THE GENERALIZED LAPPED BIORTHOGONAL TRANSFORM

Trac D. Tran[1] *Ricardo 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

A lattice structure based on the singular value decomposition (SVD) is introduced. The lattice can also be proven to use a minimal number of delay elements and to completely span a large class of M-channel linear phase perfect reconstruction filter bank (LPPRFB): all analysis and synthesis filters have the same FIR length of $L = KM$, sharing the same center of symmetry. The lattice also structurally enforces both linear phase and perfect reconstruction properties, is capable of providing fast and efficient implementation, and avoids the costly matrix inversion problem in the optimization process. From a block transform perspective, the new lattice represents a family of generalized lapped biorthogonal transform (GLBT) with arbitrary integer overlapping factor K. The relaxation of the orthogonal constraint allows the GLBT to have significantly different analysis and synthesis basis functions which can then be tailored appropriately to fit a particular application. Several design examples are presented along with a high-performance GLBT-based progressive image coder to demonstrate the superiority of the new lapped transforms.

1. INTRODUCTION

Linear phase perfect reconstruction filter banks have been used extensively in numerous applications, especially image processing [1]. In the two-channel case, all solutions have been found whereas there are still many open problems in M-channel cases. An attractive approach to the design and implementation of LPPRFB is the parameterization by lattice structures based on the factorization of the polyphase matrices $\mathbf{E}(z)$ and $\mathbf{R}(z)$ shown in Figure 1. The lattice structure offers fast and efficient implementation, retains both LP and PR properties regardless of coefficient quantization, and (if it is general enough) guarantees that no optimal solution will be excluded in the optimization process. Complete and minimal two-channel PR lattice structure has been reported in [2]. M-channel lattices have been presented for the more restricted paraunitary case [3], resulting in the generalized lapped orthogonal transform (GenLOT) [4]. No general lattice has been reported for the biorthogonal case (defined as $\mathbf{R}(z)\mathbf{E}(z) = z^{-K}\mathbf{I}$). Only several particular solutions were proposed so far: Chan replaced some orthogonal matrices in [4] by cascades of invertible block diagonal matrices [5]; Malvar suggested a simple scaling of the first antisymmetric basis function of the initial block (which was chosen to be the DCT) [6].

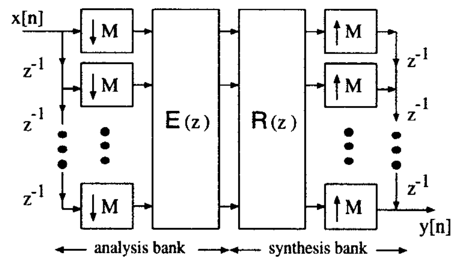


Figure 1. Polyphase representation of LPPRFB.

In this paper, we first introduce a general structure that propagates the linear phase property. Further imposition of the biorthogonal property on this structure results in a complete and minimal LPPRFB lattice whose invertible coefficient matrices are then parameterized by the SVD. In the lapped transform language, the structure is interpreted as a robust characterization of the GLBT. Several lapped transforms obtained from the nonlinear optimization of the lattice coefficients are presented. Finally, the superiority and the potential of the new LT family are illustrated through an image coding example. The GLBT-based embedded coder consistently outperforms the wavelet-based version SPIHT [7] by a large margin. The improvement in PSNR can be up to an astouding 2.65 dB.

Notation-wise, vectors and matrices are denoted by boldfaced characters. Special matrices have reserved symbols: $\mathbf{I}$, $\mathbf{J}$, $\mathbf{0}$, $\mathbf{D}$ represents, respectively, the identity matrix, the reversal matrix, the null matrix, and the diagonal matrix whose entry is $+1$ when the corresponding filter is symmetric and -1 when the corresponding filter is antisymmetric. Capital letters M, L, K denote respectively the number of channels, the filter length, and the overlapping factor.

2. LATTICE STRUCTURE

2.1. General LP-propagating Structure

Consider an M-channel FIR LPPRFB with all analysis and synthesis filters of length $L = KM$, having the same center of symmetry ($M \times L$ GLBT). The associated polyphase matrix $\mathbf{E}(z)$ has to satisfy the LP property [3]

$$\mathbf{E}(z) = z^{-(K-1)}\,\mathbf{D}\,\mathbf{E}(z^{-1})\,\mathbf{J}. \tag{1}$$

Define the order-$(K-1+N)$ polyphase matrix $\mathbf{F}(z) \triangleq \mathbf{G}(z)\mathbf{E}(z)$ where the all-zero order-N $\mathbf{G}(z)$ is the propagating structure. We are adding on or peeling off a block in the lattice depending on causal or anticausal $\mathbf{G}(z)$. The order N is purposely chosen to be arbitrary so that $\mathbf{G}(z)$ can cover all classes of FB that may be unfactorizable with order-1 structures (odd-channel, for instance). The following theorem introduces a general structure for $\mathbf{G}(z)$ to propagate the LP property.

