A Progressive Transmission Image Coder Using Linear Phase Uniform Filterbanks as Block Transforms

Trac D. Tran and Truong Q. Nguyen

***Abstract*— This paper presents a novel image coding scheme using M-channel linear phase perfect reconstruction filterbanks (LPPRFB's) in the embedded zerotree wavelet (EZW) framework introduced by Shapiro [1]. The innovation here is to replace the EZW's dyadic wavelet transform by M-channel uniform-band maximally decimated LPPRFB's, which offer finer frequency spectrum partitioning and higher energy compaction. The transform stage can now be implemented as a block transform which supports parallel processing mode and facilitates region-of-interest coding/decoding. For hardware implementation, the transform boasts efficient lattice structures, which employ a minimal number of delay elements and are robust under the quantization of lattice coefficients. The resulted compression algorithm also retains all attractive properties of the EZW coder and its variations such as progressive image transmission, embedded quantization, exact bit rate control, and idempotency. Despite its simplicity, our new coder outperforms some of the best image coders published recently in literature [1]–[4], for almost all test images (especially natural, hard-to-code ones) at almost all bit rates.**

***Index Terms*— Block transform coding, image coding, filterbanks, wavelet transform.**

I. INTRODUCTION

BLOCK transform coding and subband coding have been two dominant techniques in existing image compression standards and implementations. Both methods actually exhibit many similarities: relying on a certain transform to convert the input image to a more decorrelated representation, then utilizing the same basic building blocks such as bit allocator, quantizer, and entropy coder to achieve compression.

Block transform coders enjoyed success first due to their low complexity in implementation and their reasonable performance. The most popular block transform coder leads to the current image compression standard JPEG [5] which utilizes the 8×8 discrete cosine transform (DCT) [6] at its transformation stage. At high bit rates (1 b/pixel and up), JPEG offers almost visually lossless reconstruction image quality. However, when more compression is needed (i.e., at lower bit rates), annoying blocking artifacts show up because of two reasons: 1) the DCT bases are short, nonoverlapped, and have discontinuities at the ends and 2) JPEG processes each image block independently. So, interblock correlation has been completely abandoned.

The development of the lapped orthogonal transform (LOT) [7] and its generalized versions: the lapped biorthogonal transform (LBT) [8], the generalized LOT (GenLOT) [9]–[11], and the generalized LBT (GLBT) [12]–[14] helps solve the blocking problem by borrowing pixels from the adjacent blocks to produce the transform coefficients of the current block. It has long been recognized that lapped transforms belong to a subclass of linear phase perfect reconstruction filter banks (LPPRFB's): M-channel systems with filter lengths $N > M$ [7]. The DCT has M channels and all filters of length M; hence, its basis functions do not overlap. Lapped transform outperforms the DCT on two counts: 1) from the analysis viewpoint, it takes into account interblock correlation, hence, provides better energy compaction that leads to more efficient entropy coding of the coefficients and 2) from the synthesis viewpoint, its basis functions decay asymptotically to zero at the ends, reducing blocking discontinuities drastically. However, earlier lapped-transform-based image coders [7], [10], [15] have not utilized global information to their full advantage: the quantization and the entropy coding of transform coefficients are still independent from block to block.

Recently, subband coding has emerged as the leading standardization candidate in future image compression systems thanks to the development of the discrete wavelet transform. Wavelet representation with implicit overlapping and variable-length basis functions produces smoother and more perceptually pleasant reconstructed images. Moreover, wavelet's multiresolution characteristics have created an intuitive foundation on which simple, yet sophisticated, methods of encoding the transform coefficients are developed. Exploiting the relationship between the parent and the offspring coefficients in a wavelet tree, progressive wavelet coders [1]–[3] can effectively order the coefficients by bit planes and transmit more significant bits first. This coding scheme results in an embedded bit stream along with many other advantages such as exact bit rate control and near-idempotency (perfect idempotency is obtained when the transform can map integers to integers). In these subband coders, global information is taken into account fully.

Manuscript received June 4, 1997; revised March 4, 1998. This work was supported in part by the National Science Foundation under Grant MIP-9626563. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Christine Podilchuk.

T. D. Tran was with the Department of Electrical and Computer Engineering, University of Wisconsin, Madison, WI 53706 USA. He is now with the Department of Electrical and Computer Engineering, The Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: ttran@ece.jhu.edu).

T. Q. Nguyen is with the Electrical and Computer Engineering Department, Boston University, Boston, MA 02215 USA.

Publisher Item Identifier S 1057-7149(99)08747-3.

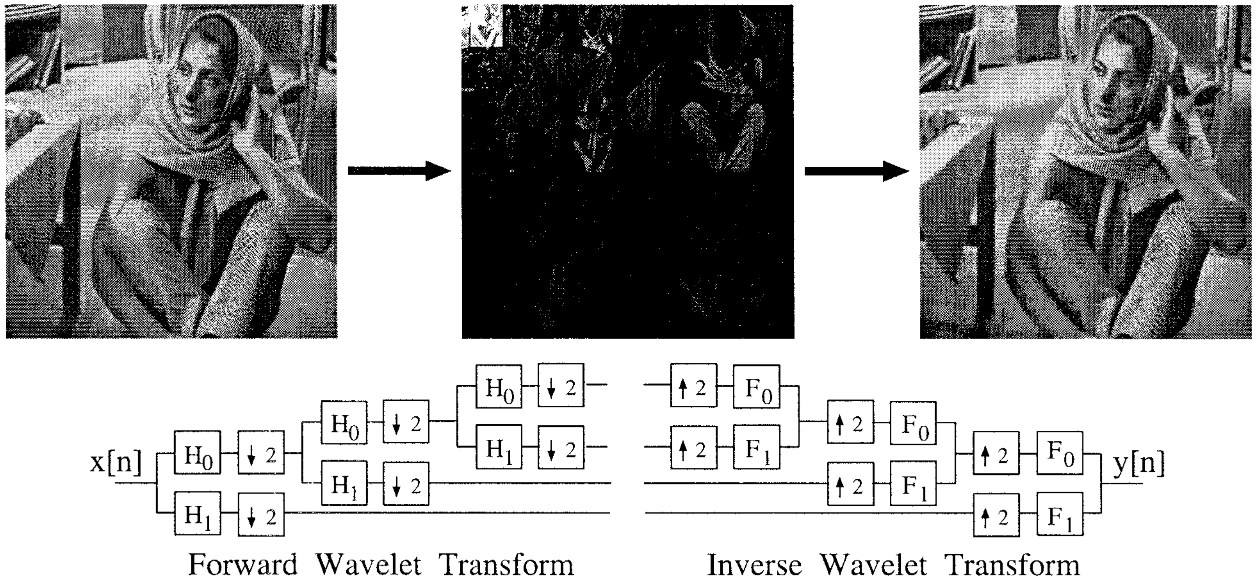


Fig. 1. Dyadic wavelet transform and its corresponding image decomposition.

From a frequency domain point of view, the wavelet transform simply provides an octave-band representation of signals. The dyadic wavelet transform can be thought of as a nonuniform-band lapped transform. It can sufficiently decorrelate smooth images; however, it has problems with images with well-localized high-frequency components—having low energy compaction in these cases. Several solutions have been proposed to solve this problem, most notably the wavelet packet approach [16] which yielded significant gains in objective performance.

In this work, we shall confirm that the embedded framework is not only limited to the wavelet transform and the wavelet packet; it can be utilized with multichannel uniform-band filterbanks as well. In fact, a judicious choice of M-channel LPPRFB coupled with several levels of wavelet decomposition of the dc band can provide much finer frequency spectrum partitioning, leading to significant improvement over current wavelet coders. This paper also attempts to shed some light onto a deeper understanding of wavelets, lapped transforms, their relation, and their performance in image compression from a multirate filterbank perspective.

The outline of the paper is as follows. In Section II, we offer a brief review of LPPRFB—its lattice structures and its implementation as block transform—as well as the wavelet transform and its role in progressive image coding. Section III describes the wavelet—block transform analogy, which leads to a general zerotree data structure for block transform coefficients and details of the design of high-performance transforms to take full advantage of the new coding scheme. Section IV presents many coding examples to confirm the validity of the theory. An extensive performance comparison between the new coder and existing state-of-the-art ones is also included. Finally, the conclusions are drawn in Section V.

A. Notations

Boldfaced characters are used to denote vectors and matrices. $\mathbf{A}^T$ and $\mathbf{A}^{-1}$ denote, respectively, the transpose and the inverse of the matrix $\mathbf{A}$. Special matrices used extensively are the identity matrix $\mathbf{I}$, the reversal matrix $\mathbf{J}$, and the null matrix $\mathbf{0}$. When the size of a matrix is not clear from context, subscripts will be included to indicate its size. For example, $\mathbf{J}_M$ denotes the $M \times M$ reversal matrix, and $\mathbf{0}_{M\times N}$ stands for the $M \times N$ null matrix. The impulse response of a filter and its discrete z-transform are represented by $h[n]$ and $H(z)$. For abbreviations, we use LP, PR, PU, and FB to denote, respectively, *linear phase, perfect reconstruction, paraunitary* (or *orthogonal*), and *filterbanks*. The letters L, M, N are reserved for the number of wavelet decomposition levels, the number of channels, and the filter's (or the input's) length, respectively. The terms *LPPRFB, block transform,* and *lapped transform* are used interchangeably in the paper.

II. Review

A. The Wavelet Transform and Progressive Image Transmission

From a filterbank viewpoint, the wavelet transform is an octave-band representation for signals; the discrete dyadic wavelet transform can be obtained by iterating on the lowpass output of a PR two-channel filterbank with enough regularity [17]–[19] as shown in Fig. 1. For a true wavelet decomposition, one iterates on the lowpass output only, whereas for a wavelet-packet decomposition, one may iterate on any output. The wavelet transform is, intuitively, a multiresolution decomposition of a signal into its coarse and detailed components. In the case of images, the wavelet representation is well-matched to psychovisual models, and it has given rise to numerous compression systems with superior subjective and objective quality to others at medium and high compression ratios [19], [20].

