THE VARIABLE-LENGTH GENERALIZED LAPPED BIORTHOGONAL TRANSFORM

Trac D. Tran[1] *Ricardo L. de Queiroz*[2] *Truong Q. Nguyen*[3]

[1]University of Wisconsin, ECE Department, Madison, WI, 53706
[2]Xerox Corporation, Corporate Research and Technology, Webster, NY, 14580
[3]Boston University, ECE Department, Boston, MA 02215
trac@saigon.ece.wisc.edu, queiroz@wrc.xerox.com, nguyent@bu.edu

ABSTRACT

This paper introduces a class of linear phase lapped biorthogonal transforms with basis functions of variable length (VLGLBT). The transform can be characterized by a lattice which structurally enforces both linear phase and perfect reconstruction properties as well as provides a fast and efficient transform implementation. Our main motivation of the new transform is its application in image coding. The VLGLBT has several long overlapped basis functions for representing smooth signals to avoid blocking artifacts. The rest of the bases covering high-frequency bands are constrained to be short to limit ringing artifacts. The relaxation of the orthogonal constraint allows the VLGLBT to have significantly different analysis and synthesis banks which can be tailored appropriately to obtain high-quality reconstructed images. Most importantly, the variable-length property allows us to design very fast and low-complexity transforms. Comparing to the popular DCT, a fast VLGLBT named FLT only requires 6 more multiplications and 8 more additions. Yet, image coding examples show that the FLT is far superior than the DCT and is close to the 9/7-tap biorthogonal wavelet in both objective and subjective coding performance.

1. INTRODUCTION

Multi-channel block transforms have long found application in image coding. For instance, the JPEG image compression standard [1] employs the 8 × 8 discrete cosine transform (DCT) at its transformation stage. At high bit rates, JPEG offers almost visually lossless reconstruction image quality. However, when more compression is needed, annoying blocking artifacts appear since the DCT bases are short and do not overlap, thus have discontinuities at block ends. The development of the lapped orthogonal transform [2], its generalized version GenLOT [3], and the extensions to biorthogonality [4],[5], [6] help solve the blocking problem by borrowing pixels from the adjacent blocks to produce the transform coefficients of the current block. Lapped transforms outperform the DCT on two counts: (i) from the analysis viewpoint, it takes into account inter-block correlation, hence, provides better energy compaction; (ii) from the synthesis viewpoint, its basis functions decay asymptotically to zero at the ends, reducing blocking discontinuities drastically.

Nevertheless, lapped transforms have not yet been able to replace the DCT in international standards. One reason is the increase in computational complexity. In this paper, we introduce the first generalized lapped biorthogonal transform with basis functions of variable length called the VLGLBT. Since blocking is most noticeable in smooth image regions, in order to reduce blocking artifacts, filters covering high-frequency bands do not have to be long and overlapped. Long basis functions are only needed for low-frequency components. Moreover, the short basis functions reserving for high-frequency signal components can effectively limit the ringing artifacts. The transform's variable-length property allows us to design fast lapped transforms. Comparing to the DCT, the novel FLT only requires 6 more multiplications, 8 more additions, and 2 more delays (borrowing 4 low-frequency coefficients from 2 neighboring blocks). Despite its simplicity, the FLT provides a significant improvement in image quality over the traditional DCT – little blocking and ringing artifacts at medium and high compression ratios.

2. LATTICE STRUCTURE FOR THE VLGLBT

2.1. Review

In this paper's context, lapped transforms are M-channel uniform linear phase perfect reconstruction filter banks (LPPRFB) whose polyphase representation is depicted in Figure 1 below.

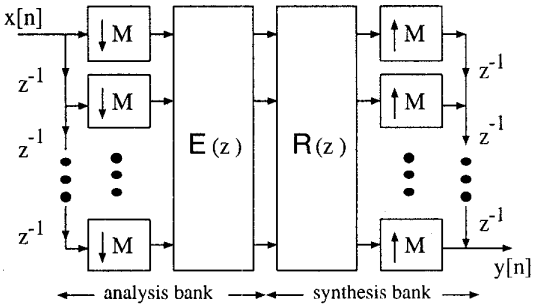

Figure 1. Polyphase representation of an LPPRFB.

