A Class of M-Channel Linear-Phase Biorthogonal Filter Banks and Their Applications to Subband Coding

S. C. Chan, A. Nallanathan, T. S. Ng, and P. C. K. Kwok

***Abstract*—This correspondence presents a new factorization for linear-phase biorthogonal perfect reconstruction (PR) FIR filter banks. Using this factorization, we propose a new family of lapped transform called the generalized lapped transform (GLT). Since the analysis and synthesis filters of the GLT are not restricted to be the time reverses of each other, they can offer more freedom to avoid blocking artifacts and improve coding gain in subband coding applications. The GLT is found to have higher coding gain and smoother synthesis basis functions than the lapped orthogonal transform (LOT). Simulation results also demonstrated that the GLT has significantly less blocking artifacts, higher peak signal-to-noise ratio (PSNR), and better visual quality than the LOT in image coding. Simplified GLT with different complexity/performance tradeoff is also studied.**

***Index Terms*—Filter banks, signal processing, transforms.**

I. INTRODUCTION

Subband and transform coding are two most widely used methods for data compression of image and audio signals. The performance of a subband coding system is significantly affected by the filter bank and the quantization scheme used. The theory and design of M-channel FIR filter banks have been studied extensively [7]. Due to the large number of parameters and constraints in the optimization problem, the design of general M-channel FIR filter banks is usually very difficult and time consuming. More recently, efficient filter bank structures such as the modulated lapped transform (MLT) [18], the cosine modulated filter banks (CMFB's) [17], the extended lapped transform (ELT) [3], and the lapped orthogonal transform (LOT) [2] have been proposed to reduce the design and arithmetic complexities of PR FIR filter banks. Most of these filter structures considered are orthogonal in nature, and the analysis and synthesis filters must be time reverse of each other. It has been observed by Aase and Ramstad [13] that for image coding, the functions of the analysis filter bank should maximize the energy compaction, whereas the synthesis filter bank should provide blocking-free reconstruction. In addition, the synthesis filters should be short to avoid excessive ringing and should be smooth enough to reduce blocking effects. Since the analysis and synthesis filters of an orthogonal system must be the time reverse of each other, it is very difficult to achieve these objectives simultaneously. On the other hand, biorthogonal filter banks do not suffer from this restriction, and they are more appropriate for this application.

In this correspondence, we propose a new factorization of linear-phase biorthogonal perfect reconstruction (PR) FIR filter banks. Using this factorization, it is possible to obtain linear-phase orthogonal and biorthogonal lapped transform with greater overlap. We also propose a new family of biorthogonal lapped transform called the generalized lapped transform (GLT) for subband coding applications. Like the LOT, the GLT has relatively low complexity of implementation and is based on the well-known discrete cosine transform (DCT), where many fast algorithms are also available [10], [11]. It should be noted that complete factorization of the orthogonal linear-phase PR filter banks has previously been obtained by Soman *et al.* [6]. Using this factorization, Queiroz *et al.* [5] have extended the LOT to length greater than $2M$. The paper is organized as follows. In Section II, we will introduce the factorization for linear-phase biorthogonal PR filter banks. The family of biorthogonal lapped transform and the generalized lapped transform (GLT) are discussed in Section III. Section IV is devoted to the design of the GLT and the design examples. The performance of the GLT for image coding will be presented in Section V. Finally, we summarize our work in the conclusion.

Manuscript received February 15, 1997; revised April 7, 1998. This work was supported by the Research Grants Council and the CRCG of the University of Hong Kong. The associate editor coordinating the review of this paper and approving it for publication was Prof. Ali N. Akansu.

The authors are with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong.

Publisher Item Identifier S 1053-587X(99)00745-X.

II. THE LINEAR-PHASE PR CASCADE

Fig. 1 shows the structure of an M-channel uniform filter bank (or analysis-synthesis system) with $f_i(n)$ and $g_i(n)$ the analysis and synthesis filters, respectively. The incoming signal is split into frequency bands by filtering with the analysis filters. Each subband signal is then maximally decimated by a factor of M. After processing in the subband domain, the M decimated signals will be interpolated, filtered by the synthesis filters, and added together to reconstruct the signal. In a perfect reconstruction (PR) filter bank, the input and output are identical, except for some delay [i.e., $y(n) = x(n - n_d)$]. For perfect reconstruction, $f_i(n)$ and $g_i(n)$ have to satisfy certain conditions. Let $F_{i,k}(z^M)$ and $G_{i,k}(z^M)$ be the type-I and type-II polyphase components of $F_i(z)$ and $G_i(z)$, respectively.

$$\begin{aligned} F_i(z) &= \sum_{k=0}^{M-1} z^{-k} F_{i,k}(z^M) \\ G_i(z) &= \sum_{k=0}^{M-1} z^{-(M-k-1)} G_{i,k}(z^M). \end{aligned} \tag{2.1}$$

The filter bank is PR if [7]

$$\boldsymbol{R}(z)\boldsymbol{E}(z) = z^{-d}\boldsymbol{I}_M \tag{2.2}$$

where d is a constant, $\boldsymbol{I}_M$ is the $(M \times M)$ identity matrix, and $\boldsymbol{E}(z)$ and $\boldsymbol{R}(z)$ are the polyphase matrix of the analysis and synthesis filters and are given by

$$[\boldsymbol{E}(z)]_{i,k} = F_{i,k}(z); \quad [\boldsymbol{R}(z)]_{i,k} = G_{i,k}(z). \tag{2.3}$$

In paraunitary FIR filter banks, both $\boldsymbol{E}(z)$ and $\boldsymbol{R}(z)$ are FIR matrices, and $\boldsymbol{R}(z)$ is chosen to be $\tilde{\boldsymbol{E}}(z) = \boldsymbol{E}_*^T(z^{-1})$, where $*$ and T denote the conjugation of coefficients and matrix transposition, respectively. Therefore, the impulse responses of the analysis and synthesis filters are the time reverses of each other. In biorthogonal FIR filter banks, both $\boldsymbol{E}(z)$ and $\boldsymbol{R}(z)$ are FIR matrices satisfying the PR condition. In addition, the analysis and synthesis filters are not restricted to be time reverses of each other. This provides additional freedom in coding applications, as we shall see in Section V. In addition, they can usually provide higher coding gain than their orthogonal counterparts. The linear-phase structural PR cascade introduced here [9] is motivated from the structure of the cosine modulated filter banks (CMFB's) in [3] and [17]. In CMFB's, the analysis filters $f_k(n)$ and the synthesis filters $g_k(n)$ are obtained by modulating

