Express Letters

Block Wavelet Transforms for Image Coding

A. Enis Çetin, Ömer N. Gerek, and Şennur Ulukuş

Abstract—In this paper, a new class of block transforms is presented. These transforms are constructed from subband decomposition filter banks corresponding to regular wavelets. New transforms are compared to the discrete cosine transform (DCT). Image coding schemes that employ the block wavelet transform (BWT) are developed. BWT's can be implemented by fast ($O(N \log N)$) algorithms.

I. INTRODUCTION

Frequency domain waveform coding methods— subband and block transform coding—are widely used in practice [1], [2]. In these methods one takes advantage of nonuniform distribution of bits to frequency components. In almost all transform coding methods [1], [3] the signal is first divided into blocks or vectors that are linearly transformed into another domain. The resultant set of transform domain coefficients is then quantized for transmission or storage. The efficiency of the transform coding system depends on the type of linear transform and the nature of bit allocation to the transform domain coefficients.

Recently, the relation between wavelet theory and subband decomposition has been established [4], [5], and it is shown that the wavelet orthonormal bases serve to provide a useful multiresolution signal representation. Corresponding to each wavelet orthonormal basis there exists a subband decomposition filter bank realizing the multiresolution signal representation. In this paper, it is shown that a new class of linear block transforms, block wavelet transforms (BWT's), can be obtained from subband based multiresolution signal decomposition. BWT's are constructed from subband filter banks corresponding to regular [4]–[7] wavelets. Image coding schemes that employ the BWT's are described.

II. BLOCK WAVELET TRANSFORM (BWT)

Let $H_0(\omega)$ and $H_1(\omega)$ be the low-pass and high-pass filters, respectively, of a perfect reconstruction filter bank. In a two-band partition, the input signal $x[n]$ is filtered by $h_0[n]$ and $h_1[n]$ and the resultant signals are downsampled by a factor of two. In this way two subsignals, $x_i[n] = \sum_k h_i[k]x[2n - k]$, $i = 0, 1$, are obtained. The subsignal $x_0[n](x_1[n])$ contains the low-pass (high-pass) information of the original signal $x[n]$. In many signal and image coding methods this signal decomposition operation is repeated in a tree-like structure, and the resultant subsignals are compressed by various coding schemes [1], [2].

Let us first construct a BWT for a vector of size $N = 2$. We assume that the filters $h_0[n]$ and $h_1[n]$ are FIR perfect reconstruction filter pairs [4]–[7] corresponding to regular wavelets. Let $x[n]$ be a finite extent signal of duration $N = 2$. The signals $x[n] * h_0[n]$ and $x[n] * h_1[n]$ have durations $K + N - 1 = K + 1$

if the durations of the filters, $h_0[n]$ and $h_1[n]$, are K. In spite of downsampling by two, $x_0[n]$ and $x_1[n]$ contain more than two nonzero samples in order to achieve perfect reconstruction for practical filter banks. Let us define the periodic sequence $\tilde{x}[n]$ as a periodic extension of $x[n]$, i.e., $\tilde{x}[n] = \sum_m x[n + Nm]$. The signals $\tilde{x}[n] * h_0[n]$ and $\tilde{x}[n] * h_1[n]$ are also periodic, with period $N = 2$. Therefore, the subsignals $\tilde{x}_0[n]$ and $\tilde{x}_1[n]$ are also periodic signals, with period $N/2 = 1$. We define $\tilde{x}_0[0]$ and $\tilde{x}_1[0]$ as BWT coefficients of the vector $x = [x[0] \ x[1]]^T$. Since convolution and downsampling operations are linear, the BWT is also a linear transform. Let $A_2 = [a_{ij}]_{2,2}$ be the transform matrix. The entries of the 2×2 BWT matrix A_2 can be determined by using the basis vectors $e_1 = [1 \ 0]^T$ and $e_2 = [0 \ 1]^T$. This is equivalent to applying the periodic signals $\tilde{e}_1[n] = \sum_m \delta[n + 2m]$ and $\tilde{e}_2[n] = \sum_m \delta[n - 1 + 2m]$ to the subband decomposition filter bank. If the input is $\tilde{e}_1$, then the output of the low-pass (high-pass) branch is $a_{11}(a_{21})$. Similarly, a_{12} and a_{22} are also obtained by using $\tilde{e}_2$.

