsymbol{T}} \tag{1}$$

Where r is the rank of $\boldsymbol{A}$

$$\begin{aligned} \sum{}^{T}\sum &= diag(\lambda_1, \lambda_2 ......\lambda_r, 0,0,0,0) \\ &= diag(\sigma_1^2, \sigma_2^2, ..., \sigma_r^2, 0,0,0,0) \\ \lambda_1 \geq \lambda_2 ... \geq \lambda_r &> \lambda_{r+1} = ... = \lambda_N = 0 \end{aligned} \tag{2}$$

We can calculate the columns $\boldsymbol{v}(n)$ of $\boldsymbol{V}$ by solving

$$(\boldsymbol{A}'-\lambda(n)\boldsymbol{I})\,\boldsymbol{v}(n)=0 \qquad n=1,.....,r \tag{3}$$

Where $\boldsymbol{A}' = \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}$. The columns of $\boldsymbol{U}$ are

$$\boldsymbol{u}(n) = \frac{1}{\sqrt{\lambda(n)}}\boldsymbol{Av}(n) \qquad n=1,.....,r \tag{4}$$

The original block can be estimated by retaining the q largest eigenvalues and corresponding eigenvectors

$$\hat{A} = \sum_{n=1}^{q} \sqrt{\lambda(n)}\, \boldsymbol{u}(n)\, \boldsymbol{v}^{\boldsymbol{T}}(n) \qquad q \leq r \tag{5}$$

The square error is equal to the sum of the discarded eigenvalues

$$\sum_{m=1}^{r} \sum_{n=1}^{r} \left| X(m,n) - \hat{X}(m,n) \right|^2 = \sum_{n=q+1}^{r} \lambda(n) \tag{6}$$

Where $X(m,n)$ is the original sample and $\hat{X}(m,n)$ is the reconstructed sample of a subblock. The energy contained in q retained eigenvalues is

$$E = \sum_{n=1}^{q} \lambda(n) \tag{7}$$

SVD yields two matrices of eigenvectors and one of eigenvalues. VQ techniques have been used successfully to encode the eigenvectors and scalar quantization techniques (SQ) to encode the eigenvalues [9] [11].

3. The proposed algorithm

Figure 1 shows the proposed Hybrid DWT-SVD system. The decision threshold on what transform to use is based on the average standard deviation criterion (ASTD).

A. Encoder

Before encoding, the source image is divided into tiles of 64x64 pixels. Each tile is encoded independently. The ASTD is first calculated on the source tile by calculating the average of the standard deviations of 8x8 subblocks of a tile (64 subblocks). If the ASTD is high the mean of the 8x8 subblock is subtracted and SVD is applied to decompose the subblock into two orthogonal matrices ($\boldsymbol{U}$ and $\boldsymbol{V}^{\boldsymbol{T}}$), containing the eigenvectors, and a diagonal matrix (Σ) containing the square eigenvalues as shown by equation 1. The eigenvectors and correspondent eigenvalues are rearranged in decreasing energy order. The mean of the subblock is quantized uniformly with 8 bits.

In the adaptive reconstruction stage, the eigenvalues are discarded starting with the eigenvalue containing less energy. Each time an eigenvalue is discarded, the MSE is calculated until a MSE threshold is met. When the MSE is below the threshold, the eigenvalues are coded using scalar quantization and the eigenvectors are passed to the decision stage.

In the decision stage, the first three eigenvectors of matrices are vector quantized using codebooks of 256, 128 and 32 codewords. If the MSE of the indexed codeword is below 0.01, 0.1 and 0.4 respectively, with respect to the original eigenvectors, then the indices are transmitted. If the MSE is above these thresholds, the components of the first eigenvector are uniformly quantized to 7 bits and the components of the second and third eigenvectors are uniformly quantized to 5 bits each and sent to the encoded file. One extra bit is included to indicate to the decoder that the data corresponds to an index of the codebook or to the quantized eigenvector. The codebooks for the fourth and fifth eigenvectors are of 32 code words each. The sixth, seventh and eighth codebooks are of length 2.