The most general lattice for M-channel lapped biorthogonal transforms (GLBT) is presented in [6]. The polyphase matrix $\mathbf{E}(z)$ can be factorized as

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

$$\mathbf{G}_i(z) = \frac{1}{2}\begin{bmatrix} \mathbf{U}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_i \end{bmatrix}\begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{I} & -\mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & z^{-1}\mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{I} & -\mathbf{I} \end{bmatrix}$$

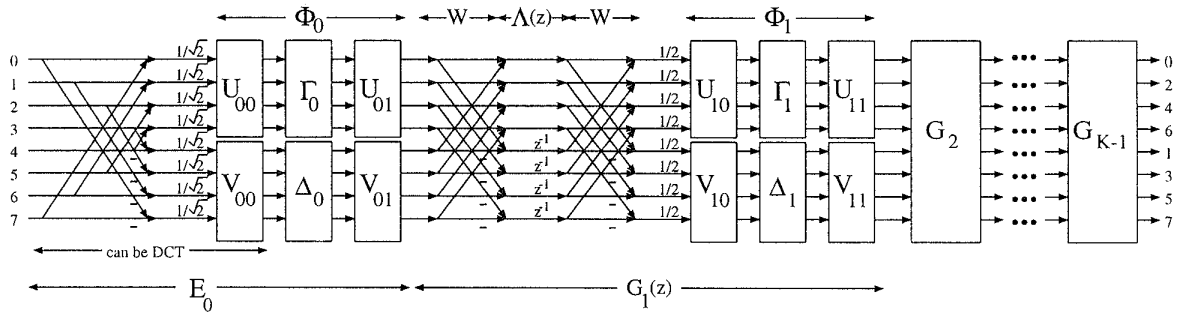

Figure 2. The most general lattice structure for linear phase lapped transforms with filter length $L = KM$.

$$\triangleq \tfrac{1}{2}\,\mathbf{\Phi}_i\,\mathbf{W}\,\mathbf{\Lambda}(z)\,\mathbf{W}, \quad \text{and} \tag{2}$$

$$\mathbf{E}_0 \;=\; \tfrac{1}{\sqrt{2}}\begin{bmatrix} \mathbf{U}_0 & \mathbf{U}_0\,\mathbf{J}_{\frac{M}{2}} \\ \mathbf{V}_0\,\mathbf{J}_{\frac{M}{2}} & -\mathbf{V}_0 \end{bmatrix}. \tag{3}$$

This lattice results in all filters having length $L = KM$. Each cascading structure $\mathbf{G}_i(z)$ increases the filter length by M. All $\mathbf{U}_i$ and $\mathbf{V}_i$, $i = 0, 1, \ldots, K-1$, are arbitrary $\frac{M}{2} \times \frac{M}{2}$ invertible matrices, and they can be completely parameterized by their SVD decomposition, i.e., $\mathbf{U}_i = \mathbf{U}_{i0}\mathbf{\Gamma}_i\mathbf{U}_{i1}$ and $\mathbf{V}_i = \mathbf{V}_{i0}\mathbf{\Delta}_i\mathbf{V}_{i1}$, where $\mathbf{U}_{i0}$, $\mathbf{U}_{i1}$, $\mathbf{V}_{i0}$, $\mathbf{V}_{i1}$ are orthogonal matrices, and $\mathbf{\Gamma}_i$, $\mathbf{\Delta}_i$ are diagonal matrices with positive elements. The parameterization is shown in Figure 3. For fast implementations, the initial stage $\mathbf{E}_0$ can be replaced by the DCT. Further trade-off between the FB's speed and performance can be elegantly carried out by setting some of the diagonal multipliers to 1 or some of the rotation angles to 0.

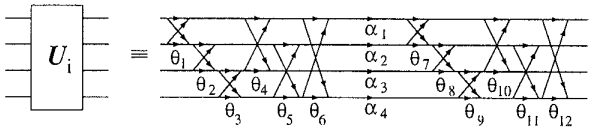

Figure 3. Parameterization of an invertible matrix.

2.2. Existence Conditions

Let us first consider an M-channel LPPRFB with variable-length filters: M is even, N filters of length MK, $(M-N)$ filters of length $M(K-1)$, each analysis filter $h_i[n]$ and the corresponding synthesis filter $f_i[n]$ have the same length $L_i, 0 \le i \le M-1$. The following theorems describe the class of all possible solutions in terms of the FB's symmetry polarity and filter length.