Let us now construct the BWT matrix for $N = 2^l$ where l is a positive integer. Let $x[n]$ be a finite extent signal of length N and $x = [x[0] \ x[1] \ \cdots \ x[N - 1]]^T$ be a vector of size N. We consider an N band partition of the frequency domain. In order to achieve such a partition one can use an l stage subband decomposition, as shown in Fig. 1. If the finite signal $x[n]$ is applied to the structure of Fig. 1, then the subsignals $x_0[n], x_1[n], \cdots, x_{N-1}[n]$ contain more than N samples due to filtering in spite of downsampling by N. Let $\tilde{x}[n] = \sum_{m=-\infty}^{\infty} x[n + Nm]$. Since $\tilde{x}[n]$ is a periodic signal with period N, the subsignals $\tilde{x}_i$ are also periodic with period 1. As in the case of $N = 2$, we define the BW transform vector $\theta = [\theta_0, \theta_1 \ \cdots \ \theta_{N-1}]^T$ as follows:

$$\theta_p = \tilde{x}_p[0], \qquad p = 0, 1, \cdots, N - 1. \tag{1}$$

The jth column of the $N \times N$ transform matrix $A_N = [a_{ij}]_{N,N}$ can be obtained by applying $\tilde{e}_j[n] = \sum_m \delta[n - (j - 1) + Nm]$ as the input to the N-band subband decomposition structure of Fig. 1. In this case, the branch outputs are periodic signals with period 1 and the value at the ith branch is the (i, j)th entry a_{ij} of the matrix A_N.

BWT's obtained as shown above are orthogonal transforms, i.e., $A_N A_N^T = I_N$. Fast implementation of the transform can be carried out in l stages by using the structure of Fig. 1. Since at each stage half of the samples are dropped, and $l = \log N$, one can develop a fast algorithm to implement a BWT. It requires Order($N \log N$) (if the FIR filters, h_i, $i = 0, 1$ are Kth order filters, then $\min(K, N) \times N \log N$ in the most general case) multiplications to get N BWT coefficients. In most filter banks, filters h_0 and h_1 are related to each other. By taking advantage of this fact computational complexity can further be reduced.

We now present two BWT's constructed from the Haar orthogonal basis [5] and perfect reconstruction (PR) filter banks of Daubechies wavelets [4], [6].

Haar BWT: This transform is constructed from the PR filter bank: $h_0[0] = h_0[1] = 1/2$, $h_0[n] = 0$ for $n \neq 1, 2$, and $h_1[n] = (-1)^n h_0[n]$. This filter bank structure produces the Haar (wave-
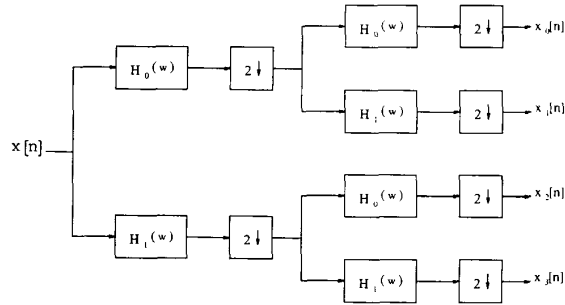
Fig. 1. $N = 4$ band ($l = 2$ stage) subband decomposition filter bank structure.

let) basis of $L^2(R)$ [5]. The l stage operation performed in every branch of Fig. 1 is equivalent to a single stage operation consisting of filtering the input by $h^l[n]$ whose frequency response is

$$H^k(\omega) = H_{i_0}(\omega)H_{i_1}(2\omega) \cdots H_{i_{N-1}}(2^{l-1}\omega),$$

$$k = 0, 1, \cdots, N - 1; i_j = 0, 1 \qquad (2)$$

where $k = i_0 2^{l-1} + i_1 2^{l-2} + \cdots + i_{N-1}$ and downsampling by $N = 2^l$. Since the filters h_0 and h_1 are second-order FIR filters, the filters $h^k[n]$, $l = 0, 1, \cdots, N - 1$ are Nth order filters. Because of this, the rows of the $N \times N$ transform matrix A_N are the impulse responses of the filters $h^k[n]$ and the transform matrix is nothing but the well-known Hadamard transform matrix [1].