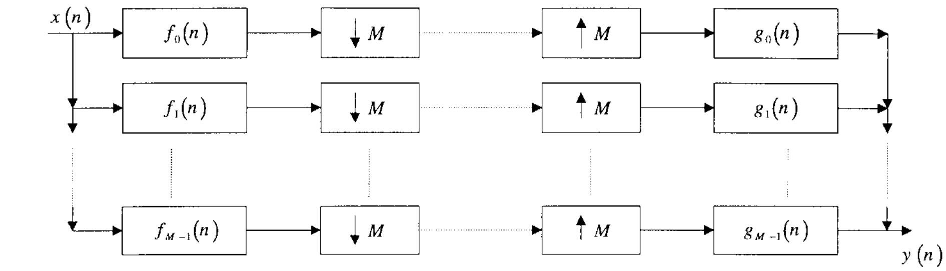
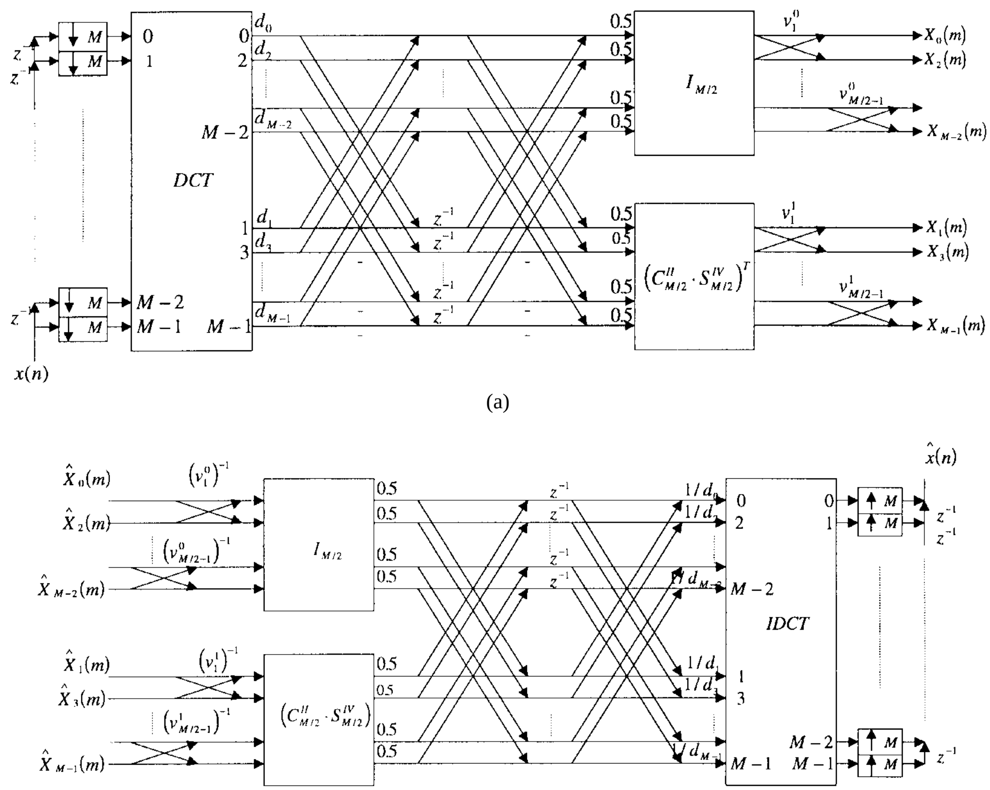


Fig. 1. M-channel uniform filter bank.



(a)

(b)

Fig. 2. (a) Flowgraph of the forward GLT. (b) Flowgraph of the inverse GLT.

TABLE I
CODING GAIN COMPARISON OF LOT AND GLT'S

	Coding Gain (dB)							
M	LOT	GLT	GLT ($N_D = 64$)	GLT ($N_D = 32$)	GLT ($N_D = 16$)	GLT ($N_D = 8$)	GLT ($N_D = 4$)	GLT ($N_D = 2$)
8	9.2189	9.6131	--	--	--	9.6118	9.6048	9.5745
16	9.7593	9.9562	--	--	9.9555	9.9527	9.9469	9.9312
32	9.9729	10.0636	--	10.0632	10.0626	10.0585	10.0536	10.0510
64	10.0541	10.0929	10.0928	10.0922	10.0913	10.0900	10.0884	10.0876

prototype filters $h(n)$ and $s(n)$ with the modulation sequences $c_{k,n}$ and $\bar{c}_{k,n}$, respectively:

$$\begin{aligned} f_k(n) &= h(n)c_{k,n} \\ g_k(n) &= s(n)\bar{c}_{k,n}, \qquad k = 0, 1, \cdots, M-1 \\ & \qquad n = 0, 1, \cdots, N-1 \end{aligned} \tag{2.4}$$

where M is the number of channels, and N is the length of the filters.

Two possible modulations are given by

$$c_{k,n} = 2\cos\left((2k+1)\,\frac{\pi}{M}\left(n - \frac{N-1}{2}\right) + (-1)^k\,\frac{k}{4}\right) \tag{2.5a}$$

or

$$c_{k,n} = \sqrt{\frac{2}{M}}\,\cos\left((2k+1)\,\frac{\pi}{M}\left(n + \frac{M+1}{2}\right)\right). \tag{2.5b}$$