If the ASTD of the source tile is low, the mean of the tile is subtracted and quantized to 8 bits. The resulting tile is encoded using a 9/7 Daubechies filter bank factored using lifting steps to help reduce computational complexity [4]. Before each filtering stage, the tile, or the corresponding subband, is extended using symmetric periodic extension to reduce the block-artifact in the reconstructed tile [4] and to avoid coefficients expansion. Each 64x64 tile is decomposed into a maximum of three-levels; the level zero or last decomposition level is an 8x8 block. The subband coefficients are non-integer. Therefore they are rounded to the nearest integer, which causes a minimal loss of PSNR [7].

The 9/7 Daubechies filter bank was chosen because it is biorthogonal and uses symmetric odd length filters (linear phase), which is required to preserve the symmetry of the data along subbands. The transform is separable. Therefore, filtering along rows followed by filtering along columns is performed. The resulting subbands, after one, two or a maximum of three levels of decomposition, are encoded using HC-RIOT. Each encoded tile will start with a ‘1’ if it was encoded using SVD or ‘0’ if DWT was used. Choice of the DWT or SVD codebooks sizes for eigenvectors (also VQ or SQ) etc, have been decided after extensive simulations using the test data.

B. Decoder

At the decoder side the decision to decode using HC-RIOT or inverse VQ/SQ is performed by reading the first bit of each encoded tile. If the bit is a ‘1’ then the 8x8 eigenvectors and eigenvalue matrices $\boldsymbol{U}$, $\boldsymbol{V}^{\boldsymbol{T}}$, and Σ are constructed. The reconstruction process starts by reading a second bit. If it is a ‘0’ then the entry indices for the first codebook is recovered, if it is a ‘1’ the quantized eigenvector is recovered and dequantized. This process is repeated for the second and third eigenvectors. The remaining eigenvectors are reconstructed by recovering the entry indexes as well as the mean value. This process is repeated until the 64x64 tile is recovered.

If the first bit of the encoded tile is a ‘0’ the decoding process uses HC-RIOT. The decoding process starts with the maximum threshold and after each pass the threshold is divided by 2 until the minimum threshold is reached. The entire process can be stopped after decoding the total number of bits indicated by the encoder or after reaching a desired distortion in the recovered 64x64 tile. A synthesis filter bank is applied to the recovered subbands in order to recover the tile. An approximation of the tile is recovered after adding back the mean, which is part of the header of the encoded tile.

4. Results

In this section we present the performance of the HDWTSVD and compare with the SVD and JPEG baseline. We used the PSNR as a measure of the distortion in terms of the mean square error (MSE).

$$\mathrm{PSNR} = 10\log_{10}\frac{(255)^2}{\mathrm{MSE}} \quad \mathrm{dB} \qquad (8)$$

(a)

(b)

(c)

(d)

Figure 2. Recovered Lena image compressed (a) at 1.08 bpp, 36.51 dB and (b) error image, (c) compressed at 0.55 bpp, 34.20 dB and (d) error image.

We used 512 x 512 8-bit monochromatic image "Lena". The recovered and the error images are shown in figure 2 (a) and (b) for hight bit rates (above 1 bpp) and for low bit rates in figure 2 (c) and (d).

Eq. (9) and (10) were used to calculate the MSE and the PSNR respectively.

$$MSE = \frac{1}{3xMxN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\Big(\big[r(m,n)-\hat{r}(m,n)\big]^2 + \big[g(m,n)-\hat{g}(m,n)\big]^2 + \big[b(m,n)-\hat{b}(m,n)\big]^2\Big) \quad (9)$$

$$\text{PSNR} = 10\log_{10}\left(\frac{255^2}{MSE}\right) \; dB \quad (10)$$

$r(m,n),\ g(m,n),\ b(m,n)$ = Original samples of the R, G, B components.

$\hat{r}(m,n),\ \hat{g}(m,n),\ \hat{b}(m,n)$ = Reconstructed samples of the R, G, B components.

M = Number of rows.

N = Number of columns.

5. Conclusions

We have presented a system for color image compression based on the HDWTSVD [12]. Simulation results show that this system gives good results at both low and high bit rates. We can achieve even more compression at low bitrates by implementing an adaptive multistage VQ to encode the eigenvectors. An adaptive multistage VQ will help us to reduce the bitrate at low resolutions in the areas compressed by SVD while keeping the same bitrate and quality of tiles compressed by DWT.