Daubechies BWT: Daubechies filter banks [5] correspond to regular wavelets. For the filter pair

$$h_0[n] = \{ \cdots, 0, .230378, .714847, .630881, -.279838, -.187035,$$

$$.0308414, .0328830, -.0105974, 0, \cdots \}$$

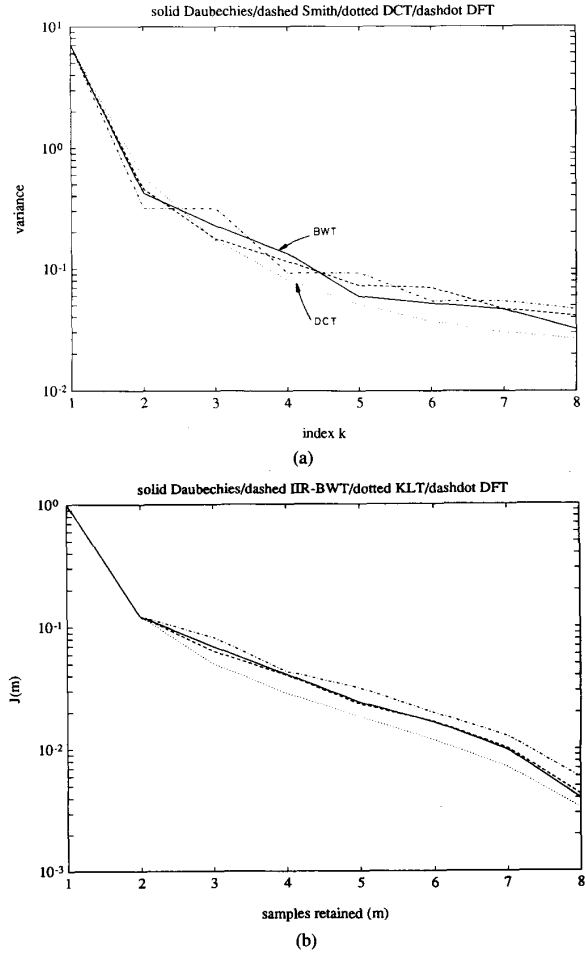and $h_1[n] = (-1)^n h_0[-n + 1]$, the BWT matrix A_8 is given as follows:

Fig. 2. (a) Distribution of variances of the transform coefficients (in decreasing order) of an AR(1) process with $\rho = 0.95$. (b) Performance of various transforms with respect to basis restriction error, $J(m) = \sum_{k=m}^{N-1} \sigma_k^2 / \sum_{k=0}^{N-1} \sigma_k^2$, versus the number of basis for an AR(1) process with $\rho = 0.95$.

$$A_8 = \begin{bmatrix} 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 \\ 0.0309 & 0.2904 & 0.5147 & 0.3870 & -0.0309 & -0.2904 & -0.5147 & -0.3870 \\ 0.3102 & 0.3921 & -0.3102 & -0.3921 & 0.3102 & 0.3921 & -0.3102 & -0.3921 \\ -0.5147 & -0.3870 & 0.0309 & 0.2904 & 0.5147 & 0.3870 & -0.0309 & -0.2904 \\ 0.3536 & -0.3536 & 0.3536 & -0.3536 & 0.3536 & -0.3536 & 0.3536 & -0.3536 \\ -0.3717 & 0.1485 & 0.3097 & -0.4938 & 0.3717 & -0.1485 & -0.3097 & 0.4938 \\ -0.3921 & 0.3102 & 0.3921 & -0.3102 & -0.3921 & 0.3102 & 0.3921 & -0.3102 \\ -0.3097 & 0.4938 & -0.3717 & 0.1485 & 0.3097 & -0.4938 & 0.3717 & -0.1485 \end{bmatrix} \qquad (3)$$

Note that the BWT matrix A_8 is orthogonal. One can also obtain block transforms from Smith and Barnwell (SB) filters [7] and from the Butterworth-IIR filter banks [8].

Two-dimensional BWT's are simply obtained in a separable manner. The transform U of the $N \times N$ matrix C is given as follows

$$U = A_N C A_N^T. \qquad (4)$$

BWT-based image (video) coders are developed just like the JPEG (MPEG) [3] image (video) coding standards. BWT coders are obtained by replacing the DCT unit of the JPEG coder with a BWT unit.