Equation (2.5a) is the CMFB proposed in [17], whereas (2.5b) is

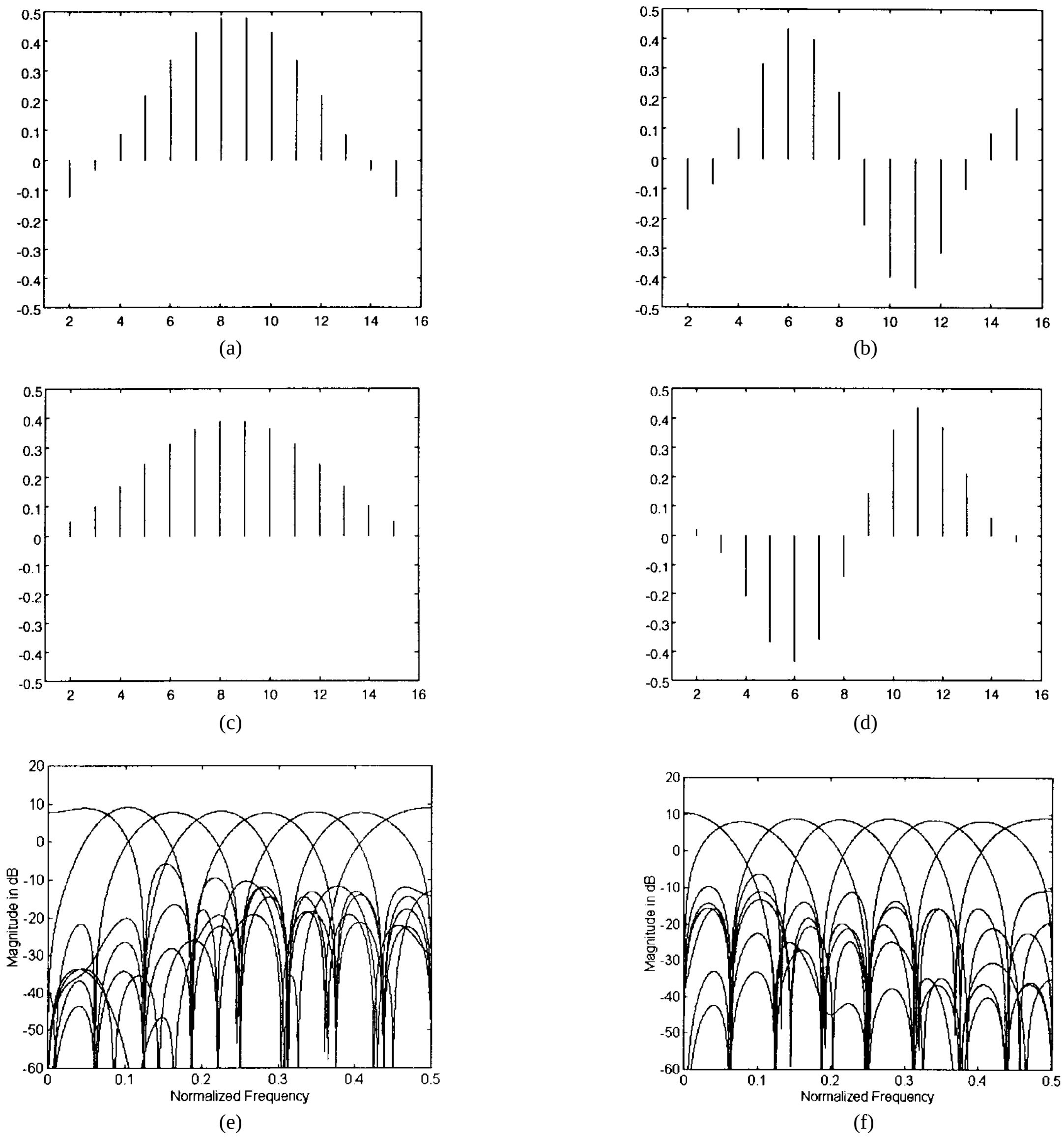


Fig. 3. Eight-channel GLT. (a) Impulse response of analysis filter $f_0(n)$. (b) Impulse response of analysis filter $f_1(n)$. (c) Impulse response of synthesis filter $g_0(n)$. (d) Impulse response of synthesis filter $g_1(n)$. (e) Frequency response of analysis filters. (f) Frequency response of synthesis filters.

the extended lapped transform (ELT) proposed in [3]. The analysis and synthesis modulations are time reverses of each other. Due to the special structure of the CMFB, it has very low design and implementation complexity. Let $H(z) = \Sigma_{q=0}^{2M-1}\ z^{-q} H_q(z^{2M})$ be the type I polyphase decomposition of the prototype filter. It can be shown that the polyphase matrix of the CMFB can be written in the form

$$\boldsymbol{E}(z) = \sqrt{M}\boldsymbol{U}_M[\boldsymbol{p}(z) + \boldsymbol{J}_M \boldsymbol{q}(z)] = \sqrt{M}\boldsymbol{U}_M \Lambda \qquad (2.6)$$

where

$$\begin{aligned}
\boldsymbol{p}(z) &= \boldsymbol{h}_0(z^2) \pm z^{-1}\boldsymbol{h}_1(z^2)\\
\boldsymbol{q}(z) &= -(-1)^m(\boldsymbol{h}_0(z^2) \mp z^{-1}\boldsymbol{h}_1(z^2))\\
\boldsymbol{h}_0(z) &= \text{diag}[H_0(-z), H_1(-z) \cdots H_{M-1}(-z)], \quad \text{and}\\
\boldsymbol{h}_1(z) &= \text{diag}[H_M(-z), H_{M+1}(-z) \cdots H_{2M-1}(-z)].
\end{aligned}$$

Here, $\boldsymbol{I_n}$ and $\boldsymbol{J_n}$ stand, respectively, for the $(n \times n)$ identity matrix and the counter identity matrix. $\boldsymbol{U}_M$ is a unitary matrix related to the modulation used. As $\boldsymbol{p}(z)$ and $\boldsymbol{q}(z)$ are diagonal matrices, matrix Λ is zero except only along its diagonal and anti-diagonal entries. Therefore, rows (columns) k and $M-k$ of Λ are orthogonal to the remaining rows (columns), and the PR condition can be simplified considerably. Without loss of generality, we assume that M is even. Results for odd M are similar. Let $p_{k,k}(z)$ and $p_{k,k}(z)$ be the kth diagonal elements of $\boldsymbol{p}(z)$ and $\boldsymbol{q}(z)$. The system is PR if the determinant of the following submatrices is equal to some delay

$$\begin{aligned}
\boldsymbol{S}_k &= \begin{bmatrix} \Lambda_{k,k} & \Lambda_{k,M-k-1} \\ \Lambda_{M-k-1,k} & \Lambda_{M-k-1,M-k-1} \end{bmatrix}\\
&= \begin{bmatrix} p_{k,k}(z) & q_{M-k-1,M-k-1} \\ q_{k,k}(z) & p_{M-k-1,M-k-1} \end{bmatrix}. \qquad (2.7)
\end{aligned}$$

Equivalently, we have

$$\det \boldsymbol{S}_k = z^{-\alpha_k}, \qquad k = 0, \cdots, (M/2) - 1. \qquad (2.8)$$

This is equivalent to the PR conditions of a set of two-channel PR filter banks. If $h(n)$ is linear phase, then $\boldsymbol{S}_k$, and it can be shown that is lossless. Therefore, it is possible to implement $\boldsymbol{S}_k$ with a two-channel lattice structure [17]. For general values of $\boldsymbol{p}(z)$ and $\boldsymbol{q}(k)$, we have the biorthogonal CMFB [19], [22], [23]. If the lossless condition on $\boldsymbol{S}_k$ is relaxed, then it is possible to use other lattice structures,