Many of these aforementioned high-performance wavelet coders also offer the capability of progressive image transmission. This coding approach relies on the fundamental idea that more important information (defined here as what decreases a

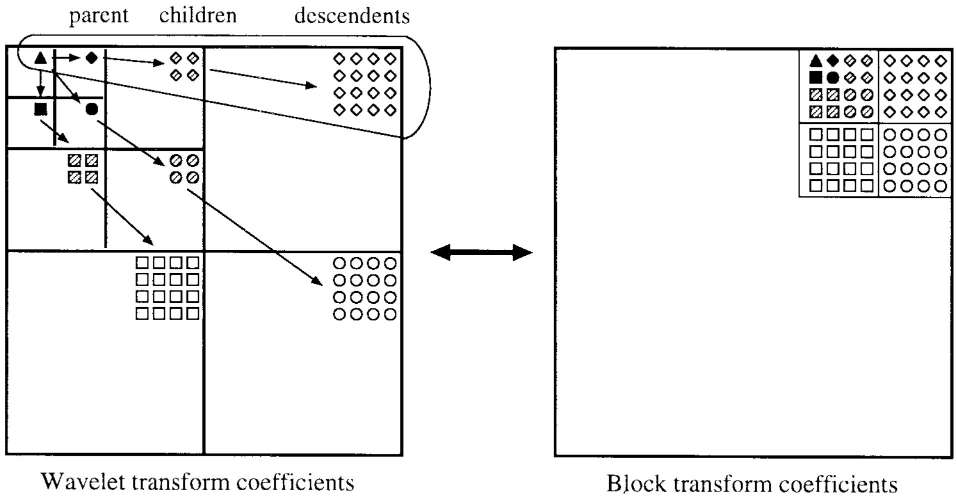


Fig. 2. Wavelet and block transform analogy.

certain distortion measure the most) should be transmitted first. Assume that the distortion measure is the mean-squared error (MSE), the transform is paraunitary, and transform coefficients $C_{i,j}$ are transmitted one by one, it can be proven that the mean squared error decreases by $|C_{i,j}|/N$, where N is the total number of pixels [21]. Therefore, larger coefficients should always be transmitted first. If one bit is transmitted at a time, this approach can be generalized to ranking the coefficients by bit planes and the most significant bits are transmitted first [22]. The progressive transmission scheme results in an embedded bit stream (i.e., it can be truncated at any point by the decoder to yield the best corresponding reconstructed image). The algorithm can be thought of as an elegant combination of a scalar quantizer with power-of-two stepsizes and an entropy coder to encode wavelet coefficients.

Embedded algorithm relies on the hierarchical coefficients' tree structure called a *wavelet tree*—a set of wavelet coefficients from different scales that belong in the same spatial locality. This is demonstrated in Fig. 2(a), where the tree in the vertical direction is circled. All of the coefficients in the lowest frequency band make up the *dc band* or the *reference signal* (located at the upper left corner). Besides these dc coefficients, in a wavelet tree of a particular direction, each lower-frequency *parent node* has four corresponding higher-frequency *offspring nodes*. All coefficients below a parent node in the same spatial locality is defined as its *descendents*. Define a coefficient $C_{i,j}$ to be *significant* with respect to a given threshold T if $|C_{i,j}| \geq T$, and *insignificant* otherwise. Meaningful image statistics have shown that if a coefficient is insignificant, it is very likely that its offspring and descendents are insignificant as well. Exploiting this fact, the most sophisticated embedded wavelet coder SPIHT can output a single binary marker to represent very efficiently a large, smooth image area (an insignificant tree). For more details on the algorithm, the reader is referred to [1]–[3].

Although the wavelet tree provides an elegant hierarchical data structure which facilitates quantization and entropy coding of the coefficients, the efficiency of the coder still depends heavily on the transform's ability in generating zerotrees. For nonsmooth images that contain a lot of texture and edges, the wavelet transform is not as efficient in signal decorrelation comparing to well-designed multichannel LPPRFB's which we shall prove to provide finer frequency selectivity and superior energy compaction.

B. Linear Phase Perfect Reconstruction Filterbanks

Two equivalent representations of an M-channel filter bank are depicted in Fig. 3 [17], [18]. In this paper, we only consider filterbanks with the following properties: perfect reconstruction, linear phase, finite impulse response (FIR), real coefficient, maximally decimated, and uniform-band. Here are several of our justifications.

- The PR property is highly desirable since it provides a lossless signal representation and it simplifies the error analysis significantly.
- In image processing, it is also crucial that all analysis and synthesis filters have linear phase. Besides the elimination of the phase distortion, linear phase systems allow us to use simple symmetric extension methods to accurately handle the boundaries of finite-length signals. Furthermore, the linear phase property can be exploited, leading to faster and more efficient FB implementation.
- The filter length should be relatively short to prevent ringing artifacts in the reconstructed images and to keep the transform fast.
- Since our interest is on compression, especially at low bit rates, we prefer maximally decimated FB's which do not expand the input signals.

From the polyphase representation in Fig. 3(b), perfect reconstruction can be loosely defined as the existence of an FIR matrix $\mathbf{R}(z)$ that satisfies the equation

$$\mathbf{R}(z)\mathbf{E}(z) = bz^{-l}\mathbf{I}, \qquad b \neq 0, \, l \geq 0. \tag{1}$$

We call these systems *biorthogonal*. An important subset of PRFB is *paraunitary* FB, where $\mathbf{R}(z)$ is chosen to be

$$\mathbf{R}(z) = z^{-K}\mathbf{E}^T(z^{-1}) \tag{2}$$

with K being the order of $\mathbf{E}(z)$.

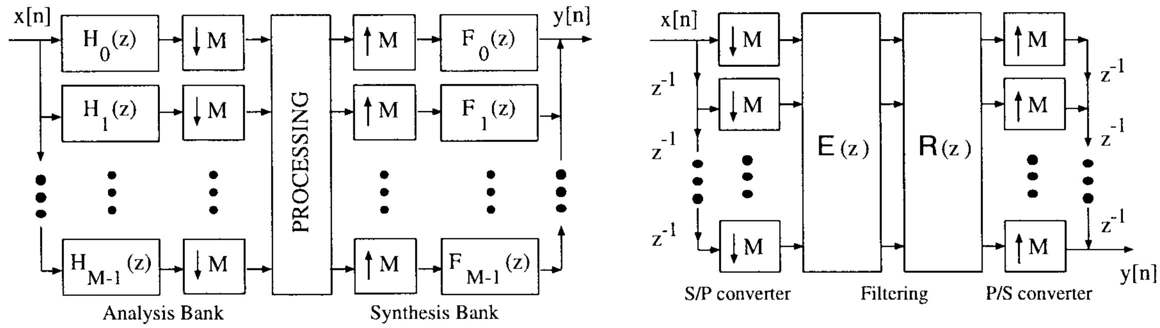


Fig. 3. Two representations of an M-channel uniform-band maximally decimated filterbank.

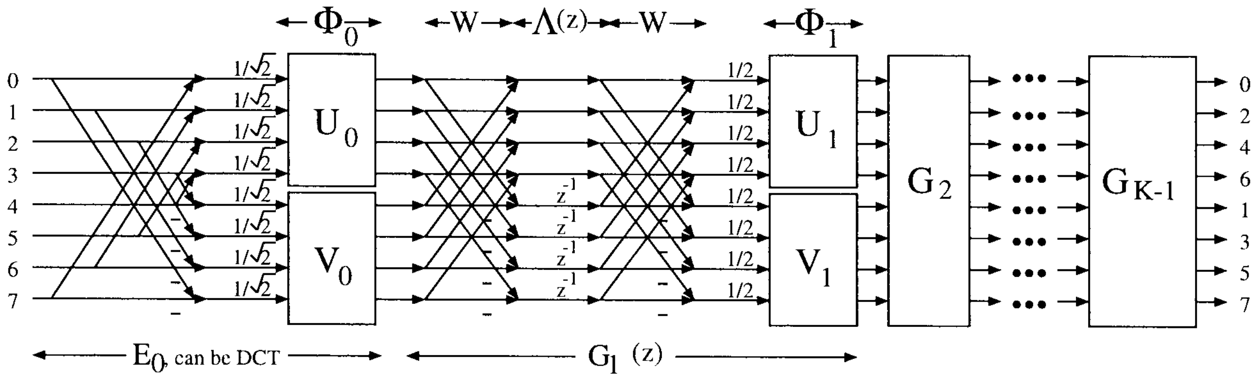


Fig. 4. Lattice structure for LPPRFB.

LPPRFB's have long found application in transform-based image coding. The DCT is an eight-channel eight-tap LPPUFB. A popular extension of the DCT is the LOT, an even-channel $2M$-tap LPPUFB that can be interpreted as an overlapping block transform. Rather than processing one block independently from the next like DCT, LOT has overlapping input windows and it elegantly solves the blocking problem in DCT-based coders by partly smoothing out the block boundaries. To reduce blocking effect further, longer overlaps might be needed. This motivates the development of the generalized lapped orthogonal transform (GenLOT) [9]–[11] and its biorthogonal versions (GLBT) [12]–[14].