Theorem I: $\mathbf{F}(z)$ has LP and PR if and only if $\mathbf{G}(z)$ is FIR invertible and takes the form $\mathbf{G}(z) = z^{-N}\,\mathbf{D}\,\mathbf{G}(z^{-1})\,\mathbf{D}$.

Proof: Using the LP property of $\mathbf{E}(z)$ and $\mathbf{F}(z)$, we have

$$\begin{aligned}\mathbf{F}(z) &= z^{-(K-1+N)}\mathbf{D}\mathbf{F}(z^{-1})\mathbf{J} \\ &= z^{-(K-1+N)}\mathbf{D}\mathbf{G}(z^{-1})\mathbf{E}(z^{-1})\mathbf{J} \\ &= z^{-N}\mathbf{D}\mathbf{G}(z^{-1})z^{-(K-1)}\mathbf{E}(z^{-1})\mathbf{J} \\ &= z^{-N}\mathbf{D}\mathbf{G}(z^{-1})\mathbf{D}z^{-(K-1)}\mathbf{D}\mathbf{E}(z^{-1})\mathbf{J} \\ &= z^{-N}\mathbf{D}\mathbf{G}(z^{-1})\mathbf{D}\mathbf{E}(z).\end{aligned}$$

For $\mathbf{E}(z)$ and $\mathbf{F}(z)$ to have FIR inverses, it is necessary and sufficient that $\mathbf{G}(z)$ is FIR invertible and

$$\mathbf{G}(z) = z^{-N}\,\mathbf{D}\,\mathbf{G}(z^{-1})\,\mathbf{D}. \qquad \square$$

Let $\mathbf{G}(z) = \sum_{i=0}^{N}\mathbf{A}_i z^{-i}$. The above form of $\mathbf{G}(z)$ imposes interesting symmetric constraints on the matrices $\mathbf{A}_i$.

$$\begin{aligned}\mathbf{G}(z) &= z^{-N}\,\mathbf{D}\,\mathbf{G}(z^{-1})\,\mathbf{D} = z^{-N}\mathbf{D}(\textstyle\sum_{i=0}^{N}\mathbf{A}_i z^{i})\mathbf{D} \\ &= \mathbf{D}(\textstyle\sum_{i=0}^{N}\mathbf{A}_i z^{i-N})\mathbf{D} = \mathbf{D}(\textstyle\sum_{i=0}^{N}\mathbf{A}_{N-i} z^{-i})\mathbf{D} \\ &= \textstyle\sum_{i=0}^{N}(\mathbf{D}\mathbf{A}_{N-i}\mathbf{D})z^{-i} \\ \Longrightarrow\ & \mathbf{A}_i = \mathbf{D}\,\mathbf{A}_{N-i}\,\mathbf{D}. \qquad (2)\end{aligned}$$

2.2. GLBT Lattice Structure Based on the SVD

Assume that M is even. In this case, there are $\frac{M}{2}$ symmetric and $\frac{M}{2}$ antisymmetric filters [3]. Furthermore, we know that LPPRFB exists for every integer $K \geq 1$ [3, 4], i.e., these FB can be factorized by order-1 $\mathbf{G}(z)$. If $N = 1$ in Eq.(2), $\mathbf{A}_1 = \mathbf{D}\mathbf{A}_0\mathbf{D}$. $\mathbf{G}(z)$ then takes the general form of $\mathbf{G}(z) = \mathbf{A}_0 + z^{-1}\mathbf{D}\mathbf{A}_0\mathbf{D}$. It can be proven that every aforementioned non-trivial LPPRFB can be factorized using the following form for $\mathbf{A}_0$ [8]: $\mathbf{A}_0 = \frac{1}{2}\begin{bmatrix}\mathbf{U} & \mathbf{U}\\ \mathbf{V} & \mathbf{V}\end{bmatrix}$, where $\mathbf{U}$ and $\mathbf{V}$ are arbitrary $\frac{M}{2}\times\frac{M}{2}$ matrices. So, $\mathbf{G}(z)$ can be factorized as follows