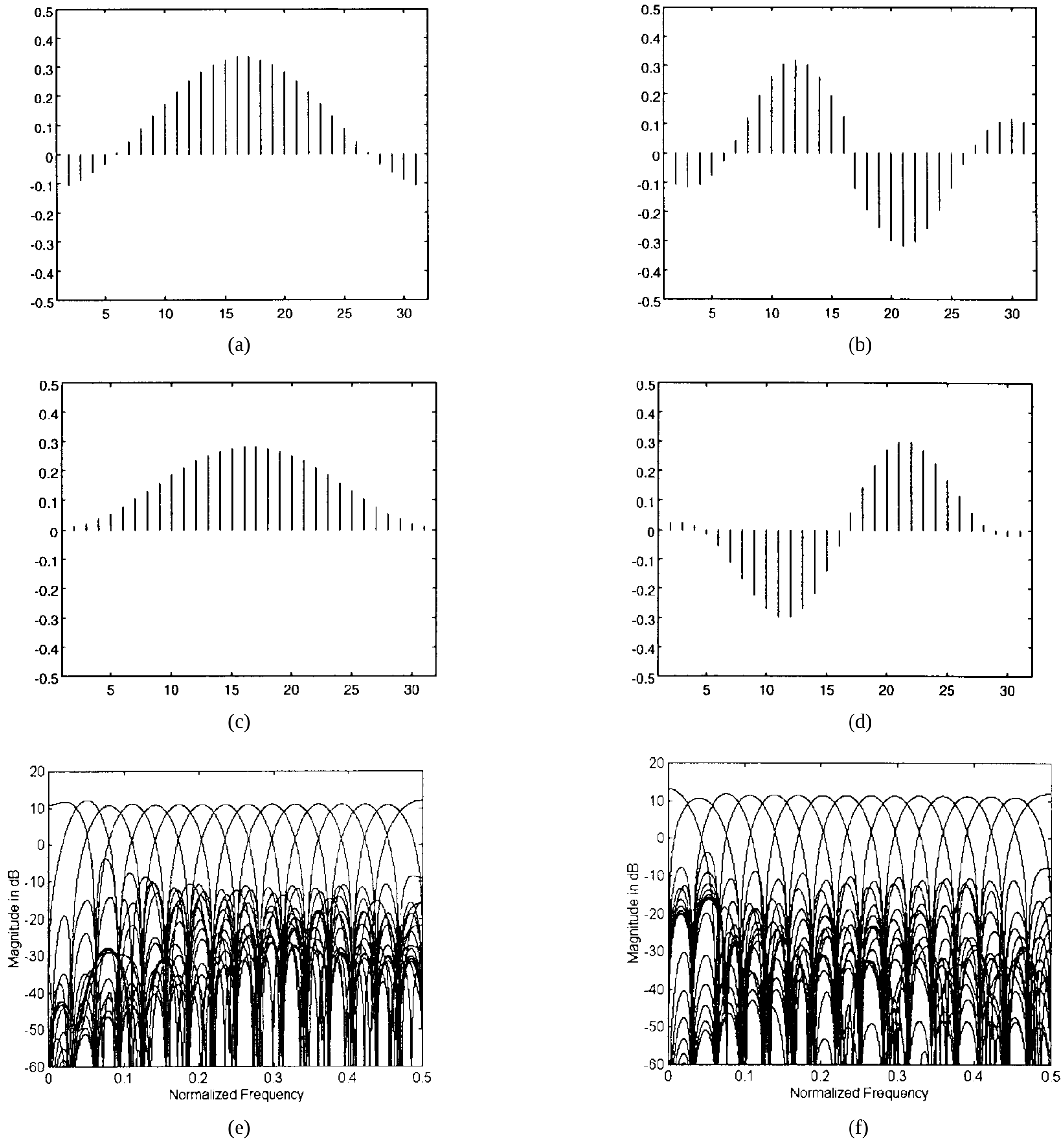


Fig. 4. Sixteen-channel GLT. (a) Impulse response of analysis filter $f_0(n)$. (b) Impulse response of analysis filter $f_1(n)$. (c) Impulse response of synthesis filter $g_0(n)$. (d) Impulse response of synthesis filter $g_1(n)$. (e) Frequency response of analysis filters. (f) Frequency response of synthesis filters.

such as the linear-phase lattice [21], as

$$S_k(z) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \prod_{\ell=1}^{K-1} \begin{pmatrix} 1 & 0 \\ 0 & z^{-1} \end{pmatrix} \begin{pmatrix} 1 & \alpha^{(k)} \\ \alpha_\ell^{(k)} & 1 \end{pmatrix}. \quad (2.9)$$

The use of the linear-phase factorization does not guarantee that the analysis filters will be linear phase. In fact, the polyphase matrix has to satisfy the following linear-phase test [20]:

$$\begin{pmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I}_{M/2} \end{pmatrix} \cdot [z^{-n}\boldsymbol{E}(z^{-1})] \cdot \boldsymbol{J}_M = \boldsymbol{E}(z) \quad (2.10)$$

for some positive integer n. Clearly, the matrix Λ satisfies the linear-phase test if the factorizations in (2.9) is used. In order for $\boldsymbol{E}(z)$ to satisfy (2.10), we must have

$$\begin{pmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I}_{M/2} \end{pmatrix} \boldsymbol{U}_M \begin{pmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I}_{M/2} \end{pmatrix} = \boldsymbol{U}_M. \quad (2.11)$$

This is only possible if $\boldsymbol{U}_M$ is a block diagonal matrix

$$\boldsymbol{U}_M = \mathrm{diag}\{\boldsymbol{U}_{00}, \boldsymbol{U}_{11}\}$$

where $\boldsymbol{U}_{00}$ and $\boldsymbol{U}_{11}$ are $(M/2) \times (M/2)$ invertible matrices. This is rather restrictive. Fortunately, several such factors can be cascaded together to form a more general linear-phase system. Suppose that the polyphase matrix $\boldsymbol{E}'(z)$ is linear phase and PR. We want to determine the condition on matrix $\boldsymbol{H}(z)$ such that the cascade $\boldsymbol{U}_M\boldsymbol{H}(z)\boldsymbol{E}'(z)$, is also linear-phase and PR. Using the linear-phase test, we get

$$\begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I}_{M/2} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_{00} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{U}_{11} \end{bmatrix} z^{-n_d}\boldsymbol{H}(z^{-1})\boldsymbol{E}'(z^{-1})\boldsymbol{J}_M = \begin{bmatrix} \boldsymbol{U}_{00} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{U}_{11} \end{bmatrix} \boldsymbol{H}(z)\boldsymbol{E}'(z) \quad (2.12)$$

for some positive integer n_d. Partitioning $\boldsymbol{H}(z)$ into four submatrices, (2.12) can be rewritten as

$$z^{-(n_d-n)} \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I}_{M/2} \end{bmatrix} \begin{bmatrix} \boldsymbol{H}_{00}(z^{-1}) & \boldsymbol{H}_{01}(z^{-1}) \\ \boldsymbol{H}_{10}(z^{-1}) & \boldsymbol{H}_{11}(z^{-1}) \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I}_{M/2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_{00}(z) & H_{01}(z) \\ \boldsymbol{H}_{10}(z) & \boldsymbol{H}_{11}(z) \end{bmatrix}. \quad (2.13)$$