Every FB presented in this paper has an efficient lattice structure that retains both LP and PR properties under quantization of lattice coefficients. The key idea behind the lattice structure is the factorization of the filterbank's polyphase matrix $\mathbf{E}(z)$. Let $H_k(z)$ and $F_k(z)$ be the analysis and synthesis filters of length $N = MK$ in an M-channel LPPUFB. If M is even, it is necessary to have $M/2$ symmetric and $M/2$ antisymmetric filters [11]. Define

$$\mathbf{\Phi}_i = \begin{bmatrix} \mathbf{U}_i & \mathbf{0}_{M/2} \\ \mathbf{0}_{M/2} & \mathbf{V}_i \end{bmatrix}$$

where $\mathbf{U}_i$ and $\mathbf{V}_i$ are arbitrary $M/2 \times M/2$ invertible matrices, and

$$\mathbf{W} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_{M/2} & \mathbf{I}_{M/2} \\ \mathbf{I}_{M/2} & -\mathbf{I}_{M/2} \end{bmatrix}, \qquad \mathbf{\Lambda} = \begin{bmatrix} \mathbf{I}_{M/2} & \mathbf{0}_{M/2} \\ \mathbf{0}_{M/2} & z^{-1}\mathbf{I}_{M/2} \end{bmatrix}.$$

Then, the polyphase matrix $\mathbf{E}(z)$ can always be factored as follows [10], [14]:

$$\mathbf{E}(z) = \mathbf{G}_{K-1}(z)\mathbf{G}_{K-2}(z) \cdots \mathbf{G}_1(z)\mathbf{E}_0 \tag{3}$$

where $\mathbf{G}_i(z) = \mathbf{\Phi}_i \mathbf{W} \mathbf{\Lambda}(z) \mathbf{W}$, and

$$\mathbf{E}_0 = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{U}_0 & \mathbf{U}_0\mathbf{J} \\ \mathbf{V}_0\mathbf{J} & -\mathbf{V}_0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{U}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_0 \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{J} \\ \mathbf{J} & -\mathbf{I} \end{bmatrix}.$$

Again, $\mathbf{U}_0$ and $\mathbf{V}_0$ are arbitrary $M/2 \times M/2$ invertible matrices. For fast implementations, $\mathbf{E}_0$ can be replaced by the DCT. The complete lattice structure is shown in Fig. 4.

It is clear from (3) that each stage of the biorthogonal FB [either $\mathbf{G}_i(z)$ or $\mathbf{E}_0$] contains two arbitrary invertible matrices of size $M/2$. These matrices contain the free parameters, or the degrees of freedom, that can be used to fine-tune the FB in the design process. If an orthogonal FB is desired, the free matrices are restricted to be orthogonal. In this case, the $\mathbf{U}_i$ and $\mathbf{V}_i$ orthogonal matrices can be factored further into plane rotations as depicted in Fig. 5(a). In the more general biorthogonal case, the free matrices are only required to be invertible, and we can always decompose them using the singular value decomposition (SVD) as $\mathbf{U}_i = \mathbf{U}_{i0}\mathbf{\Gamma}_i\mathbf{U}_{i1}$ where $\mathbf{U}_{i0}$ and $\mathbf{U}_{i1}$ are orthogonal matrices with rotation angles θ_i, and $\mathbf{\Gamma}_i$ is a diagonal matrix with positive elements α_i as illustrated in Fig. 5(b). Invertibility is guaranteed structurally under a mild condition—as long as none of the diagonal lattice coefficients α_i representing $\mathbf{\Gamma}_i$ is quantized to zero. Unconstrained optimization can now be used to optimize the FB whose rotation angles (and the diagonal multipliers

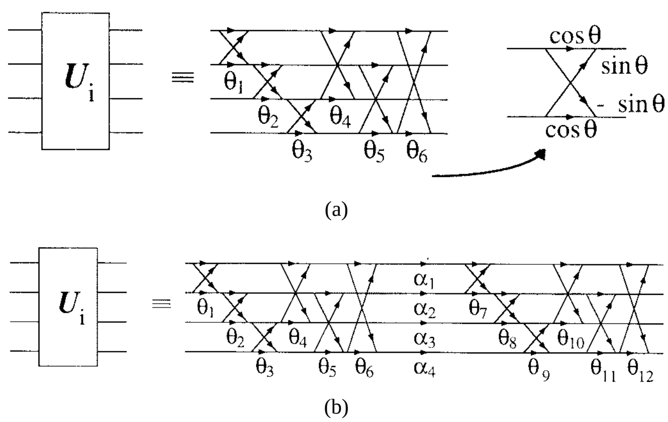


Fig. 5. Matrix parameterization: (a) orthogonal and (b) invertible.

in the biorthogonal case) are allowed to vary freely and independently.

Conceptually, each of the aforementioned LPPRFB's can be implemented as a block transform directly as depicted in Fig. 6. In the one-dimensional (1-D) direct implementation, the input signal can be blocked into sequences $x_i[n]$ of length $N = KM$, overlapped by $M(K-1)$ samples with adjacent sequences. The M columns of the transform coefficient matrix $\mathbf{P}$ hold the impulse responses of the analysis filters $h_i[n]$. The resulting M transform coefficients $X_i[n]$ can then be quantized, coded, and transmitted to the decoder where the inverse transform is performed to reconstruct the original sequences $x_i[n]$. The long basis functions which decay smoothly to zero, coupled with overlapping data blocks, have the ability to reduce blocking artifacts at high compression ratios. In some cases as we demonstrate later, blocking can be completely eliminated. In the two-dimensional (2-D) case, the input block size is $N \times N$ and the output block size is $M \times M$. Like the 2-D separable wavelet transform, filtering is applied vertically first and then horizontally (or vice versa).

III. DESIGN

A. Zerotree Data Structure for Block Transform

One of the original contributions of this paper is the novel usage of M-channel uniform LPPRFB as a replacement for the dyadic wavelet transform at the transformation stage of a zerotree coder. Instead of obtaining an octave-band signal decomposition, one can have a finer uniform-band partitioning as depicted in Fig. 7 (drawn for $M = 8$). The finer frequency partitioning increases the frequency resolution that can often generate more insignificant coefficients, leading to an enhancement in the performance of the zerotree algorithm. However, uniform filterbank also has uniform downsampling (all subbands now have the same size). A parent node would not have four offspring nodes as in the case of the wavelet representation. How would one come up with a new tree structure that still takes full advantage of the interscale correlation between the transform coefficients?

The above question can be answered by investigating an analogy between the wavelet and the block transform as illustrated in Fig. 2. The parent, the offspring, and the descendents in a wavelet tree cover the same spatial locality, and so do the coefficients of a transform block. In fact, a wavelet tree in an L-level decomposition is analogous to a 2^L-channel transform's coefficient block. The difference lies at the bases that generate these coefficients. It can be shown that a 1-D L-level wavelet decomposition, if implemented as a lapped transform, has the following coefficient matrix

$$\mathbf{P}_L = \begin{bmatrix} h_0[n] * h_0\left[\frac{n}{2}\right] * \cdots * h_0\left[\frac{n}{2^{L-2}}\right] * h_0\left[\frac{n}{2^{L-1}}\right] \\ h_0[n] * h_0\left[\frac{n}{2}\right] * \cdots * h_0\left[\frac{n}{2^{L-2}}\right] * h_1\left[\frac{n}{2^{L-1}}\right] \\ h_0[n] * h_0\left[\frac{n}{2}\right] * \cdots * h_1\left[\frac{n}{2^{L-2}}\right] \\ h_0[n] * h_0\left[\frac{n}{2}\right] * \cdots * h_1\left[\frac{n}{2^{L-2}}\right] \\ \vdots \\ h_1[n] \\ h_1[n] \\ h_1[n] \\ h_1[n] \end{bmatrix}. \tag{4}$$

From the coefficient matrix $\mathbf{P}_L$, we can observe the following interesting and important characteristics of the wavelet transform through the block transform's prism.

- The wavelet transform can be viewed as a lapped transform with filters of variable lengths. For an L-level decomposition, there are 2^L filters.
- Each basis function has linear phase; however, they do not share the same center of symmetry.
- The block size is defined by the length of the longest filter. If $h_0[n]$ is longer and has length N_0, the top filter covering the dc component turns out to be the longest, and it has a length of $(2^L - 1)(N_0 - 1) + 1$. For the biorthogonal wavelet pair with $h_0[n]$ of length 9 and $h_1[n]$ of length 7 and three levels of decomposition, the eight resulting basis functions have respective lengths of 57, 49, 21, 21, 7, 7, 7, and 7.
- For a six-level decomposition using the same 9/7-tap pair, the length of the longest basis function grows to 505! The large number of overlapping pixels explains the smoothness of the reconstructed images where blocking artifacts are completely eliminated.

Each block of lapped transform coefficients represents a spatial locality similarly to a tree of wavelet coefficients as illustrated in Fig. 2. Let $\mathcal{O}(i, j)$ be the set of coordinates of all offspring of the node (i, j) in an M-channel block transform ($0 \le i, j \le M-1$), then $\mathcal{O}(i, j)$ can be represented as follows:

$$\begin{aligned} \mathcal{O}(i, j) = \{&(2i, 2j), (2i, 2j+1), (2i+1, 2j), \\ &(2i+1, 2j+1)\}. \end{aligned} \tag{5}$$

All (0, 0) coefficients from all transform blocks form the dc band, which is similar to the wavelet transform's reference signal, and each of these nodes has only three offspring: (0, 1), (1, 0), and (1, 1). The complete tree is now available locally, i.e., we do not have to search for the offspring across the