$$\begin{aligned}\mathbf{G}(z) &= \tfrac{1}{2}\begin{bmatrix}\mathbf{U}+z^{-1}\mathbf{U} & \mathbf{U}-z^{-1}\mathbf{U}\\ \mathbf{V}-z^{-1}\mathbf{V} & \mathbf{V}+z^{-1}\mathbf{V}\end{bmatrix} \\ &= \tfrac{1}{2}\begin{bmatrix}\mathbf{U} & \mathbf{0}\\ \mathbf{0} & \mathbf{V}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{I}\\ \mathbf{I} & -\mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & z^{-1}\mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{I}\\ \mathbf{I} & -\mathbf{I}\end{bmatrix} \\ &\triangleq \tfrac{1}{2}\,\mathbf{\Phi}\,\mathbf{W}\,\mathbf{\Lambda}(z)\,\mathbf{W}. \qquad (3)\end{aligned}$$

Since $\mathbf{W}$ and $\mathbf{\Lambda}(z)$ have trivial inverses, $\mathbf{G}(z)$ is invertible if and only if $\mathbf{\Phi}$ is invertible, i.e. $\mathbf{U}$ and $\mathbf{V}$ are invertible. The polyphase matrix of any even-channel LPPRFB with filter length $L = KM$ can be realized by a cascade of $(K-1)$ $\mathbf{G}_i$ blocks and a zero-delay initial matrix $\mathbf{E}_0$:

$$\mathbf{E}(z) = \mathbf{G}_{K-1}(z)\,\mathbf{G}_{K-2}(z)\,\cdots\,\mathbf{G}_1(z)\,\mathbf{E}_0 \qquad (4)$$

The starting block $\mathbf{E}_0$ has no delay element, representing LPPRFB of length M, and was often chosen to be DCT [4, 5, 6]. General $\mathbf{E}_0$ satisfying Eq.(1) can be factorized as

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

For $\mathbf{E}_0$ to have PR, $\mathbf{U}_0$ and $\mathbf{V}_0$ again have to be invertible. This result should not come as a surprise. The factorization is very similar to the GenLOT's [4]; the only difference here is that $\mathbf{U}_i$ and $\mathbf{V}_i$ do not have to be orthogonal. Now, the difficulty arises: how do we completely characterize these $\frac{M}{2}\times\frac{M}{2}$ nonsingular matrices?

Recall that every invertible matrix has an SVD: $\mathbf{U}_i = \mathbf{U}_{i0}\mathbf{\Gamma}_i\mathbf{U}_{i1}$, where $\mathbf{U}_{i0}$ and $\mathbf{U}_{i1}$ are orthogonal matrices, and $\mathbf{\Gamma}_i$ is a diagonal matrix with positive elements [1]. Thus, $\mathbf{\Phi}_i$ can be further factorized as

$$\mathbf{\Phi}_i = \begin{bmatrix}\mathbf{U}_{i0} & \mathbf{0}\\ \mathbf{0} & \mathbf{V}_{i0}\end{bmatrix}\begin{bmatrix}\mathbf{\Gamma}_i & \mathbf{0}\\ \mathbf{0} & \mathbf{\Delta}_i\end{bmatrix}\begin{bmatrix}\mathbf{U}_{i1} & \mathbf{0}\\ \mathbf{0} & \mathbf{V}_{i1}\end{bmatrix}. \qquad (5)$$

The orthogonal matrices $\mathbf{U}_{i0}$, $\mathbf{U}_{i1}$, $\mathbf{V}_{i0}$, and $\mathbf{V}_{i1}$ are parameterized by $\frac{M(M-2)}{8}$ rotations each. The diagonal matrices $\mathbf{\Gamma}_i$ and $\mathbf{\Delta}_i$ are characterized by $\frac{M}{2}$ positive parameters each. The complete lattice structure is shown in Figure 2 (drawn for $M = 8$). The most general $M \times L$ GLBT can be parameterized by $\frac{LM}{2}$ parameters as expected from LP systems. However, the SVD parameterization in Eq.(5) has three advantages: (i) exact reconstruction is guaranteed structurally under a mild condition – as long as none of the diagonal coefficients is quantized to zero; (ii) SVD representation avoids the costly matrix inversion problem in the optimization process; (iii) it is much simpler to prevent the matrices from being singular or near-singular.