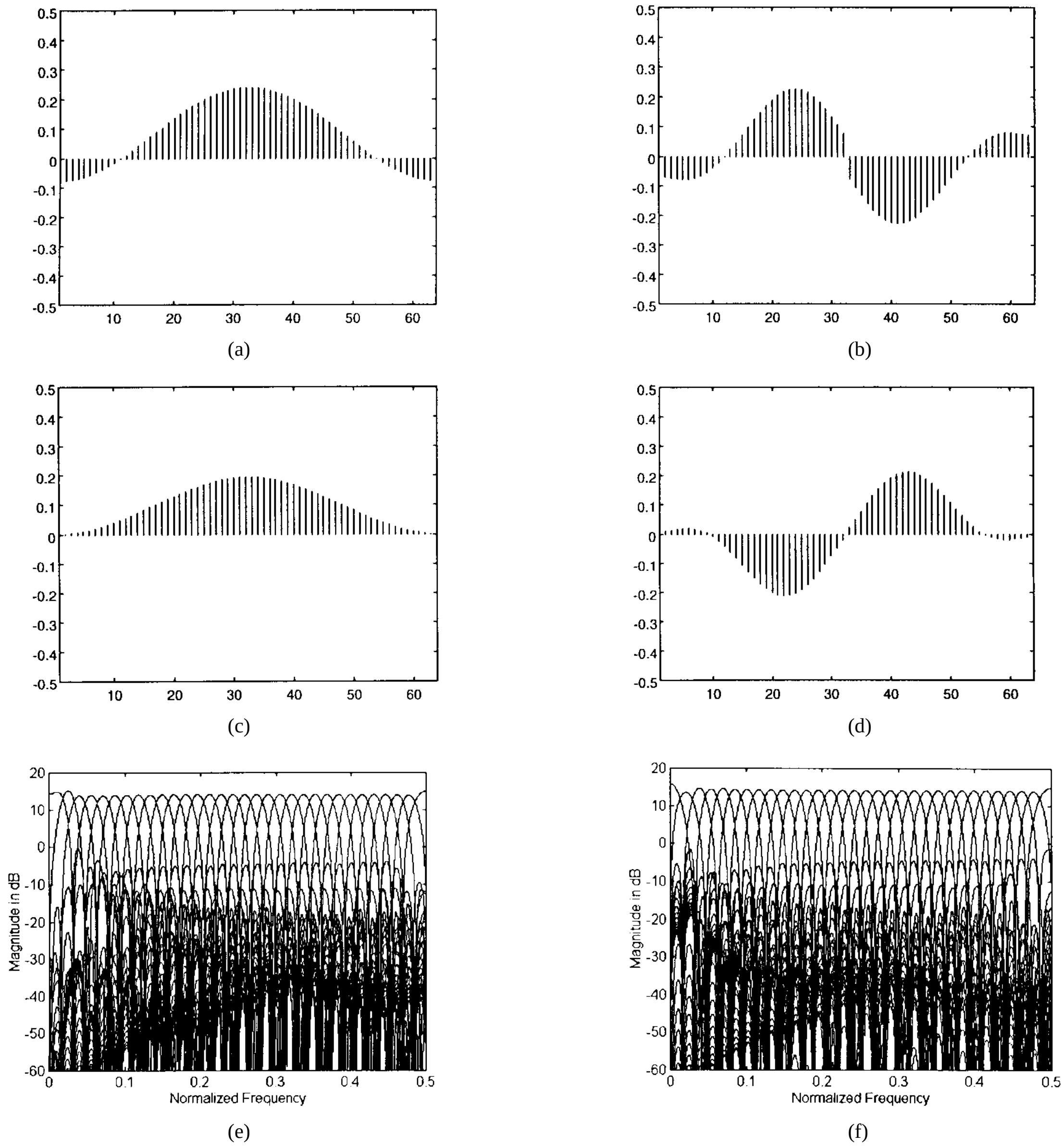


Fig. 5. Thirty-two-channel GLT. (a) Impulse response of analysis filter $f_0(n)$. (b) Impulse response of analysis filter $f_1(n)$. (c) Impulse response of synthesis filter $g_0(n)$. (d) Impulse response of synthesis filter $g_1(n)$. (e) Frequency response of analysis filters. (f) Frequency response of synthesis filters.

TABLE II
OPTIMIZED NORMALIZATION FACTORS OF THE SIMPLIFIED 8-CHANNEL GLT

N_D	d_{00}	d_{11}	d_{22}	d_{33}	d_{44}	d_{55}	d_{66}	d_{77}
2	0.871542	1.344389	--	--	--	--	--	--
4	0.851525	1.316109	0.898429	1.058310	--	--	--	--
8	0.812531	1.254635	0.858661	1.012906	0.909617	0.984497	0.953568	0.981262

From (2.13), we obtain the required condition on the submatrices as

$$\begin{aligned} z^{-(n_d-n)}\boldsymbol{H}_{00}(z^{-1}) &= \boldsymbol{H}_{00}(z) \\ z^{-(n_d-n)}\boldsymbol{H}_{01}(z^{-1}) &= -\boldsymbol{H}_{01}(z) \\ z^{-(n_d-n)}\boldsymbol{H}_{10}(z^{-1}) &= -\boldsymbol{H}_{10}(z) \\ z^{-(n_d-n)}\boldsymbol{H}_{11}(z^{-1}) &= \boldsymbol{H}_{11}(z). \end{aligned} \tag{2.14}$$