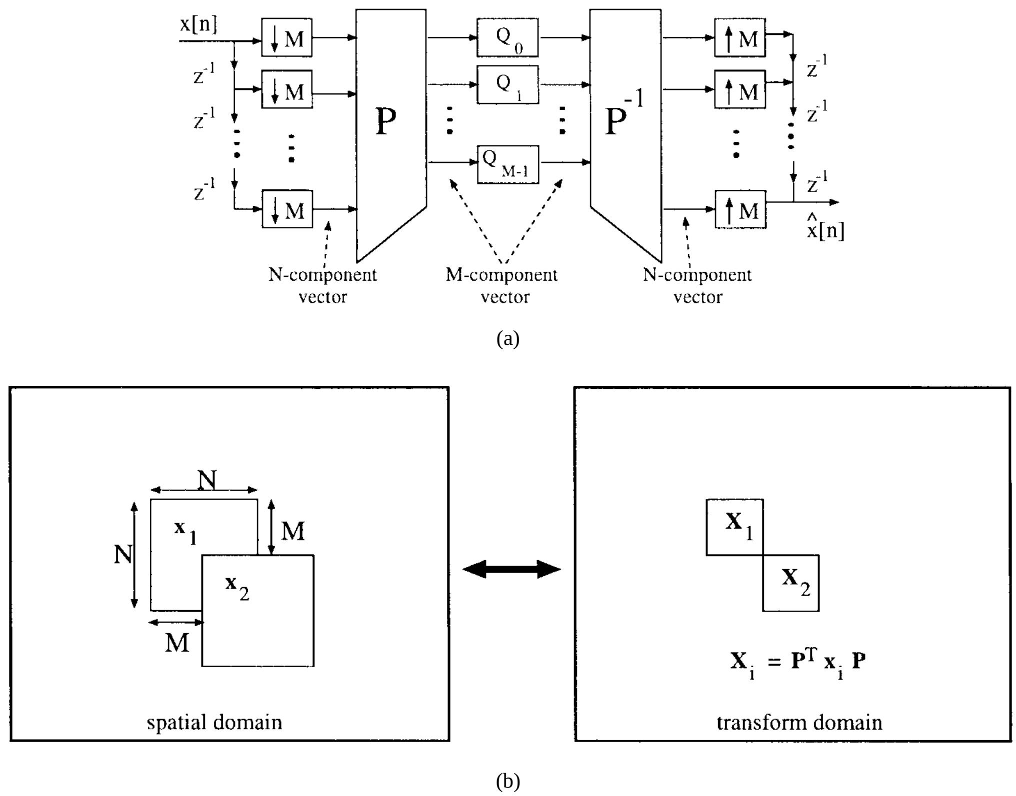


Fig. 6. M-channel LPPRFB as lapped transform: (a) direct 1-D implementation and (b) illustration in 2-D.

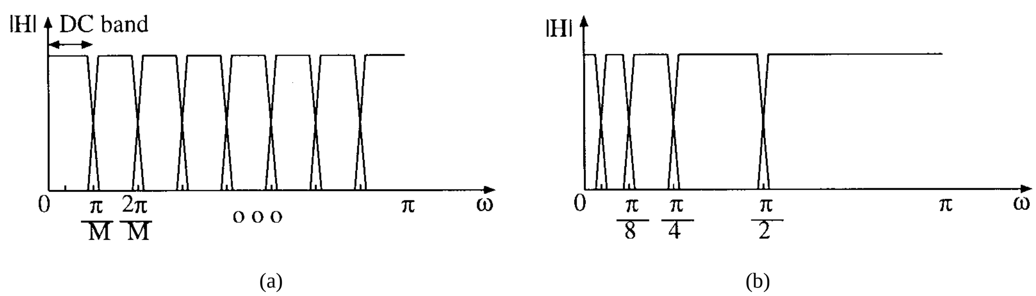


Fig. 7. Frequency spectrum partitioning: (a) M-channel uniform-band LPPRFB and (b) dyadic wavelet transform.

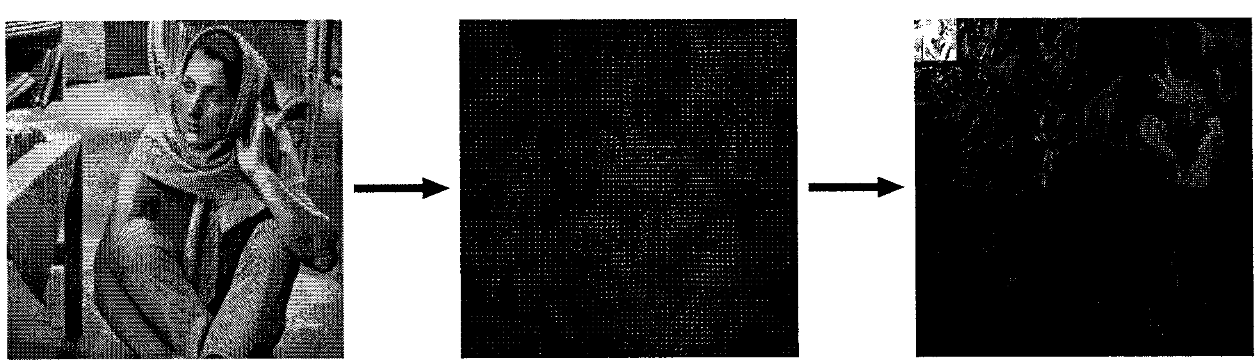


Fig. 8. Demonstration of the analogy between block transform and wavelet representation.

subbands anymore. This is a straightforward generalization of the structure first proposed for the 8×8 DCT in [23]. The only requirement here is that the number of channel M has to be a power of two.

Fig. 8 demonstrates through a simple rearrangement of the block transform coefficients that the redefined tree structure above does possess a wavelet-like multiscale representation. Table I compares the energy compaction level between the

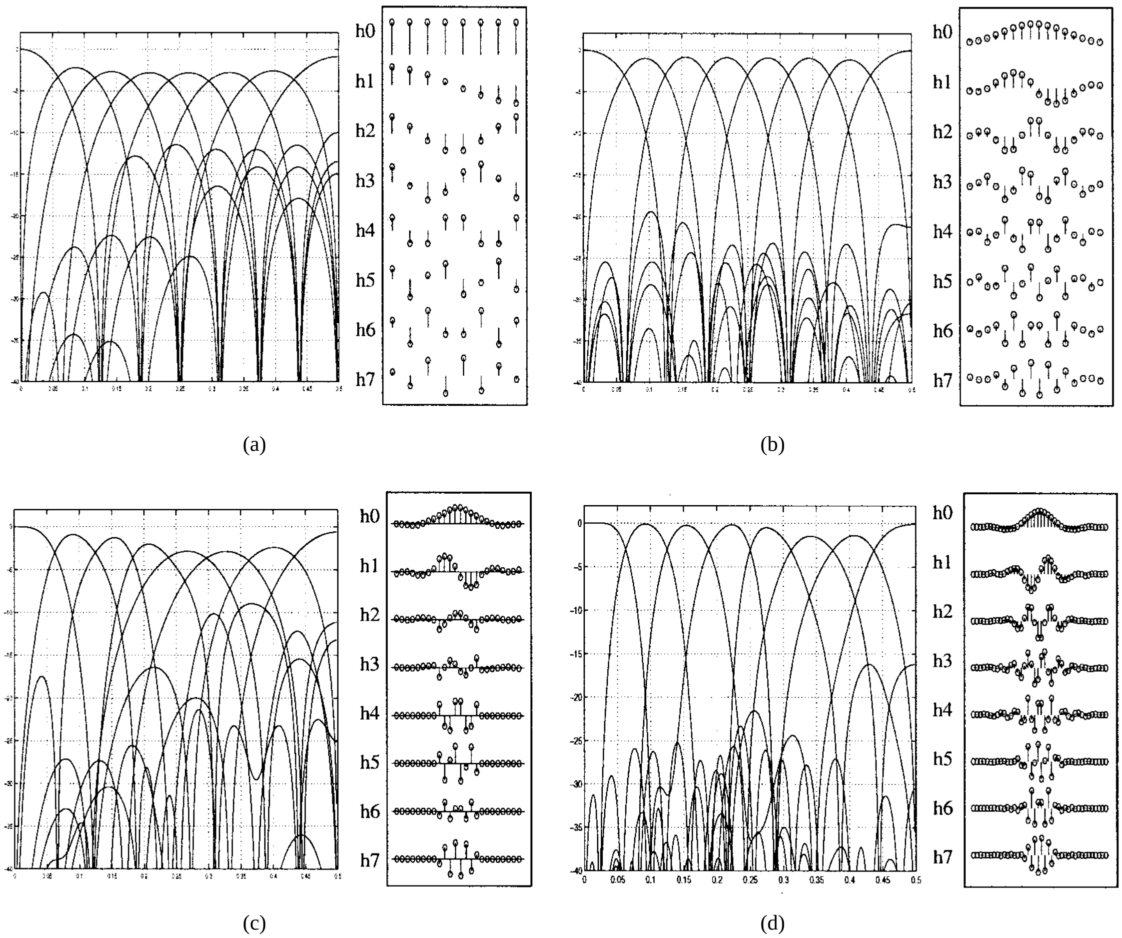


Fig. 9. Frequency and impulse responses of orthogonal transforms (a) 8 × 8 DCT. (b) 8 × 16 type-II LOT. (c) 4 × 24 4 × 8 VLLOT. (d) 8 × 40 GenLOT.

TABLE I
COMPARISON ON THE ENERGY COMPACTION LEVELS OF THE DYADIC WAVELET TRANSFORM AND UNIFORM LPPRFB'S ON THE BARBARA IMAGE

	Threshold T	Transform: 9/7 Wavelet 3-level	9/7 Wavelet 4-level	8 x 40 GenLOT	16 x 32 GLBT
% of coefficients C ij whose |Cij | < T	1	22.54	22.57	26.50	32.49
	2	40.58	40.64	46.26	52.56
	4	61.08	61.20	67.29	72.05
	8	75.88	76.10	81.67	84.63
	16	85.70	86.07	89.73	91.91
	32	92.30	92.86	94.32	96.06
	64	96.29	97.07	96.91	98.19
	128	98.15	99.12	98.27	99.20
	256	98.45	99.53	99.06	99.65
	512	98.72	99.61	99.58	99.86
	1024	99.31	99.67	100	99.99