It is also very easy to verify that all previously reported LPPRFB's lattice structures are special cases of the new GLBT lattice. For examples, the GLT design example in [5] has $M = 8, K = 2$, $\mathbf{U}_{00}$ and $\mathbf{V}_{00}$ from the DCT, $\mathbf{U}_{01} = \mathbf{V}_{01} = \mathbf{\Gamma}_0 = \mathbf{\Delta}_0 = \mathbf{I}$, and $\mathbf{V}_1$ parameterized as a cascade of block diagonal matrices. The LBT in [6] has $M = 8, K = 2$, $\mathbf{U}_{00}$ and $\mathbf{V}_{00}$ from the DCT, $\mathbf{U}_{01} = \mathbf{V}_{01} = \mathbf{\Gamma}_0 = \mathbf{I}$, $\mathbf{\Delta}_0 = diag[\sqrt{2}\ \ 1\ \ 1\ \ 1]$, and $\mathbf{U}_1$, $\mathbf{V}_1$ to be orthogonal. When orthogonality is imposed, we get back GenLOT [4]. When $M = 2$, the lattice turns into a simplified form of Type-A system lattice in [2]. Formal proofs of the lattice's completeness and minimality will be presented in [8].

3. DESIGN EXAMPLES AND GLBT APPLICATION IN IMAGE CODING

Figure 3 presents several GLBT design examples obtained from nonlinear optimization of the new lattice coefficients with various cost functions. The analysis banks are on top. Designs in Figure 3(a), (b), and (d) are DCT-based. While increasing the GLBT length does not improve the coding gain much, it helps in the case of stopband attenuation (where longer filters are desired) as testified in Figure 3(c).

One of the most popular applications of the GLBT is image compression. Overlapping analysis filters reduce inter-block redundancy, providing higher coding efficiency of the transform coefficients, while overlapping synthesis filters whose ends decay asymptotically to zero eliminate blocking artifacts. Each bank can now be designed appropriately.

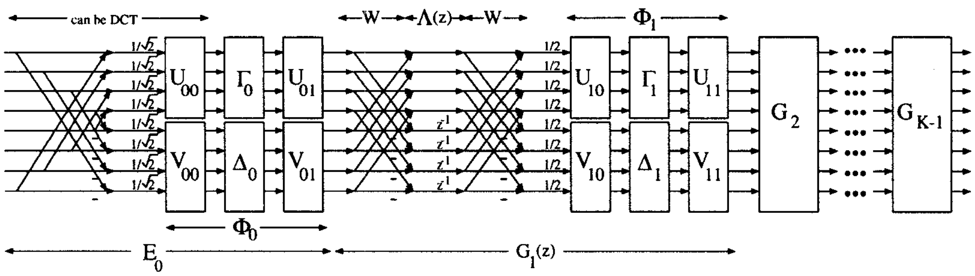


Figure 2. Lattice structure for the GLBT.