Using the concept of structural constraints as in Section II, we let $\boldsymbol{H}_{ij}(z)$ be diagonal matrices. Since the kth and the $(k+M/2)$th rows (columns) of $\boldsymbol{H}(z)$ are orthogonal to all other rows (columns), the PR condition is simplified to the PR condition of the following submatrices:

$$\boldsymbol{R}_k(z) = \begin{bmatrix} H_{00}^k(z) & H_{01}^k(k) \\ H_{10}^k(z) & H_{11}^k(z) \end{bmatrix}, \qquad k = 0, 1, \cdots, M/2 \tag{2.15}$$

where $H_{ij}^k(z)$ are the kth diagonal element of $\boldsymbol{H}_{ij}(z)$. It can be shown that a possible solution of $\boldsymbol{R}_k(z)$ is

$$\boldsymbol{R}_k(z) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \boldsymbol{S}_k^T(z) \quad \text{and}$$
$$\boldsymbol{H}(z) = \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{I}_{M/2} \\ \boldsymbol{I}_{M/2} & -\boldsymbol{I}_{M/2} \end{bmatrix} \Gamma^T(z) \tag{2.16}$$

where the $(k, k), (k, k+M/2), (k+M/2, k)$, and $(k+M/2, k+M/2)$ entries of $\Gamma^T(z)$ are obtained from matrix $\boldsymbol{S}_k(z)$. Using (2.16), we can cascade several of these factors together to obtain the factorization

$$\boldsymbol{E}(z) = \left[\prod_{\ell=0}^{K-1} \begin{bmatrix} \boldsymbol{U}_{00}^\ell & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{U}_{00}^\ell \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{I}_{M/2} \\ \boldsymbol{I}_{M/2} & -\boldsymbol{I}_{M/2} \end{bmatrix} \Gamma_\ell^T(z) \right] \boldsymbol{E}'(z). \tag{2.17}$$

A simple matrix for $\boldsymbol{E}'(z)$ is

$$\boldsymbol{E}'(z) = \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{J}_{M/2} \\ \boldsymbol{I}_{M/2} & -\boldsymbol{J}_{M/2} \end{bmatrix} \boldsymbol{R} \tag{2.18}$$

where $\boldsymbol{R}$ is a persymmetric matrix satisfying $\boldsymbol{R} = \boldsymbol{JRJ}$. If $\Gamma_\ell(z)$ is obtained from (2.9) with $K = 2$, (2.17) represents a cascade structure of linear-phase PR filter bank

$$\begin{aligned} \boldsymbol{E}(z) = \Bigg[\prod_{\ell=0}^{K-1} & \begin{bmatrix} \boldsymbol{U}_{00}^\ell & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{U}_{00}^\ell \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{I}_{M/2} \\ \boldsymbol{I}_{M/2} & -\boldsymbol{I}_{M/2} \end{bmatrix} \\ & \cdot \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{A}_{M/2} \\ \boldsymbol{A}_{M/2} & -\boldsymbol{I}_{M/2} \end{bmatrix} \\ & \cdot \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{0} \\ \boldsymbol{0} & z^{-1}\boldsymbol{I}_{M/2} \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{I}_{M/2} \\ \boldsymbol{I}_{M/2} & -\boldsymbol{I}_{M/2} \end{bmatrix} \Bigg] \\ & \cdot \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{J}_{M/2} \\ \boldsymbol{I}_{M/2} & -\boldsymbol{J}_{M/2} \end{bmatrix} \boldsymbol{R} \end{aligned} \tag{2.19}$$

where $\boldsymbol{A}_{M/2}$ are diagonal matrices containing the multipliers α's. It can be shown that the matrix

$$\begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{A}_{M/2} \\ \boldsymbol{A}_{M/2} & -\boldsymbol{I}_{M/2} \end{bmatrix}$$

can be absorbed into the block diagonal matrix diag$[\boldsymbol{U}_{00}^\ell \boldsymbol{U}_{11}^\ell]$. Therefore, it can be removed from (2.19). Furthermore, if $\boldsymbol{U}_{ii}^\ell, i = 1, 2$ are unitary (orthogonal) matrices, the filter bank will be paraunitary (orthogonal) as well. In next section, we shall propose a family of biorthogonal lapped transforms using this factorization.

III. A Family of Biorthogonal Lapped Transform

An example of the orthogonal representation in (2.19) is the lapped orthogonal transform (LOT), which is an orthogonal filter bank with length $2M$. The LOT was originally proposed to reduce the blocking artifacts in traditional transform coding of images using the discrete cosine transform (DCT). The corresponding polyphase matrix is given by

$$\begin{aligned} \boldsymbol{E}(z) = \frac{1}{2} \boldsymbol{P} & \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{0}_{M/2} \\ \boldsymbol{0}_{M/2} & (\boldsymbol{C}_{M/2}^{II} \boldsymbol{S}_{M/2}^{IV})^T \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{I}_{M/2} \\ \boldsymbol{I}_{M/2} & -\boldsymbol{I}_{M/2} \end{bmatrix} \\ & \cdot \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{0}_{M/2} \\ \boldsymbol{0}_{M/2} & z^{-1}\boldsymbol{I}_{M/2} \end{bmatrix} \boldsymbol{R}_0 \end{aligned} \tag{3.1}$$

where $\boldsymbol{R}_0 = \boldsymbol{P}' \mathrm{diag}\{\boldsymbol{B}_2, \cdots, \boldsymbol{B}_2\} \boldsymbol{C}_M^{II} \boldsymbol{J}_M$, and

$$\boldsymbol{B}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

$\boldsymbol{C}_M^k$ and $\boldsymbol{S}_M^k$ denote the type-k length-M discrete cosine and sine transforms, respectively. $\boldsymbol{P}$ is a permutation matrix that permutes the kth and the $(k+M/2)$th rows to the $(2k)$th and $(2k+1)$th $(k = 0, \cdots, M/2-1)$ rows, respectively. $\boldsymbol{P}'$ is a permutation matrix that permutes the $(2k)$th and $(2k+1)$th rows to the kth and $(k+M/2)$th $(k = 0, \cdots, M/2-1)$ rows, respectively. Using the linear-phase factorization in (2.19), we introduce a biorthogonal generalization of the LOT called the generalized lapped transform (GLT). The polyphase matrix of the GLT is defined as

$$\begin{aligned} \boldsymbol{E}(z) = \frac{1}{2} \boldsymbol{P} & \begin{bmatrix} \boldsymbol{U}_{00} & \boldsymbol{0}_{M/2} \\ \boldsymbol{0}_{M/2} & \boldsymbol{U}_{11} \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{0}_{M/2} \\ \boldsymbol{0}_{M/2} & (\boldsymbol{C}_{M/2}^{II} \boldsymbol{S}_{M/2}^{IV} \end{bmatrix} \\ & \cdot \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{I}_{M/2} \\ \boldsymbol{I}_{M/2} & -\boldsymbol{I}_{M/2} \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_{M/2} & \boldsymbol{0}_{M/2} \\ \boldsymbol{0}_{M/2} & z^{-1}\boldsymbol{I}_{M/2} \end{bmatrix} \boldsymbol{R}_I \end{aligned} \tag{3.2}$$

where $\boldsymbol{U}_{00}$ and $\boldsymbol{U}_{11}$ are block diagonal invertible matrices $\boldsymbol{R}_1 = \boldsymbol{P}' \mathrm{diag}\{\boldsymbol{B}_2 \cdots \boldsymbol{B}_2\} \boldsymbol{D} \boldsymbol{C}_M^{II} \boldsymbol{J}_M$, and $\boldsymbol{D}$ is a diagonal matrix.