III. SIMULATION EXAMPLES AND CONCLUSIONS

Consider an $AR(1)$ random process with $\rho = 0.95$. Fig. 2(a) [Fig. 2(b)] shows the distribution of variances (basis restriction error [1]) of the transform coefficients for DCT, DFT, BWT's obtained from Butterworth-IIR, Daubechies, and SB filter banks. In this case, it is well known that the performance of DCT is very close to KL transform [1]. It can be observed from Fig. 2 that the performance of Daubechies BWT is the second best. Comparable performance can be obtained for the regular wavelet and subband filter-based decompositions.

Fig. 3. The test image Barbara (672 × 560 with 8 b/pel).

(a)

(b)

Fig. 4. (a) Details of the low-resolution image recovered from the first 4 × 4 DCT coefficients of the 8 × 8 BW transformed Barbara. (b) Details of the low-resolution image recovered from the first 4 × 4 DCT coefficients of the 8 × 8 discrete cosine transformed Barbara. Aliasing effects are more disturbing than Figure 4(a) (see the scarf of Barbara).

Consider the test image Barbara (672 × 560 with 8 b/pel) shown in Fig. 3. The Barbara image is coded by both an implementation of the JPEG standard and a BWT coder. The BWT coder is actually the JPEG coder, which employs a BWT instead of the DCT. In both coding schemes, 8 × 8 image subblocks and the default weighting matrix of the JPEG standard [3] are used. The JPEG (IIR-BWT) [Daubechies-BWT] coder compressed the image to 1.349 (1.44) [1.487] b/pel. All of the coded images are visually indistinguishable from the original one (figures are omitted).

A 1-D DCT (Butterworth-IIR-BWT) [Daubechies-BWT] {SB-BWT} of size $N = 8$ can be implemented by performing 12 (8) [22] {62} multiplications.

In many image coding applications, scalability of the coded bit stream is a desirable property [9]. If a low-resolution signal can be recovered from the bit stream, it can be displayed in low-resolution display. The low-resolution images (336 × 280) recovered from the first 4 × 4 BWT coefficients are compared to the low-resolution signal that can be extracted from a DCT-based scheme. In Fig. 4 the low-resolution images are shown. It is clear that the low-resolution image obtained from BWT [Fig. 4(a)] is better than the one obtained from DCT [Fig. 4(b)]. Aliasing effects are more disturbing in Fig. 4(b). If a scalable coded bit stream is desired, then the BWT is more suitable than DCT. One can extract better quality low-resolution images in BWT-coded bit streams than in DCT-coded ones.

From any given PR filter bank one can construct a linear block transform. Since the class of PR filter banks is quite large, the corresponding BWT class is also large. The choice of BWT for a given application remains as an interesting problem.

REFERENCES

[1] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
[2] J. W. Woods, Ed., *Subband Image Coding*. Norwood, MA: Kluwer, 1991.
[3] Didier Le Gall, "Digital multimedia systems: Digital image and video standards," *Commun. Ass. Comput. Mach.*, vol. 34, pp. 47–58, 1991.
[4] I. Daubechies, "Orthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.* vol. 41, pp. 909–996, 1988.
[5] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 2, pp. 674–693, July 1989.
[6] R. Ansari, C. Guillemot, and J. F. Kaiser, "Wavelet construction using Lagrange halfband filters," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 1116–1118, 1991.
[7] M. J. T. Smith and T. Barnwell, "Exact reconstruction techniques for tree-structured subband coders," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 434–441, June 1986.
[8] R. Ansari and B. Liu, "A class of low-noise computationally efficient recursive digital filters with applications to sampling rate alterations," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 90–97, Feb. 1985.
[9] C. Gonzales and E. Viscito, "Flexibly scalable digital video coding," *Image Commun.*, submitted for publication.

Vector Quantization of Images Using Input-Dependent Weighted Square Error Distortion

Anamitra Makur

I. INTRODUCTION

The measure of quantization distortion in a vector quantization (VQ) [1] coder for image is vital to the perceptual coder performance. The distortion function should ideally quantify human visual discomfort towards quantization errors. Mean