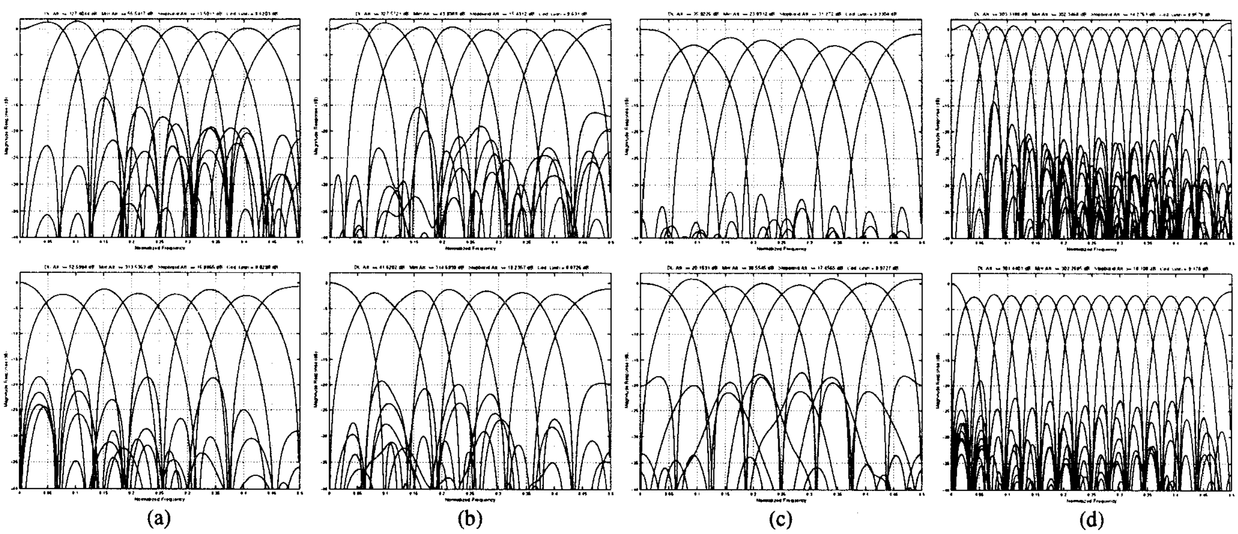


Figure 3. GLBT design examples (a) $M = 8$ $L = 16$ optimized for coding gain, DC attenuation and mirror frequency attenuation (b) $M = 8$ $L = 32$ optimized for coding gain and DC attenuation (c) $M = 8$ $L = 32$ optimized for analysis stopband attenuation (d) $M = 16$ $L = 32$ optimized for coding gain, DC and mirror frequency attenuation.

For the case of $M = 8$ $L = 16$, GLBT optimized for pure coding gain can attain 9.63 dB. However, we trade off 0.01 dB of coding gain in Figure 3(a) for high attenuation at DC, near-DC, and mirror frequencies to ensure high visual quality in the reconstructed images as well. The 16×32 GLBT in Figure 3(d) achieves an impressive coding gain of 9.96 dB. When the new transforms are incorporated into the block-transform progressive coding framework described in [9], the resulting GLBT-based embedded coders provide unrivaled objective and subjective performance as indicated in Table 1 and Figure 4(b)-(d). For a smooth image like Lena which the wavelet transform can sufficiently decorrelates, the best wavelet-based embedded coder SPIHT [7] provides a comparable performance. However, for a highly-textured image like Barbara, 16×32 GLBT coder can provide a PSNR gain of around 2.5 dB over a wide range of bit rates. The visual reconstructed image quality is also superior: texture is beautifully preserved, blocking is completely eliminated, and ringing is barely noticeable. Comparing to the optimal 8×40 GenLOT in [9], the 8×16 GLBT in Figure 3(a) already offers comparable performance at much lower computational complexity. More objective and subjective evaluation of GLBT-based progressive coders can be found at *http://saigon.ece.wisc.edu/~waveweb/Coder/index.html* along with more details on the coding scheme.

4. CONCLUSIONS

We have presented in this paper a general, albeit minimal, lattice structure for M-channel KM-length LPPRFB. The novel lattice based on the SVD provides a robust implementation and a friendly design procedure for all lapped transforms with arbitrary integer overlapping factor K.

REFERENCES

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

[2] T. Q. Nguyen and P. P. Vaidyanathan, "Two channel PR FIR QMF structures which yield linear-phase analysis and synthesis filters", *IEEE Trans. on ASSP*, pp. 676-690, May 1989.

[3] A. K. Soman, P. P. Vaidyanathan, and T. Q. Nguyen, "Linear-phase paraunitary filter banks: theory, factorizations and applications," *IEEE Trans. on SP*, vol. 41, Dec. 1993.