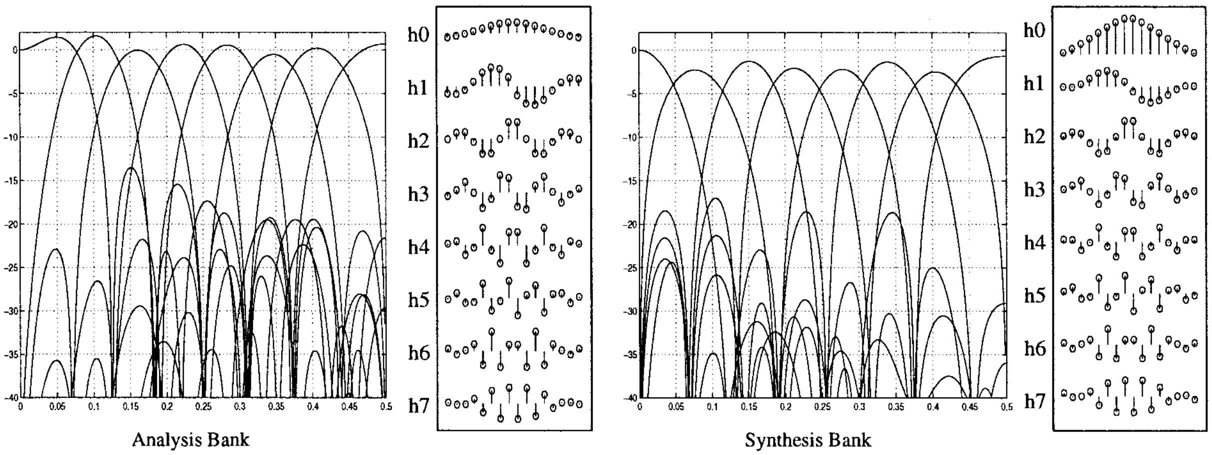
wavelet transform and two high-performance block transforms, the 8 × 40 GenLOT and the 16 × 32 GLBT (which will be presented in more details in the next section), for the Barbara image. The two block transforms consistently generate a higher percentage of small-value coefficients, hence creates a significant increase in the number of zerotrees. This holds the key to our coder's superior performance.

B. Transform Design Issues

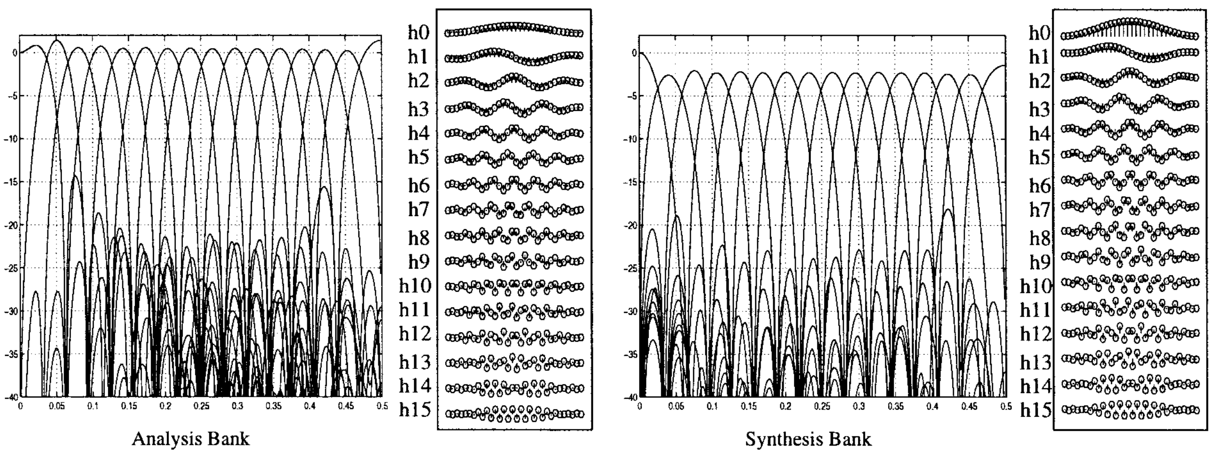
Transform-based coders rely on multirate filterbanks to generate the frequency coefficients that can be quantized and entropy coded. In the decoders, filterbanks are again used to combine and reconstruct the signal. Therefore, from our viewpoint, well-optimized filterbanks play an integral role in the coder's performance. As previously mentioned in Section II, any realization of the lattice coefficient set $\{\theta_i, \alpha_i\}$ in the previous section results in an LPPRFB. The degrees of freedom in the lattice coefficient set can be exploited to obtain other desirable properties for the FB's. The cost function used in this paper is a weighted linear combination of coding gain, dc leakage, attenuation around mirror frequencies, and stopband attenuation—all of which are well-known properties in yielding the best reconstructed image quality [18], [24]:

$$C_{\text{overall}} = k_1 C_{\text{coding gain}} + k_2 C_{\text{DC}} + k_3 C_{\text{mirror}} + k_4 C_{\text{analysis stopband}} + k_5 C_{\text{synthesis stopband}}. \quad (6)$$

1) Coding Gain: Coding gain is an approximate measure of the transform's energy compaction capability. Since our coder has progressive transmission, higher coding gain almost



(a)



(b)

Fig. 10. Frequency and impulse responses of biorthogonal transforms: (a) 8 × 16 GLBT and (b) 16 × 32 GLBT.

TABLE II
COMPARISON OF OBJECTIVE TRANSFORM PROPERTIES

Transform Property	Transform					
	8x8 DCT	8x16 LOT	4x24 4x8 VLLOT	8x40 GenLOT	8x16 GLBT	16x32 GLBT
Coding Gain (dB)	8.83	9.22	9.26	9.52	9.62	9.96
DC Attenuation (-dB)	310.62	312.56	322.10	322.10	327.40	303.32
Stopband Attenuation (-dB)	9.96	19.38	8.06	16.18	13.50	14.28
Mirror Freq. Attenuation (-dB)	322.10	317.24	318.58	317.24	55.54	302.35

always translates to higher image quality in the mean-squared sense. Transforms with higher coding gain tend to compact more energy into a fewer number of coefficients, and the more significant bits of those coefficients always get transmitted first. All new FB's presented in this paper are obtained with a version of the generalized coding gain formula [25]:

$$C_{\text{coding gain}} = 10 \log_{10} \frac{\sigma_x^2}{\left(\prod_{k=0}^{M-1} \sigma_{x_i}^2 ||f_i||^2 \right)^{1/M}} \tag{7}$$

where σ_x^2 is the variance of the input signal, $\sigma_{x_i}^2$ is the variance of the ith subband, and $||f_i||^2$ is the norm of the ith synthesis filter. The signal $x[n]$ is the commonly-used AR(1) process with intersample autocorrelation coefficient $\rho = 0.95$ [7]. For orthogonal (paraunitary) FB's, the synthesis filters are simply the time-reversed analysis filters. The reader should note that we have never attempted to optimize the FB's to match the statistics of any input image. The only image model employed in the design process is the AR(1) model. We do recognize that even higher coding performance can be achieved when

TABLE III
OBJECTIVE CODING RESULTS (PSNR IN DECIBELS): (a) LENA, (b) GOLDHILL, AND (c) BARBARA

Lena	Progressive Transmission Image Coders						
Comp. Ratio	SPIHT (9/7 WL)	8 x 8 DCT	8 x 16 LOT	8 x 40 GenLOT	4x24 4x8 VLLOT	8 x 16 GLBT	16 x 32 GLBT
1:8	40.41	39.91	40.05	40.43	40.18	40.35	40.43
1:16	37.21	36.38	36.72	37.32	36.85	37.28	37.33
1:32	34.11	32.90	33.56	34.23	33.61	34.14	34.27
1:64	31.10	29.67	30.48	31.16	30.48	31.04	31.18
1:100	29.35	27.80	28.62	29.31	28.62	29.14	29.38
1:128	28.38	26.91	27.61	28.35	27.64	28.19	28.39

(a)

Goldhill	Progressive Transmission Image Coders						
Comp. Ratio	SPIHT (9/7 WL)	8 x 8 DCT	8 x 16 LOT	8 x 40 GenLOT	4x24 4x8 VLLOT	8 x 16 GLBT	16 x 32 GLBT
1:8	36.55	36.25	36.63	36.80	36.49	36.69	36.78
1:16	33.13	32.76	33.18	33.36	33.06	33.31	33.42
1:32	30.56	30.07	30.56	30.79	30.51	30.70	30.84
1:64	28.48	27.93	28.36	28.60	28.35	28.58	28.74
1:100	27.38	26.65	27.09	27.40	27.10	27.33	27.62
1:128	26.73	26.01	26.48	26.79	26.46	26.71	26.96

(b)

Barbara	Progressive Transmission Image Coders						
Comp. Ratio	SPIHT (9/7 WL)	8 x 8 DCT	8 x 16 LOT	8 x 40 GenLOT	4x24 4x8 VLLOT	8 x 16 GLBT	16 x 32 GLBT
1:8	36.41	36.31	37.43	38.08	36.83	37.84	38.43
1:16	31.40	31.11	32.70	33.47	31.86	33.02	33.94
1:32	27.58	27.28	28.80	29.53	27.99	29.04	30.18
1:64	24.86	24.58	25.70	26.37	25.10	26.00	27.13
1:100	23.76	23.42	24.34	24.95	23.96	24.55	25.39
1:128	23.35	22.68	23.37	24.01	23.24	23.49	24.56

(c)

the transform is image-dependent. However, the coder's level of complexity is also significantly higher.

2) Low DC Leakage: The dc leakage cost function measures the amount of dc energy that leaks out to the bandpass and highpass subbands. The main idea is to concentrate all signal energy at dc into the dc coefficients. This proves to be advantageous in both signal decorrelation and in the prevention of discontinuities in the reconstructed signals. Low dc leakage can prevent the annoying checkerboard artifact that usually occurs when high-frequency bands are severely quantized [18]. This problem is more troublesome in traditional block transform coders because high-frequency bands are usually more coarsely quantized. The dc cost function is defined as

$$C_{\text{DC}} = \sum_{i=1}^{M-1} \sum_{n=0}^{L-1} h_i[n]. \tag{8}$$

Notice that all antisymmetric filters have a zero at dc. Therefore, the above formula only needs to apply to symmetric filters to reduce the complexity of the optimization process.