The lapped transform with greater overlap can similarly be defined. In this correspondence, we parameterize the matrix $\boldsymbol{U}_{ii}$ by products of block diagonal (2×2) invertible matrices

$$\boldsymbol{U}_{ii} = \prod_{k=1}^{M/2-1} \boldsymbol{v}_k^i \tag{3.3}$$

where

$$\boldsymbol{v}_k^i = \begin{bmatrix} \boldsymbol{I}_{k-1} & & & \boldsymbol{0} \\ & x_k^i & y_k^i & \\ & y_k^i & x_k^i & \\ \boldsymbol{0} & & & \boldsymbol{I}_{M-k-1} \end{bmatrix} \quad \text{and}$$
$$(\boldsymbol{v}_k^i)^{-1} = \frac{1}{((x_k^i)^2 - (y_k^i)^2)} \cdot \begin{bmatrix} \boldsymbol{I}_{k-1} & & & \boldsymbol{0} \\ & x_k^i & -y_k^i & \\ & -y_k^i & x_k^i & \\ \boldsymbol{0} & & & \boldsymbol{I}_{M-k-1} \end{bmatrix}.$$

The signal flow graphs for the M-channel length $2M$ forward and inverse GLT are shown in Fig. 2(a) and (b). Parameterization using the LU and Gauss–Jordan factorizations are also possible, but the present choice has the advantage that $\boldsymbol{U}_{ii}$ is simple and diagonal dominance. The parameters x_k^i, y_k^i, d_{ii} (normalization factors) can be used to maximize the objective function such as the coding gain.

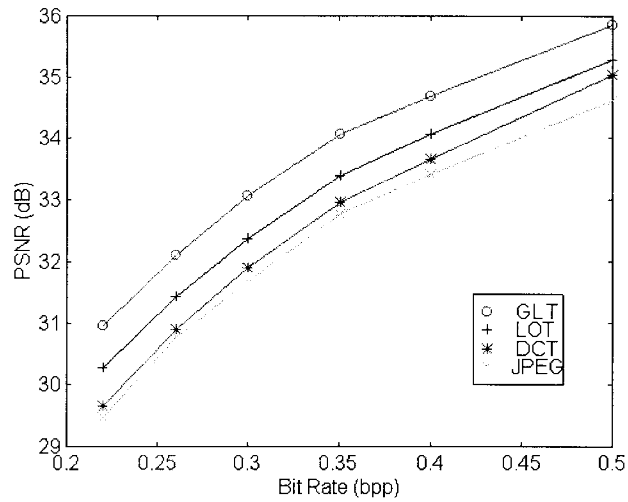


Fig. 6. PSNR comparison of JPEG, DCT (with bit allocation), LOT and GLT for coding the image "Lena."

IV. Design Procedure and Examples

An important issue in transform and subband coding is the efficiency of the transform or filter bank employed. The coding gain is

(a)

(b)

(c)

(d)

Fig. 7. Performance of different algorithms on image "Lena." (a) Original. (b) 0.3 b/pixel using DCT. (c) 0.3 b/pixel using LOT. (d) 0.3 b/pixel using GLT $(N_D = 8)$.

frequently used as an effective measure of transform efficiency. For the orthogonal transform, the coding gain G_{TC} is given by

$$G_{TC} = \frac{\frac{1}{M}\sum_{i=0}^{M-1}\sigma_i^2}{\left(\sum_{i=0}^{M-1}\sigma_i^2\right)^{1/M}} \tag{4.1}$$

where σ_i^2 is the variance of the ith transform coefficients. For image coding, the input is usually modeled as a first-order autoregressive process. The corresponding covariance matrix $\boldsymbol{R}_{xx}$ is given by

$$[\boldsymbol{R}_{xx}]_{ij} = \rho^{|i-j|}, \qquad 0 \le i, j \le M-1 \tag{4.2}$$

where ρ is the correlation coefficient, which is usually taken as 0.95. For biorthogonal filter bank, the coding gain formula [15] is

$$G_{SBC} = \frac{1}{M\left[\prod_{k=0}^{M-1} A_k B_k\right]^{1/M}} \tag{4.3}$$

where

$$A_k = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} h_k(i) l h_k(j)\rho^{|i-j|}$$

and

$$B_k = \frac{1}{M}\sum_{j=0}^{L-1} g_k^2(j) \qquad 0 \le k \le M-1.$$

$h_k(n)$ and $g_k(n)$ are the impulse responses of the kth analysis and synthesis filters of length L. In addition to the coding gain, we can incorporate other desirable properties in the design process. For example, it is possible to reduce the blocking effects by minimizing the stopband energies of the synthesis lowpass filter $g_0(n)$ with an interval of $\pm I$ (e.g., 0.002) at frequencies $(i/M), i = 1, \cdots, M/2$ [14]. We can also force the synthesis filters to assume small values at the ends of the impulse responses to improve their smoothness. Next, we shall present some design examples to be used in subband coding of images and audio.

All examples presented here were obtained using the constrained optimization program, *NCONF*, in the *IMSL* library. For audio coding, the number of channels can be as large as 32. Therefore, we present the design of GLT for eight, 16, 32, and 64 channels. In each case, the coding gain is calculated using an AR(1) process with correlation coefficient $\rho = 0.95$. The parameters x_k^i, y_k^i, and d_{ii} are obtained by maximizing the coding gain of the system. Table I shows the resulting coding gains (G_{SBC}) after the optimization. The coding gain of the LOT is also given as a comparison.

It can be seen that the GLT has a higher coding gain than the LOT. Figs. 3–5 show the impulse responses of the first two analysis and synthesis filters and the frequency responses of the GLT with $M = 8, 16,$ and 32, respectively. It can be seen that the synthesis filters are much smoother than the analysis filters and decay to zero at both ends to reduce the blocking artifacts.