Theorem I. For the class of LPPRFBs described above, the number of long filters N and the number of short filters $(M-N)$ must both be even.

Proof. This result can be trivially established from the permissible solutions of the M-channel LPPRFB whose filters have lengths $L_i = K_iM + \beta$ [7]. The sum of all the filter lengths has to be even when M is even. Hence, N cannot be odd. □

Theorem II. Furthermore, half of the long filters are symmetric, and half of the short filters are symmetric.

The proof is omitted here due to the lack of space. It is constructed analogously to that of orthogonal systems presented in [8]. The complete proof can be found in [9] and will appear in the full version of this paper [10].

2.3. Variable-length Lattice

Let $\mathbf{E}_L(z)$ be the $N \times M$ polyphase matrix of order $(K-1)$, representing the long analysis filters, and $\mathbf{E}_S(z)$ be the $(M-N) \times M$ polyphase matrix of order $(K-2)$, representing the shorter analysis filters. Similarly, let $\mathbf{R}_L(z)$ and $\mathbf{R}_S(z)$ represent the long and the short synthesis filters respectively. Without any loss of generality, the long filters are permuted to be on top. The following factorization establishes the completeness of our solution.

Since all filters have linear phase, $\mathbf{E}(z)$ also has to satisfy the LP property [7]:

$$\begin{cases} \mathbf{E}_L(z) &= z^{-(K-1)}\,\mathbf{D}_L\,\mathbf{E}_L(z^{-1})\,\mathbf{J}_M \\ \mathbf{E}_S(z) &= z^{-(K-2)}\,\mathbf{D}_S\,\mathbf{E}_S(z^{-1})\,\mathbf{J}_M. \end{cases} \tag{4}$$

where $N \times N$ $\mathbf{D}_L$ and $(M-N) \times (M-N)$ $\mathbf{D}_S$ are diagonal matrices whose entries are $+1$ when the corresponding filter is symmetric and -1 when the corresponding filter is antisymmetric. $\mathbf{E}_L(z)$ now forms a remarkably similar system to an N-channel order-$(K-1)$ GLBT [6].

From [6],[11], there always exists a factorization similar to the one shown in Eq.(1) that reduces the order of the polyphase matrix $\mathbf{E}_L(z)$ by one. Hence, the VLGLBT's polyphase matrix $\mathbf{E}(z)$ can always be factorized as follows (the $\frac{N}{2}$ long symmetric filters are arranged on top)

$$\mathbf{E}(z) \;=\; \hat{\mathbf{G}}_0(z)\,\mathbf{E}_{K-2}(z), \tag{5}$$

where

$$\hat{\mathbf{G}}_0(z) \;\triangleq\; \tfrac{1}{2}\,\hat{\mathbf{\Phi}}_0\,\hat{\mathbf{W}}\,\hat{\mathbf{\Lambda}}(z)\,\hat{\mathbf{W}}, \quad \text{and} \tag{6}$$

$$\hat{\mathbf{W}} = \begin{bmatrix} \mathbf{I}_{\frac{N}{2}} & \mathbf{I}_{\frac{N}{2}} & \mathbf{0} \\ \mathbf{I}_{\frac{N}{2}} & -\mathbf{I}_{\frac{N}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}, \; \hat{\mathbf{\Lambda}}(z) = \begin{bmatrix} \mathbf{I}_{\frac{N}{2}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & z^{-1}\mathbf{I}_{\frac{N}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix},$$

$$\hat{\mathbf{\Phi}}_0 \;=\; \begin{bmatrix} \hat{\mathbf{U}}_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{V}}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}.$$

The above factorization leaves $\mathbf{E}_S(z)$ untouched, it reduces the length of the longer filters by M, so all filters now have the same length of $M(K-1)$. $\mathbf{E}_{K-2}(z)$ is the familiar polyphase matrix of an order-$(K-2)$ GLBT, and it can be factorized into the familiar cascade structure in Eq.(1). The complete factorization is

$$\mathbf{E}(z) \;=\; \hat{\mathbf{G}}_0(z)\,\mathbf{G}_{K-2}(z)\,\cdots\,\mathbf{G}_1(z)\,\mathbf{E}_0, \tag{7}$$