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

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

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

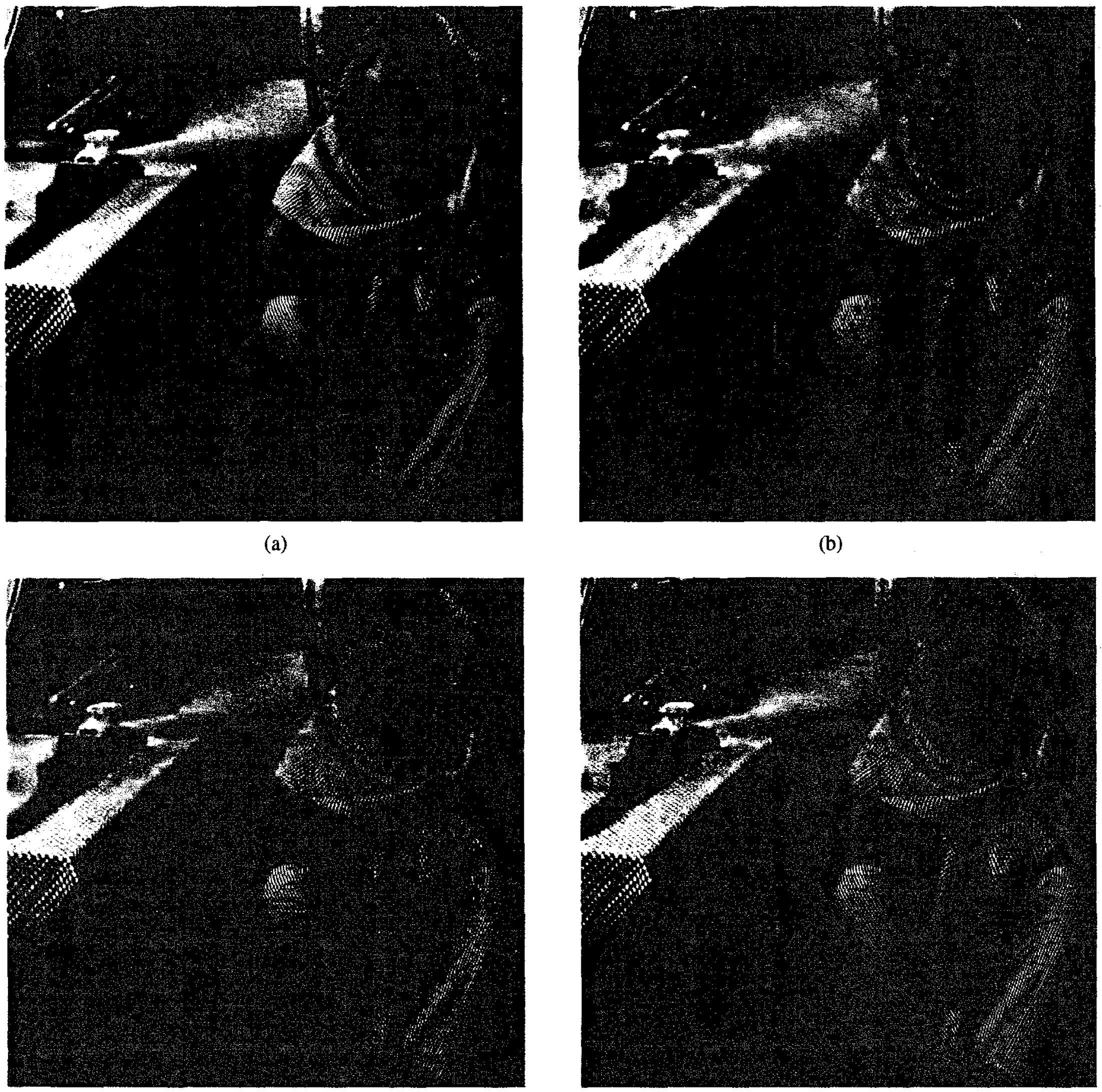
[7] A. Said and W. A. Pearlman, "A new fast and efficient image codec on set partitioning in hierarchical trees," *IEEE Trans on Circuits Syst. Video Tech.*, vol. 6, pp. 243-250, June 1996.

[8] T. D. Tran, R. de Queiroz, and T. Q. Nguyen, "Lattice structure for linear phase perfect reconstruction filter bank," to be submitted to *Trans. on Signal Processing*, 1997.

[9] T. D. Tran and T. Q. Nguyen, "A progressive transmission image coder using linear phase paraunitary filter banks," *31st Asilomar Conference on SSC*, Pacific Grove, CA, Nov. 1997.

	Lena				Goldhill				Barbara			
Comp. Ratio	SPIHT	8 x 40 GenLOT	8 x 16 GLBT	16 x 32 GLBT	SPIHT	8 x 40 GenLOT	8 x 16 GLBT	16 x 32 GLBT	SPIHT	8 x 40 GenLOT	8 x 16 GLBT	16 x 32 GLBT
1:8	40.41	40.43	40.35	40.43	36.55	36.80	36.69	36.78	36.41	38.08	37.84	38.43
1:16	37.21	37.32	37.28	37.33	33.13	33.36	33.31	33.42	31.40	33.47	33.02	33.94
1:32	34.11	34.23	34.14	34.27	30.56	30.79	30.70	30.84	27.58	29.53	29.04	30.18
1:64	31.10	31.16	31.04	31.18	28.48	28.60	28.58	28.74	24.86	26.37	26.00	27.13
1:100	29.35	29.31	29.14	29.38	27.38	27.40	27.33	27.62	23.76	24.95	24.55	25.39
1:128	28.38	28.35	28.19	28.39	26.73	26.79	26.71	26.96	23.35	24.01	23.49	24.56

Table 1. Objective coding result comparison



(a) (b)

(c) (d)

Figure 4. Embedded coding results of Barbara at 1:32 compression ratio (a) original image (b) SPIHT, 27.58 dB (c) embedded 8×16 GLBT, 29.04 dB (d) embedded 16×32 GLBT, 30.18 dB .