3) Attenuation at Mirror Frequencies: The mirror frequency cost function is a generalization of C_{DC}. The concern is now at every aliasing frequencies $\omega_m = 2\pi m/M$, $m \in \mathcal{Z}$, $1 \leq m \leq M/2$. Ramstad *et al.* show that frequency attenuation at mirror frequencies are very important in the further reduction of blocking artifacts: the filter responses should be small at these mirror frequencies as well [24]. The corresponding cost function is

$$C_{\text{mirror}} = \sum_{i=0}^{M-1} \left|H_i(e^{j\omega_m})\right|^2,$$
$$\omega_m = \frac{2\pi m}{M}, \qquad 1 \le m \le \frac{M}{2}. \tag{9}$$

Low dc leakage and high attenuation near the mirror frequencies are not as essential to the coder's objective performance as coding gain. However, they do improve the visual quality of the reconstructed image significantly.

4) Stopband Attenuation: Stopband attenuation of the filters is a classical performance criterion in filter design. In this paper, the stopband attenuation criterion measures the sum of all of the filters' energy outside the designated passbands:

$$C_{\text{analysis stopband}} = \sum_{i=0}^{M-1} \int_{\omega \in \Omega_{\text{stopband}}} \cdot W_i^a(e^{j\omega})\left|H_i(e^{j\omega})\right|^2 d\omega \tag{10}$$
$$C_{\text{synthesis stopband}} = \sum_{i=0}^{M-1} \int_{\omega \in \Omega_{\text{stopband}}} \cdot W_i^s(e^{j\omega})\left|F_i(e^{j\omega})\right|^2 d\omega. \tag{11}$$

In the LPPUFB case, $C_{\text{analysis stopband}} = C_{\text{synthesis stopband}}$. The biorthogonal FB's offer more flexibility. In the analysis bank, the stopband attenuation cost helps in improving the signal decorrelation and decreasing the amount of aliasing. In meaningful images, we know *a priori* that most of the energy is concentrated in low frequency region. Hence, high stopband attenuation in this part of the frequency spectrum becomes extremely desirable. In the synthesis bank, the reverse is true. Synthesis filters covering low-frequency bands need to have high stopband attenuation near and/or at $\omega = \pi$ to enhance their smoothness. The biased weighting can be enforced using two simple linear functions $W_i^a(e^{j\omega})$ and $W_i^a(e^{j\omega})$ as shown in (10) and (11).

5) Transforms with Variable-Length Basis Functions: The elegant factorization in Section II results in all filters of equal length. For images that contain a lot of strong edges, the long basis functions covering high-frequency bands can cause excessive ringing at low bit rates. On the other hand, the longer the filter becomes, the higher the complexity of the FB gets. Since blocking is most noticeable in smooth image regions, in order to reduce blocking artifacts, filters covering high-frequency bands do not need long overlapping windows. In fact, they may not have to be overlapped at all. If the filter length can be restricted mathematically, i.e., these coefficients are structurally enforced to exact zeros, the complexity of the resulting FB can be reduced significantly. Efforts to reduce ringing artifacts and to minimize the transform complexity can be found in [26]–[28]. This class of transforms represents low-frequency components by longer overlapped basis functions to prevent blocking, while reserving shorter basis functions for high-frequency components to minimize ringing. One of

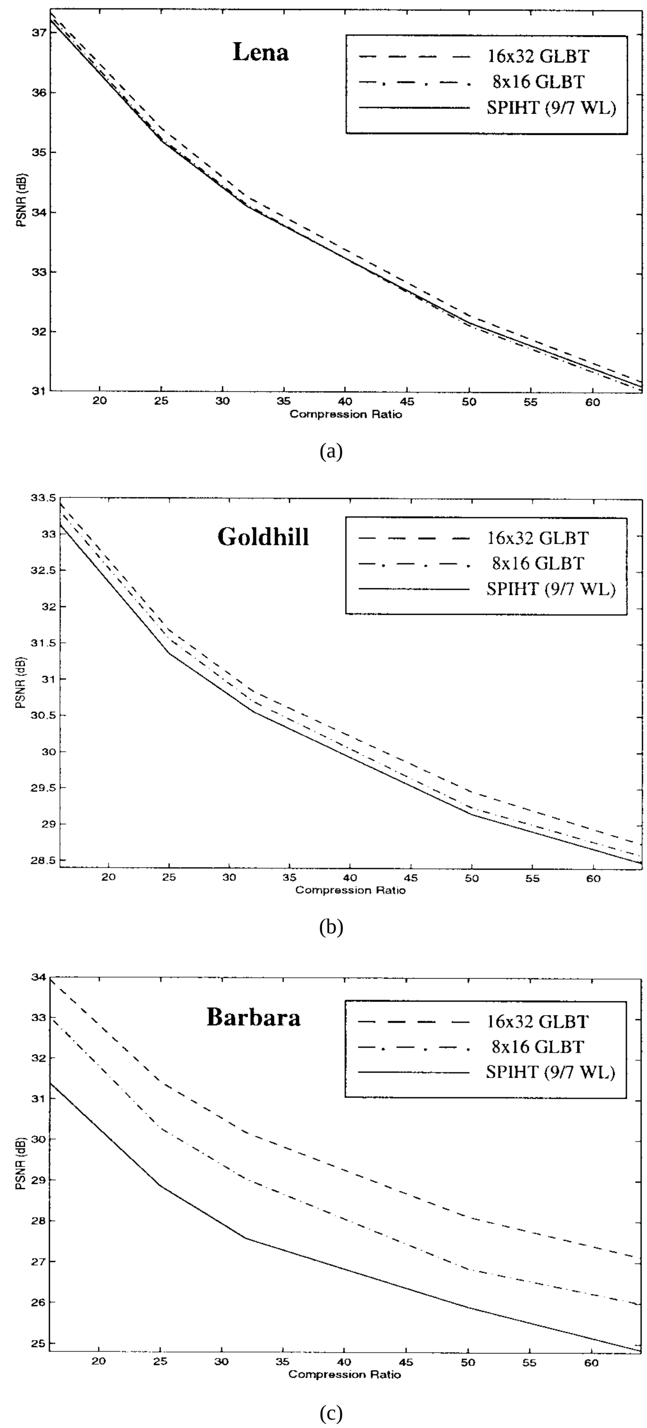


Fig. 11. Rate-distortion curves of image coding examples (a) Lena, (b) Goldhill, and (c) Barbara.

such transform named VLLOT with four 24-tap and four eight-tap filters is presented in this paper to validate the flexibility of the general zerotree coding scheme. Ringing can also be minimized by adding a time-constrained objective to high-frequency bandpass filters to force the tails of their impulse responses to have very small values (not necessarily zeros). This constraint does not limit the search space of the optimization routine, so it tends to yield better filterbanks. It does not reduce the transform and inverse transform cost however.

(a) (b) (c) (d)

Fig. 12. The 512 $\times$ 512 Barbara image coded at 1 : 32 using various transforms: (a) 8 $\times$ 8 DCT, (b) 8 $\times$ 16 LOT, (c) 8 $\times$ 16 GLBT, and (d) 16 $\times$ 32 GLBT.

All FB's presented in this paper are obtained from the multivariable nonlinear optimization routine *simplex* in Matlab. To initialize the lattice, we set the matrices containing the free parameters ($\mathbf{U}_i$ and $\mathbf{V}_i$) to either $\mathbf{I}$ or $-\mathbf{I}$. More specifically, the rotation angles θ_i are initialized to either zero or π, whereas the diagonal multipliers α_i are all initialized to 1. A set of weighting factors that we have found to provide a reasonable tradeoff between various transform properties is $\{k_i\} = \{10.0, 1.0, 0.1, 0.5, 0.5\}$.

The frequency and impulse responses of several orthogonal are shown in Fig. 9(a)–(d). The 8 $\times$ 8 DCT [6] and the 8 $\times$ 16 quasioptimal type-II fast LOT [7] in Fig. 9(a) and (b), respectively, are from previous works. They are included to serve as comparative yardsticks of how important a good choice of FB is in image coding applications. The 4 $\times$ 24 4 $\times$ 8 VLLOT and the 8 $\times$ 40 GenLOT are designed following the guidelines presented in this section. The biorthogonal LPFB's are depicted in Fig. 10(a) and (b). The objective performance measures of the DCT, the LOT, and the new block transforms are tabulated and compared in Table II. Notice that the GenLOT and GLBT all have high coding gain, nonequiripple frequency responses that decay to zero at dc, high mirror frequency attenuation, and their impulse responses decay smoothly to zero at two ends—another crucial factor in reducing blocking artifacts [7], [8].

C. Treatment of DC Band

Fig. 8 shows that there still exists correlation between dc coefficients. To decorrelate the dc band even more, several levels of wavelet decomposition can be used depending on the

(a)

(b)

(c)

(d)

Fig. 13. The 512 $\times$ 512 Goldhill image coded by the 16 $\times$ 32 GLBT: (a) 1 : 16, 33.42 dB, (b) 1 : 32, 30.84 dB, (c) 1 : 64, 28.74 dB, and (d) 1 : 100, 27.62 dB.