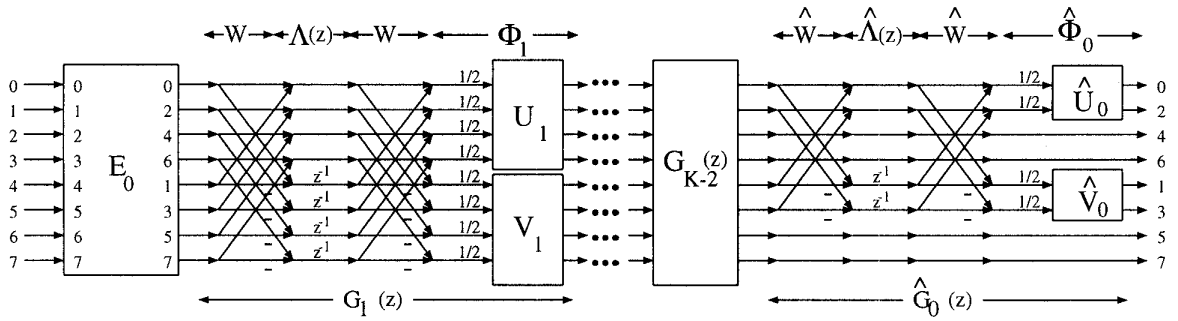

Figure 4. Detailed lattice structure for a VLGLBT (drawn for $M = 8$ and $N = 4$).

depicted in Figure 4. The inverse transform is obtained by inverting each building block and causality can be achieved by a z^{-1} shift, i.e.,

$$\mathbf{R}(z) = \mathbf{E}_0^{-1} z^{-1} \mathbf{G}_1^{-1}(z) \cdots z^{-1} \mathbf{G}_{K-2}^{-1}(z) z^{-1} \hat{\mathbf{G}}_0(z).$$

We should also mention that the factorization in Eq.(7) can be proven to be minimal, i.e., the resulting lattice employs the least number of delays in the implementation [9].

Of course, more VL structure $\hat{\mathbf{G}}_i(z)$ can be added to increase the frequency resolution of the long filters. Each $\hat{\mathbf{G}}_i(z)$ block increases the length of N_i filters by M and leaves the rest intact. The most general lattice for the VLGLBT is shown in Figure 5. The invertible matrices $\hat{\mathbf{U}}_i$, $\hat{\mathbf{V}}_i$, $\mathbf{U}_i$, and $\mathbf{V}_i$ are parameterized by the SVD as illustrated in Figure 3. The rotation angles $\hat{\theta}_i$, θ_i and the diagonal multipliers $\hat{\alpha}_i$, α_i are the free parameters (also called lattice coefficients) that can be tuned to optimize the VLGLBT. Invertibility is guaranteed structurally under a mild condition – as long as none of the diagonal coefficients $\hat{\alpha}_i$ and α_i is quantized to zero. If orthogonality is desired, all diagonal coefficients can be set to unity and we are back to the VLLOT [8].

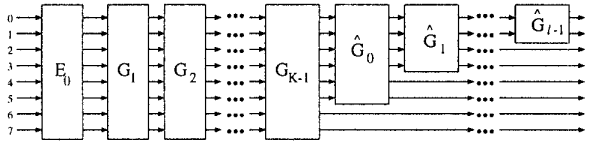

Figure 5. General lattice structure for the VLGLBT.

2.4. The Fast VLGLBT

Despite significant performance improvements, lapped transforms have not yet been able to replace the DCT in practical systems. One obvious reason is the increase in computational complexity. The rest of this section is devoted to the design of a high-performance, yet low-complexity, lapped transform named FLT to replace the DCT in the near future where the newly-found flexibility of the VLGLBT is exploited to our advantage.