We have also investigated a simplified version of the GLT, where the matrices $\boldsymbol{U}_{ii}$ are chosen as identity matrix. In this case, only the parameters d_{ii}'s are present, and the simplified GLT would require M more multiplications than the LOT. Since the higher order basis functions are less important in reducing the blocking artifacts and improving the coding gain than the lower order basis functions, we can further reduce the arithmetic complexity by setting their normalization factors (d_{ii}) to 1. This will considerably reduce the arithmetic complexity of the GLT for real-time applications. These remaining free parameters $d_{ii}, i = 0, \cdots, N_D - 1$, are used to optimize the coding gain of the system. Here, N_D is the number of normalization factors that we have retained in the GLT. The coding gains of the simplified GLT with different values of N_D are shown in Table I. The optimized normalization factors for the eight-channel simplified GLT are shown in Table II. It can be seen that the coding gains of the simplified GLT's are quite close to each other. In addition, the impulse responses of the simplified GLT's are also similar to those of the GLT.

V. Coding Performance

In this section, we will compare the performance of the DCT, LOT and GLT in image coding. (8 × 8) DCT and separable eight-channel length-16 LOT and GLT $(N_D = 8)$ are used for simulation. For a fair comparison between the systems, we follow the JPEG quantization scheme with the DCT replaced by the LOT and the GLT. The quantization table is obtained using the bit allocation algorithm proposed in [24], and no human visual model has been used. Fig. 7 shows the images "Lena" (512 × 512 8-bit grey scale) encoded to 0.3 b/pixel using the various algorithms. The PSNR comparison of the various algorithms is shown in Fig. 6. The JPEG algorithm using the default quantization table is also given as a comparison. It can be seen that the image encoded using the GLT has significantly less blocking artifact and higher PSNR than that of the DCT and the LOT.

VI. Conclusion

In this correspondence, we have presented a new factorization for linear-phase biorthogonal perfect reconstruction (PR) FIR filter banks. Using this factorization, it is possible to obtain linear-phase orthogonal and biorthogonal lapped transform with greater overlap. We have also proposed a new family of biorthogonal lapped transform called the generalized lapped transform (GLT) for subband coding applications. The GLT is found to have higher coding gain and smoother synthesis basis functions than the lapped orthogonal transform (LOT). Simulation results also demonstrated that the GLT has significantly less blocking artifacts, higher peak signal to noise ratio (PSNR), and better visual quality than the LOT in image coding. Simplified GLT with different complexity/performance tradeoff is also given to further reduce the implementation complexity.

Acknowledgment

The authors would like to thank Dr. W. C. Fong of the Department of Electrical and Electronic Engineering, The University of Hong Kong, for her help in developing the comparison in Section V of this work. The first author would like to thank Dr. H. S. Malvar for introducing to him a preprint of the paper [25], which leads to the study of the simplified GLT.

References

[1] H. S. Malvar and D. H. Stalin, "The LOT: Transform coding without blocking effects," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 553–559, Apr. 1989.

[2] H. S. Malvar, "Lapped transforms for efficient transform/subband coding," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 969–978, June 1990.

[3] ——, "Extended lapped transforms: Properties, applications, and fast algorithms," *IEEE Trans. Signal Processing*, vol. 40, pp. 2703–2714, Nov. 1992.

[4] ——, *Signal Processing with Lapped Transform*. Norwood, MA: Artech House, 1992.

[5] R. L. de Queiroz and T. Q. Nguyen " The GenLOT: Genaralized linear phase lapped orthogonal transform," *IEEE Trans. Signal Processing*, vol. 44, pp. 497–507, Mar. 1996.

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

[7] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs: Printice-Hall, 1993.

[8] A. N. Akansu and M. J. T. Smith, *Subband and Wavelet Transforms Design and Applications*. Boston, MA: Kluwer, 1996.

[9] S. C. Chan, "Quadrature modulated filter banks," *Proc IEEE Int. Symp. Circuits Syst.*, London, U.K., May 1994, vol. 2, pp. 501–504.

[10] S. C. Chan and K. L. Ho, "Fast algorithms for computing the discrete cosine transform," *IEEE Trans. Circuits Syst. II*, vol. 39, pp. 185–190, Mar. 1992.

[11] ——, "Direct algorithms for computing the discrete sinusoidal Transform," *Proc. Inst. Elect. Eng. F*, vol. 137, pp. 433–442, Dec. 1990.

[12] S. C. Chan, "The generalized lapped transform (GLT) for subband coding applications," in *Proc ICASSP*, Detroit, MI, 1995, pp. 1508–1511.

[13] S. O. Aase and T. A. Ramstad, "Parallel FIR filter banks for robust subband image coding," *Proc IEEE Int Conf. Acoust., Speech, Signal Process.*, Minneapolis, MN, Apr. 1993, vol. V, pp. 566–569.

[14] S. O. Aase, "Image subband coding artifacts: analysis and remedies," Ph.D. dissertation, Norwegian Inst. Technology, Trondheim, 1993.

[15] J. Katto and Y. Yasuda, "Performance evaluation of subband coding and optimization of its filter coefficients" in *Proc. SPIE's Visual Commun. Image Process.*, Nov. 1991, pp. 95–106.

[16] S. O. Aase and T. A. Ramstad, " On the optimality of nonunitary filter banks in subband coders," *IEEE Trans. Image Processing*, vol. 4, pp. 1585–1591, Dec. 1995.

[17] R. D. Koilpillai and P. P. Vaidyanathan, "Cosine-modulated FIR filter banks satisfying perfect reconstruction," *IEEE Trans. Signal Processing*, vol. 41, pp. 770–783, Apr. 1992.

[18] J. P. Princen and A. P. Bradley, "Analysis-synthesis filter bank design based on time domain aliasing cancelation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1153–1161, Oct. 1986.

[19] T. Q. Nguyen, "A class of generalized cosine-modulated filter bank," in *Proc. IEEE ISCAS*, 1992, pp. 943–946.

[20] M. Vetterli and D. L. Gall, "Perfect reconstruction FIR filterbanks: some properties and factorizations," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 1057–1071, July 1989.

[21] T. Q. Nguyen and P. P. Vaidyanathan, "Two-channel perfect reconstruction FIR QMF structures which yield linear phase PR analysis and synthesis filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 676–690, May 1989.

[22] T. Q. Nguyen and P. N. Heller, "Biorthogonal cosine-modulated filter bank," *Proc. IEEE ICASSP*, 1996, pp. 1471–1474.

[23] Yi. Luo, S. C. Chan, and K. L. Ho, "Theory and Design of Arbitrary-Length Biorthogonal Cosine-Modulated Filter Banks," *Proc. IEEE ISCAS*, 1997, Hong Kong.

[24] W. C. Fong, S. C. Chan, and K. L. Ho, " Designing JPEG quantization matrix using rate-distortion approach and human visual system model," *Proc. IEEE ICC*, 1997.

[25] H. S. Malvar, "Biorthogonal and nonuniform lapped transform for transform coding with reduced blocking and ringing artifacts," preprint, 1997.