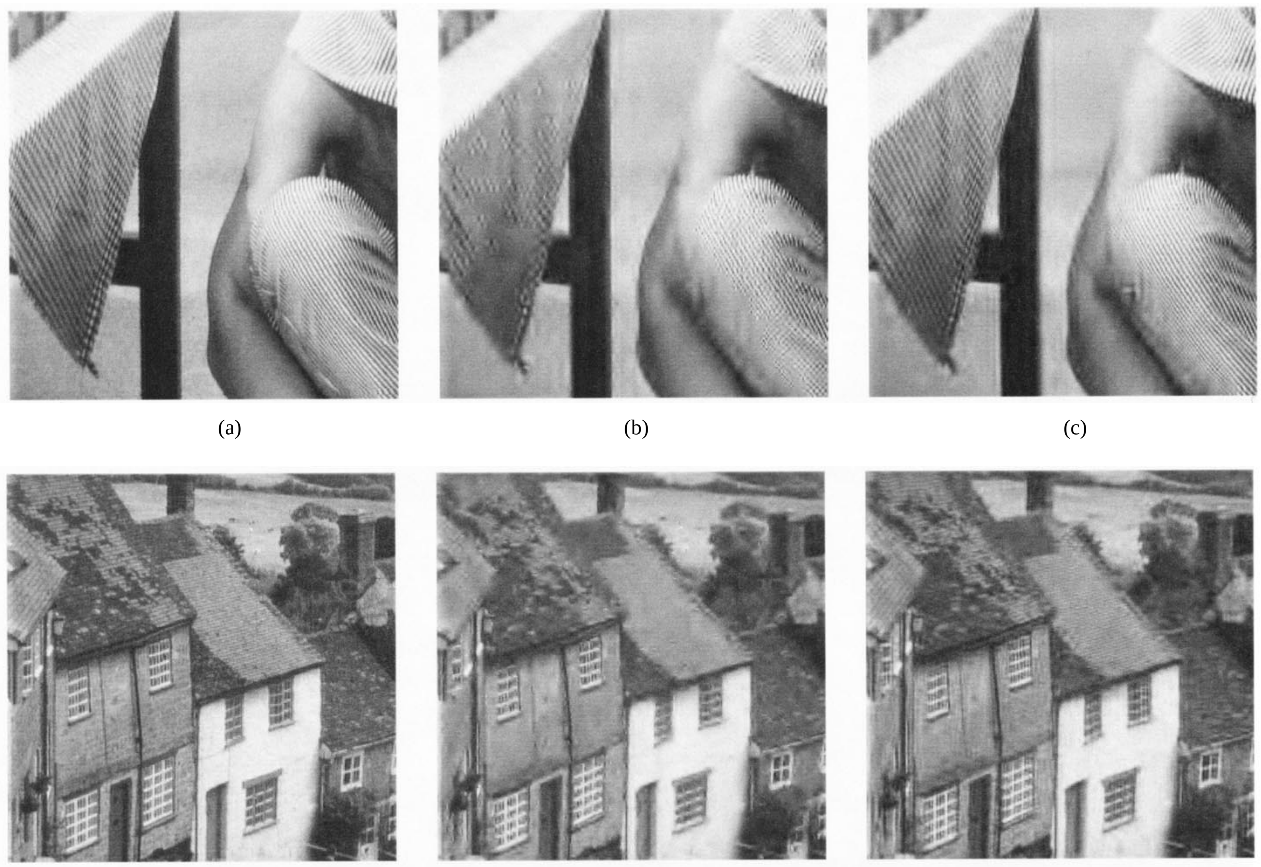
input image size. Besides the obvious increase in the coding efficiency of dc coefficients thanks to a deeper coefficient trees, wavelets provide variably longer bases for the signal's dc component, leading to smoother reconstructed images, i.e., blocking artifacts are further reduced. Regularity objective can be added in the transform design process to produce M-band wavelets, and a wavelet-like iteration can be carried out as well. In all results presented later in this paper, we choose the popular biorthogonal 9/7-tap pair [29] to process the dc coefficients.

IV. Coding Results

The objective coding results (PSNR in dB) for standard 512 $\times$ 512, 8-b gray-scale test images Lena, Barbara, and Goldhill are tabulated in Table III. The transforms in comparison are as follows.

- 9/7-tap biorthogonal wavelet [29].
- 8 $\times$ 8 DCT [6] shown in Fig. 9(a).
- 8 $\times$ 16 LOT [7] shown in Fig. 9(b).
- 4 $\times$ 24 4 $\times$ 8 VLLOT shown in Fig. 9(c).
- 8 $\times$ 40 GenLOT shown in Fig. 9(d).
- 8 $\times$ 16 GLBT shown in Fig. 10(a).
- 16 $\times$ 32 GLBT shown in Fig. 10(b).

Except the 9/7-tap biorthogonal wavelet, all of the transforms listed above are multiband uniform LPPRFB's, and their transform coefficients are encoded as described in Section III. All computed PSNR quotes in dB are obtained from a real compressed bit stream with all overheads included. The rate-



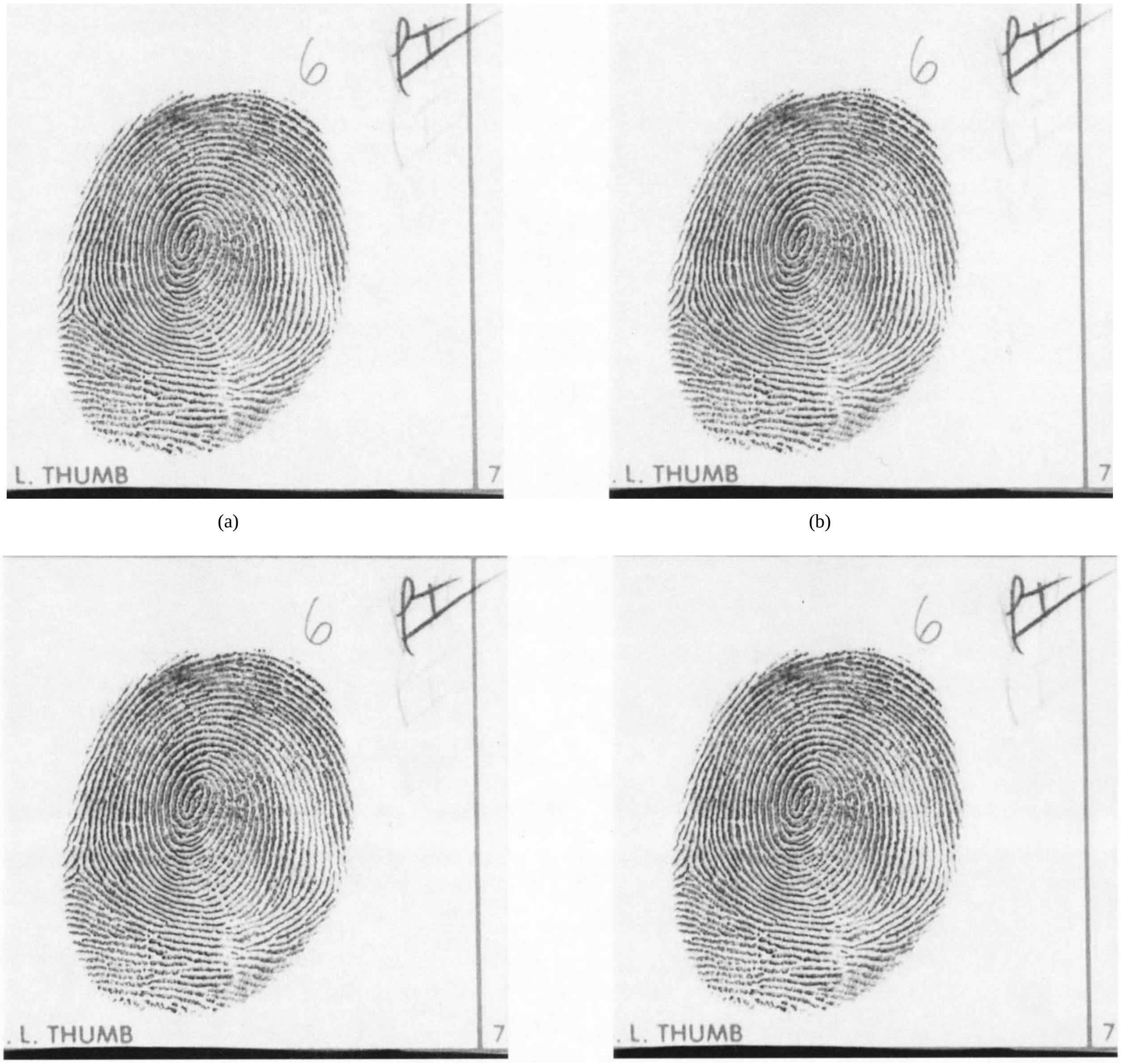
(a) (b) (c) (d) (e) (f)

Fig. 14. Perceptual comparison between the wavelet and the block transform embedded coder. Enlarged portions: (a) original Barbara image, (b) SPIHT at 1 : 32, (c) 8 × 16 GLBT embedded coder at 1 : 32, (d) original Goldhill, (e) SPIHT at 1 : 32, and (f) 8 × 16 GLBT embedded coder at 1 : 32.

distortion curves in Fig. 11 and the tabulated coding results in Table III clearly demonstrate the superiority of our block transform coder. For a smooth image like Lena where the wavelet transform can sufficiently decorrelate, SPIHT offers a comparable performance. However, for a highly-textured image like Barbara, the 8 × 40 GenLOT, the 8 × 16 GLBT, and the 16 × 32 GLBT coder can provide a PSNR gain of more than 2 dB over a wide range of bit rates. Unlike other block transform coders whose performance dramatically drops at very high compression ratios, the new progressive coders are consistent throughout as illustrated in Fig. 11. Comparing to the M-channel FB's in previous works (the DCT and the LOT), the new FB's consistently provide higher coding performances. The PSNR improvement can reach up to almost 3 dB comparing to the DCT and more than 1 dB comparing to the LOT.

Figs. 12–14 confirm the superiority of the new coders in reconstructed image quality as well. Fig. 12 shows reconstructed Barbara images at 1 : 32 using various block transforms. Comparing to JPEG, blocking artifacts are already remarkably reduced in the DCT-based coder in Fig. 12(a) and the LOT-based coder in Fig. 12(b). Blocking is completely eliminated when the DCT and the LOT are replaced by the new FB's as shown in Figs. 12(c)–(d) and 13. Even at 1 : 100, the reconstructed Goldhill image in Fig. 13(d) is still visually pleasant: no blocking and not much ringing. A closer look in Fig. 14(a)–(c) (where enlarged 256 × 256 image portions are shown so that artifacts can be more easily seen) reveals that besides blocking elimination, the 8 × 16 GLBT can preserve texture nicely (the table cloth and the clothes pattern in the Barbara image) while keeping the edges relatively clean. Comparing to the 9/7-tap wavelet, our M-channel FB's yield overall sharper reconstructed images with more defining edges and more evenly reconstructed texture regions. Although the PSNR difference is not as striking in the Goldhill image, the improvement in perceptual quality is rather significant as shown in Fig. 14(d)–(f).