To minimize the transform's complexity, we choose the fewest possible number of long filters and set the initial stage $\mathbf{E}_0$ to be the DCT itself. Only two variable-length structures $\hat{\mathbf{G}}_0(z)$ and $\hat{\mathbf{G}}_1(z)$ are employed ($N_0 = N_1 = 2$). In the orthogonal case, we can only obtain the trivial solution since the matrices $\hat{\mathbf{U}}_i$ and $\hat{\mathbf{V}}_i$ degenerate to singleton 1

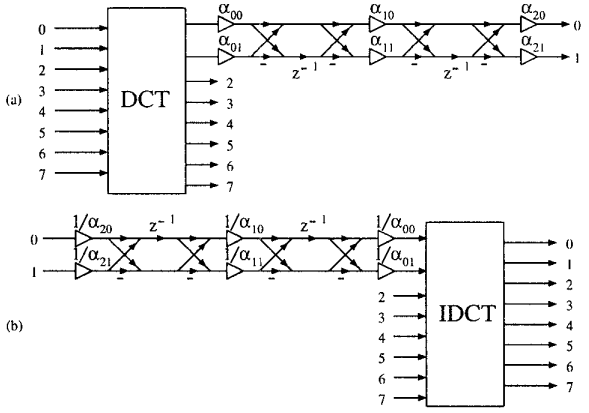

Figure 6. The FLT lattice drawn for $M = 8$. (a) Forward transform. (b) Inverse transform.

or -1, and there are no free parameters for transform optimization. In the more general biorthogonal case, nontrivial solutions exist. The invertible matrices $\hat{\mathbf{U}}_i$ and $\hat{\mathbf{V}}_i$ now becomes the lattice coefficients α_{i0} and α_{i1} as shown in the final FLT lattice in Figure 6.

3. DESIGN EXAMPLE AND APPLICATION IN IMAGE CODING

An FLT design example is obtained using unconstrained nonlinear optimization where the lattice coefficients $\{\alpha_{00}, \alpha_{01}, \alpha_{10}, \alpha_{11}, \alpha_{20}, \alpha_{21}\}$ are the free parameters. The cost function is a weighted combination of coding gain, DC leakage, stopband attenuation, and mirror frequency attenuation, all of which are desirable properties in yielding high-quality reconstructed images [14].

As usual, the system's generalized coding gain is given the highest priority. The relaxation of the orthogonal constraint allows us to tailor the FLT's analysis and synthesis filters appropriately. In particular, in the analysis bank, we opt for high stopband attenuation near DC ($\omega = 0$) since that is where most of the energy is concentrated in meaningful images. In the synthesis bank, the stopband attenuation cost function is reversed. We force the two long synthesis filters covering low-frequency bands to have high stopband attenuation near and/or at $\omega = \pi$ to enhance their smoothness.

The frequency and impulse responses of the FLT's analy-

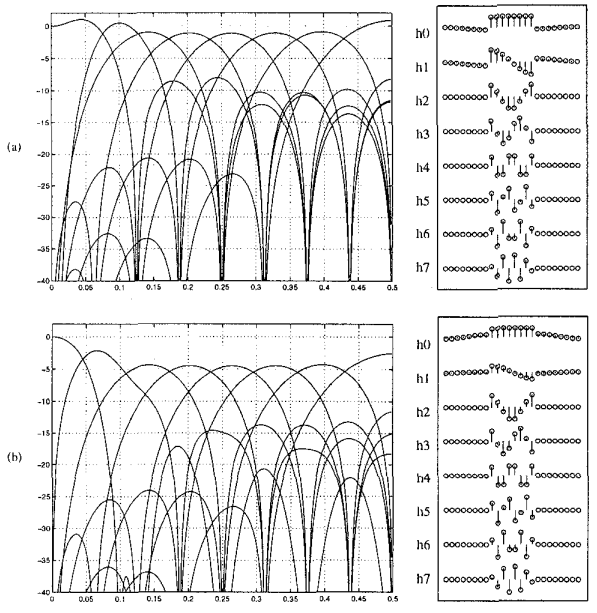

Figure 7. An FLT design example. (a) Analysis bank. (b) Synthesis bank.

sis and the synthesis filters are depicted in Figure 7(a) and Figure 7(b) respectively. Transform properties can be found in Table 1. Note that the last six filters of the FLT come straight from the DCT's.

The VLGLBT's objective coding results (PSNR in dB) as well as other popular transform's for standard 512×512 Lena, Goldhill, and Barbara test images are tabulated in Table 2. In the block-transform cases, we used the modified zerotree structure in [13]. The transforms in comparison are:

- DCT: 8-channel, all 8-tap filters
- FLT: VLGLBT with 8-channel, two 24-tap and six 8-tap filters
- LOT: 8-channel, all 16-tap filters
- GLBT: 8-channel, all 16-tap filters
- 9/7-tap wavelet, six level of decompositions.