References:

[1] J. W. Woods and S.D. O'Neil, "Subband Coding of Images," *IEEE Trans. ASSP*, vol. 34, pp.1278-1288, Oct. 1986.

[2] M. Vetterli and D. Le Gall, "Perfect Reconstruction FIR Filter Banks: Some Properties and Factorizations," *IEEE Trans. ASSP*, vol. 37, pp.1057-1071, July 1989.

[3] G. Strang and T. Nguyen, Wavelets and Filter Banks. Wellesley-Cambridge Press, 1996.

[4] I. Daubechies and W. Sweldens, "Factoring Wavelet Transforms into Lifting Steps," Princeton University, Sept. 1996.

[5] J.M. Shapiro, "Embeded Image Coding Using Zerotrees of Wavelets Coefficients," *IEEE Trans. SP*, vol. 41: 3445-3462, Dec. 1993.

[6] A. Said and W.A. Pearlman, "A New Fast and Efficient Coder Based on Set Partitioning in Hierarchical Trees," *IEEE Trans. CSVT*, vol. 6, pp.243-250, June 1996.

[7] Y. F. Syed, A Low Bit Rate Wavelet-Based Image Coder for Transmission Over Hybrid Networks. Doctoral Dissertation, UTA, 1999.

[8] S.A. Martucci, et al, "A Zerotree Wavelet Video Coder," *IEEE Trans. CSVT*, vol. 7, pp.109-118, Feb. 1997.

[9] A.K. Jain, Fundamentals of Digital Image Processing. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[10] C.M. Goldrick, W. Dowling, and A. Bury, "Image Coding Using the Singular Value Decomposition and Vector Quantization," in *Image Processing and its Applications*, pp.296-300, IEE, 1995.

[11] A. Dapena and S. Ahalt, "A Hybrid DCT-SVD Image-Coding Algorithm," *IEEE Trans. CSVT*, vol. 12, pp.114-121, Feb. 2002.

[12] H. Ochoa and K.R. Rao, "A New Modified Hybrid DWTSVD Coding System for Color Images," *WSEAS Transaction on Circuits and Systems*, vol. 4, pp. 1246-1253, Oct. 2005.

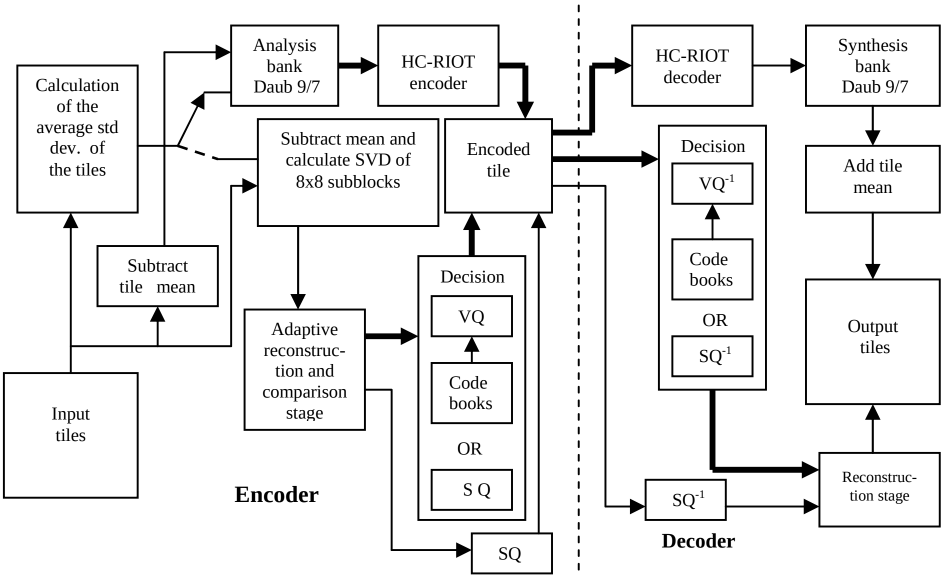

Figure 1. A New Coding System for Monochromatic Images (HDWTSVD).