As previously mentioned, the improvement over wavelets keys on the lapped transform's ability to capture and separate localized signal components in the frequency domain. In the spatial domain, this corresponds to images with directional repetitive texture patterns. To illustrate this point, the lapped-transform-based coder is compared against the FBI wavelet scalar quantization (WSQ) standard [30]. When the original 768 × 768 gray-scale fingerprint image shown in Fig. 15(a) is compressed at 1 : 13.6 (43 366 bytes) by the WSQ coder, Bradley *et al.* reported a PSNR of 36.05 dB. Using the 16 × 32 GLBT in Fig. 10(b), a PSNR of 38.09 dB can be achieved at the same compression ratio. At the same level of PSNR, the GLBT coder can compress the image down to 1 : 20 where the reconstructed image is shown in Fig. 15(b). To put this in perspective, the wavelet-packet-based SFQ



(a) (b) (c) (d)

Fig. 15. Fingerprint compression example: (a) original fingerprint image (589 824 bytes), (b) coded by the 16 × 32 GLBT coder at 1 : 20 (29 490 bytes), 36.05 dB, (c) coded by the WSQ coder at 1 : 18.036 (32 702 bytes), 34.42 dB, and (d) coded by the 16 × 32 GLBT coder at 1 : 27 (21 845 bytes), 34.42 dB.

coder in [16] reported a PSNR of only 37.30 dB at 1 : 13.6 compression ratio. At 1 : 18.036 (32 702 bytes), the WSQ's reconstructed image shown in Fig. 15(c) has a PSNR of 34.42 dB while the GLBT coder produces 36.57 dB. At the same distortion level, the GLBT coder can compress the image down to a compression ratio of 1 : 27 (21 845 bytes) as shown in Fig. 15(d). Again, the reader should note the high visual quality of the reconstructed images in Fig. 15(b) and (d): no disturbing blocking and ringing artifacts.

V. Conclusions

We have presented in this paper a novel low-complexity progressive transmission image coding scheme where M-channel uniform LPPRFB, the wavelet transform, and the zerotree entropy coder are combined to yield excellent performance in coding efficiency. As a recapitulation, the resulting coder offers the following advantages.

- It is based on multichannel block-transforms, which can provide finer frequency spectrum partitioning and higher energy compaction.
- The transform facilitates hardware implementation with efficient lattice structures which employ a minimal number of delay elements and are robust under the quantization of lattice coefficients.
- The block-based nature of the transforms facilitate region-of-interest coding/decoding.
- The transform also increases the parallelism of computation.
- The coder has progressive image transmission and all of its attractive characteristics: embedded quantization, exact bit rate control, and idempotency.

- It provides high subjective and objective performance, outperforms consistently the best progressive coders published recently in literature by a wide margin.

REFERENCES

[1] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing,* vol. 41, pp. 3445–3462, Dec. 1993.

[2] A. Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.,* vol. 6, pp. 243–250, June 1996.

[3] "Compression with reversible embedded wavelets," RICOH Co. Ltd., submission to ISO/IEC JTC1/SC29/WG1 for the JTC1.29.12 work item, 1995. Available at: http://www.crc.ricoh.com/CREW.

[4] Z. Xiong, K. Ramchandran, and M. T. Orchard, "Space-frequency quantization for wavelet image coding," *IEEE Trans. Image Processing,* vol. 6, pp. 677–693, May 1997.

[5] W. B. Pennebaker and J. L. Mitchell, *JPEG: Still Image Compression Standard.* New York: Van Nostrand Reinhold, 1993.

[6] K. R. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, Applications.* New York: Academic, 1990.

[7] H. S. Malvar, *Signal Processing with Lapped Transforms.* Boston, MA: Artech House, 1992.

[8] ——, "Biorthogonal and nonuniform lapped transforms for transform coding with reduced blocking and ringing artifacts," *IEEE Trans. Signal Processing, Spec. Issue Multirate Syst., Filter Banks, Wavelets, Applicat.,* vol. 46, pp. 1043–1053, Apr. 1998.

[9] A. Soman, P. P. Vaidyanathan, and T. Q. Nguyen, "Linear phase paraunitary filter banks," *IEEE Trans. Signal Processing,* vol. 41, pp. 3480–3496, Dec. 1993.

[10] R. L. de Queiroz, T. Q. Nguyen, and K. R. Rao, "The GenLOT: Generalized linear-phase lapped orthogonal transform," *IEEE Trans. Signal Processing,* vol. 40, pp. 497–507, Mar. 1996.

[11] T. D. Tran and T. Q. Nguyen, "On M-channel linear-phase FIR filter banks and application in image compression," *IEEE Trans. Signal Processing,* vol. 45, pp. 2175–2187, Sept. 1997.

[12] S. C. Chan, "The generalized lapped transform (GLT) for subband coding applications," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing,* Detroit, MI, May 1995.

[13] T. D. Tran, R. de Queiroz, and T. Q. Nguyen, "The generalized lapped biorthogonal transform," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing,* Seattle, WA, May 1997.

[14] ——, "Linear phase perfect reconstruction filter bank: Lattice structure, design, and application in image coding," *IEEE Trans. Signal Processing,* to be published.

[15] S. Trautmann and T. Q. Nguyen, "GenLOT—Design and application for transform-based image coding," in *Proc. Asilomar Conf.,* 1995.

[16] Z. Xiong, K. Ramchandran, and M. T. Orchard, "Wavelet packets image coding using space-frequency quantization," *IEEE Trans. Image Processing,* vol. 7, pp. 892–898, June 1998.

[17] P. P. Vaidyanathan, *Multirate Systems and Filter Banks.* Englewood Cliffs, NJ: Prentice-Hall, 1993.

[18] G. Strang and T. Q. Nguyen, *Wavelets and Filter Banks.* Cambridge, MA: Wellesley-Cambridge, 1996.

[19] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[20] P. N. Topiwala, Ed., *Wavelet Image and Video Compression.* Boston, MA: Kluwer, June 1998.

[21] R. A. DeVore, B. Jawerth, and B. J. Lucier, "Image compression through wavelet transform coding," *IEEE Trans. Inform. Theory,* vol. 38, pp. 719–746, Mar. 1992.

[22] M. Rabbani and P. W. Jones, *Digital Image Compression Techniques.* Bellingham, WA: SPIE, 1991.

[23] Z. Xiong, O. Guleryuz, and M. T. Orchard, "A DCT-based embedded image coder," *IEEE Signal Processing Lett.,* vol. 3, pp. 289–290, Nov. 1996.

[24] T. A. Ramstad, S. O. Aase, and J. H. Husoy, *Subband Compression of Images: Principles and Examples.* New York: Elsevier, 1995.

[25] J. Katto and Y. Yasuda, "Performance evaluation of subband coding and optimization of its filter coefficients," *SPIE Proc. Vis. Commun. Image Process.,* 1991.

[26] M. Ikehara, T. D. Tran, and T. Q. Nguyen, "Linear phase paraunitary filter banks with unequal-length filters," in *Proc. IEEE Int. Conf. on Image Processing,* Santa Barbara, CA, Oct. 1997.

[27] T. D. Tran, M. Ikehara, and T. Q. Nguyen, "Linear phase paraunitary filter bank with variable-length filters and its application in image compression," *IEEE Trans. Signal Processing,* vol. 47, pp. 2730–2744, Oct. 1999.

[28] T. D. Tran, R. de Queiroz, and T. Q. Nguyen, "Variable-length generalized lapped biorthogonal transform," in *Proc. IEEE Int. Conf. Image Processing,* Chicago, IL, Oct. 1998.

[29] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using the wavelet transform," *IEEE Trans. Image Processing,* vol. 1, pp. 205–220, Jan. 1992.

[30] J. N. Bradley, C. M. Brislawn, and T. Hopper, "The FBI wavelet/scalar quantization standard for gray-scale fingerprint image compression," in *Proc. VCIP,* Orlando, FL, Apr. 1993.

Trac D. Tran received the B.S. and M.S. degrees from the Massachusetts Institute of Technology, Cambridge, in 1994, and the Ph.D. degree from the University of Wisconsin, Madison, WI, in 1998, all in electrical engineering.

He joined the faculty of The Johns Hopkins University, Baltimore, MD, in July 1998, as an Assistant Professor in the Department of Electrical and Computer Engineering. His research interests are in the field of digital signal processing, particularly in multirate systems, filterbanks, wavelets, and their applications in signal representation, compression, and processing.

Truong Q. Nguyen received the B.S., M.S., and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, in 1985, 1986, and 1989, respectively.

He was with Massachusetts Institute of Technology (MIT) Lincoln Laboratory from June 1989 to July 1994, as Member of Technical Staff. During the academic year 1993–1994, he was a visiting lecturer at MIT and an Adjunct Professor at Northeastern University, Boston, MA. He was with the Electrical and Computer Engineering Department, University of Wisconsin, Madison, from August 1994 to April 1998. Since July 1996, he has been with the Electrical and Computer Engineering Department Department, Boston University. His research interests are in digital and image signal processing, image and video compression, multirate systems, wavelets and applications, biomedical signal processing, filter design, and A/D converter. He is the coauthor (with Prof. G. Strang) of the textbook *Wavelets and Filter Banks* (Wellesley, MA: Cambridge Wellesley Press).

Prof. Nguyen was an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II. He was a recipient of a fellowship from Aerojet Dynamics for advanced studies. He received the IEEE TRANSACTIONS ON SIGNAL PROCESSING Paper Award (image and multidimensional processing area) for the paper he co-authored (with P. P. Vaidyanathan) on linear-phase perfect-reconstruction filterbanks (1992). He received the NSF Career Award in 1995. He served on the DSP Technical Committee for the CAS Society. He is a member of Tau Beta Pi, Eta Kappa Nu, and Sigma Xi.