As expected, the FLT offers a $0.3 - 0.5$ dB improvement over the DCT at medium and low bit rates. It is inferior to much more complex transforms like the 9/7-tap wavelet and the 8×16 GLBT. However, we stress that it is designed mainly to improve the reconstructed image quality. Figure 8 confirms its high potential in this important criterion. Enlarged portions of reconstructed images obtained from various transforms show that our novel FLT provides a significant improvement in image quality over the traditional DCT: blocking is avoided while ringing is suppressed. In fact, the FLT is much better than the quasi-optimal type-II LOT [2] in blocking elimination. The FLT's visual performance comes quite close to those of state-of-the-art wavelets and GLBTs. For more details on the coding algorithm and comparison, the reader is referred to the web site *http://saigon.ece.wisc.edu/~waveweb/Coder/index.html.*

4. CONCLUSIONS

We have presented in this paper the theory, design, and implementation of the variable-length lapped biorthogonal transform. The FLT is based on a fast, efficient, robust, and modular lattice structure. With only 6 more multiplications, 8 more additions, and 2 more delays comparing to the DCT, our transform offers a fast, low-cost, VLSI-friendly implementation while providing high-quality reconstructed images at medium and low bit rates as demonstrated in the image coding example. Its block-based nature also supports parallel processing mode, and facilitates region-of-interest coding/decoding as well as processing large images under limited memory constraint.

REFERENCES

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

[2] H. S. Malvar, *Signal Processing with Lapped Transforms*, Artech House, 1992.

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

[4] S. C. Chan, "The generalized lapped transform (GLT) for subband coding applications," *ICASSP,* Detroit, May 1995.

[5] H. S. Malvar, "Lapped biorthogonal transforms for transform coding with reduced blocking and ringing artifacts," *ICASSP,* Munich, April 1997.

[6] T. D. Tran, R. de Queiroz, and T. Q. Nguyen, "The generalized lapped biorthogonal transform," *ICASSP,* Seattle, May 1998.

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

[8] T. D. Tran, M. Ikehara, and T. Q. Nguyen, "Linear phase paraunitary filter bank with filters of different lengths and its application in image compression," submitted to *IEEE Trans. on Signal Processing,* Dec. 1997.

[9] T. D. Tran, *Linear phase perfect reconstruction filter banks: theory, structure, design, and application in image compression,* Ph.D. thesis, University of Wisconsin, Madison, WI, May 1998.

[10] T. D. Tran, R. L. de Queiroz, and T. Q. Nguyen, "The variable-length generalized lapped biorthogonal transform," in preparation, to be submitted to *IEEE Trans. on Image Processing.*

[11] T. D. Tran, R. L. de Queiroz, and T. Q. Nguyen, "Linear phase perfect reconstruction filter bank: lattice structure, design, and application in image coding," submitted to *IEEE Trans. on Signal Processing,* Apr. 1998.

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

[13] T. D. Tran and T. Q. Nguyen, "A lapped transform embedded image coder," *ISCAS,* Monterey, May 1998.

[14] G. Strang and T. Q. Nguyen, *Wavelets and Filter Banks,* Revised Edition, Wellesley-Cambridge Press, 1997.

Transform Characteristics	Transforms			
	8 x 8 DCT	2x24 6x8 VLGLBT	8 x 16 LOT	8 x 16 GLBT
Coding Gain (dB)	8.83	9.23	9.22	9.62
DC Attenuation (- dB)	310.62	314.38	312.56	327.40
Stopband Attenuation (- dB)	9.96	8.05	19.38	13.50
Mirror Freq. Attenuation (- dB)	322.10	309.76	317.24	55.54

Table 1. Transform characteristics comparison

	Lena					Goldhill					Barbara				
Comp. Ratio	SPIHT 9/7 WL	8 x 8 DCT	2x24 6x8 VLGLBT	8 x 16 LOT	8 x 16 GLBT	SPIHT 9/7 WL	8 x 8 DCT	2x24 6x8 VLGLBT	8 x 16 LOT	8 x 16 GLBT	SPIHT 9/7 WL	8 x 8 DCT	2x24 6x8 VLGLBT	8 x 16 LOT	8 x 16 GLBT
1:8	40.41	39.91	39.89	40.05	40.35	36.55	36.25	36.22	36.63	36.69	36.41	36.31	36.22	37.43	37.84
1:16	37.21	36.38	36.51	36.72	37.28	33.13	32.76	32.76	33.18	33.31	31.40	31.11	31.12	32.70	33.02
1:32	34.11	32.90	33.25	33.56	34.14	30.56	30.07	30.25	30.56	30.70	27.58	27.28	27.42	28.80	29.04
1:64	31.10	29.67	30.15	30.48	31.04	28.48	27.93	28.17	28.36	28.58	24.86	24.58	24.86	25.70	26.00

Table 2. Objective coding result comparison (PSNR in dB)

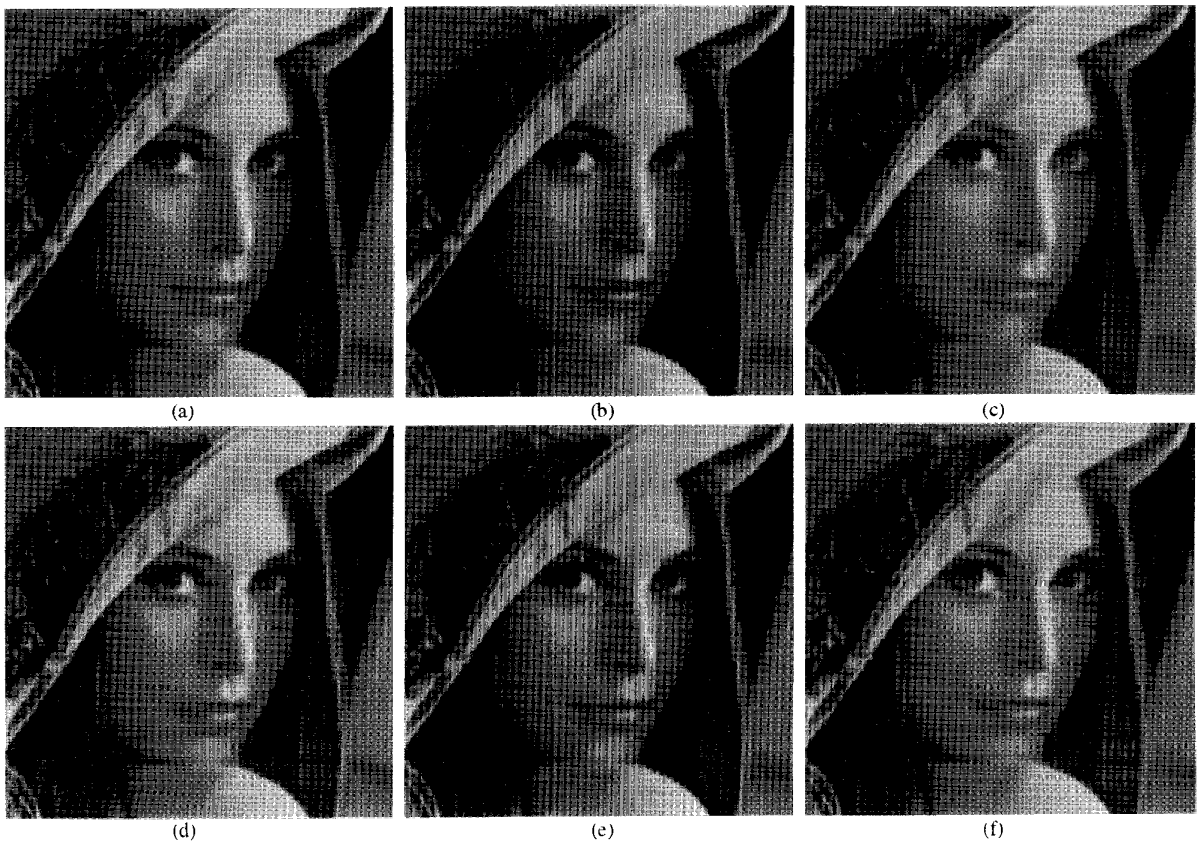

Figure 8. Coding results of Lena at 1:32 compression ratio. Enlarged portions. (a) Original image (b) DCT (c) 2×24 6×8 VLGLBT (d) 8×16 LOT (e) 8×16 GLBT (f) SPIHT, 9/7-tap biorthogonal wavelet.