The LiftLT: Fast-Lapped Transforms via Lifting Steps

Trac D. Tran, *Member, IEEE*

***Abstract*—This paper introduces a class of multiband linear phase-lapped biorthogonal transforms with fast, VLSI-friendly implementations via lifting steps called the LiftLT. The transform is based on a lattice structure that robustly enforces both linear phase and perfect reconstruction properties. The lattice coefficients are parameterized as a series of lifting steps, providing fast, efficient, in-place computation of the transform coefficients. The new transform is designed for applications in image and video coding. Compared to the popular 8 × 8 DCT, the 8 × 16 LiftLT only requires one more multiplication, 22 more additions, and six more shifting operations. However, image coding examples show that the LiftLT is far superior to the DCT in both objective and subjective coding performance. Thanks to properly designed overlapping basis functions, the LiftLT can completely eliminate annoying blocking artifacts. In fact, the novel LT's coding performance consistently surpasses that of the much more complex 9/7-tap biorthogonal wavelet with floating-point coefficients. More importantly, the transform's block-based nature facilitates one-pass sequential block coding, region-of-interest coding/decoding, and parallel processing.**

I. Introduction

MULTIBAND transforms have long found applications 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, creating discontinuities at block boundaries. The wavelet transform with long overlapping bases has elegantly solved the blocking problem. However, the transform's complexity is significantly higher than the DCT's. Except for a few special cases, the wavelet transform generally requires many more operations per output coefficient, and it may need a large memory buffer in its implementation. Another interesting alternative is the lapped transform [2], where only a small number of pixels from adjacent blocks are borrowed to produce the transform coefficients of the current block. Lapped transforms outperform the DCT on two counts. From the analysis viewpoint, it takes into account interblock correlation and hence provides better energy compaction, and from the synthesis viewpoint, its basis functions decay asymptotically to zero at the ends, reducing blocking discontinuities.

Nevertheless, lapped transforms have not yet been able to replace the DCT in international standards. One reason is that the modest improvement in coding performances is not enough to justify the increase in computational complexity. In this paper, we introduce a family of lapped biorthogonal transforms based on a minimal number of dyadic-rational lifting steps. The resulting transform called LiftLT is not only fast-computable and VLSI-suited, but it also consistently outperforms state-of-the-art wavelets given the same quantizer and entropy coder. Despite its simplicity, the LiftLT provides a significant improvement in reconstructed image quality over the traditional DCT: blocking is completely eliminated, while ringing is reasonably contained at medium and high compression ratios.

Manuscript received October 15, 1999. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. A. M. Sayeed.

The author is with the Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218-2686 USA (e-mail: ttran@ece.jhu.edu).

Publisher Item Identifier S 1070-9908(00)05101-4.

II. Review

We limit the discussions on lapped transforms to M-channel uniform linear phase-perfect reconstruction filter banks (LP-PRFB's), where analysis and synthesis filters have the same length KM. The most general lattice for M-channel linear phase-lapped biorthogonal transforms (GLBT) is presented in [3], [4]. The polyphase matrix $\boldsymbol{E}(z)$ can be factorized as

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

$$\boldsymbol{G}_i(z) = \frac{1}{2}\begin{bmatrix}\boldsymbol{U}_i & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{V}_i\end{bmatrix}\begin{bmatrix}\boldsymbol{I} & \boldsymbol{I}\\ \boldsymbol{I} & -\boldsymbol{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{0} & z^{-1}\boldsymbol{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{I} & \boldsymbol{I}\\ \boldsymbol{I} & -\boldsymbol{I}\end{bmatrix} \triangleq \frac{1}{2}\boldsymbol{\Phi}_i\boldsymbol{W}\boldsymbol{\Lambda}(z)\boldsymbol{W} \tag{2}$$

and

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

This lattice results in all filters having length $L = KM$. K is often called the overlapping factor. Each cascading structure $\boldsymbol{G}_i(z)$ increases the filter length by M. All $\boldsymbol{U}_i$ and $\boldsymbol{V}_i$, $i = 0, 1, \cdots, K-1$ are arbitrary $(M/2) \times (M/2)$ invertible matrices. These free invertible matrices hold the free design parameters, and they can be parameterized by the singular value decomposition (SVD) [3], [4]. The complete lattice of the analysis bank is depicted in Fig. 1.

III. Complete LT Lattice via Lifting

As previously mentioned, an M-channel LT with overlapping factor K can be completely parameterized by $2K$ invertible matrices of size $M/2$. Under the SVD, as in [3], [4], each invertible matrix can be completely characterized by a diagonal matrix and two orthogonal matrices. The SVD parameterization of an arbitrary invertible matrix is shown in Fig. 2(a) (drawn for $M = 8$). Since each $(M/2) \times (M/2)$ orthogonal matrix can be factorized into $(M(M-2)/8)$ plane rotations, the most general

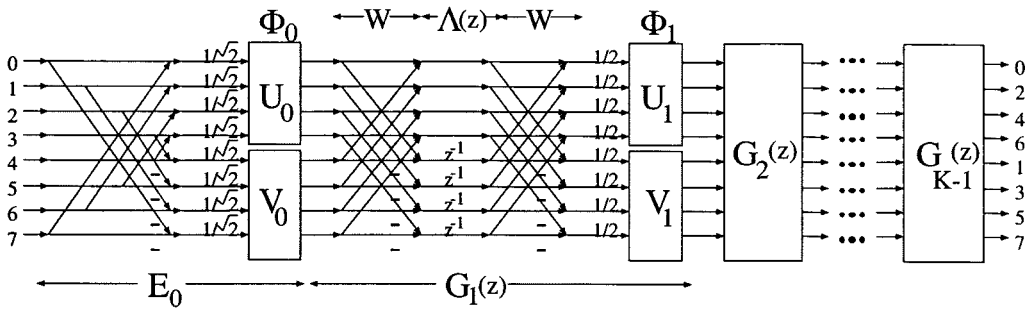

Fig. 1. The most general lattice structure for linear phase-lapped transforms with filter length $L = KM$.

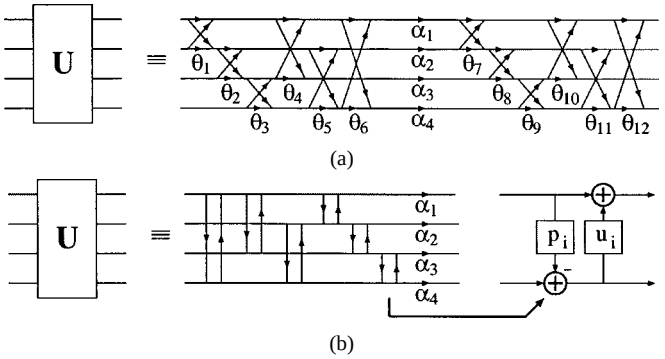

Fig. 2. Parameterization of an invertible matrix (a) via the SVD and (b) via lifting steps.

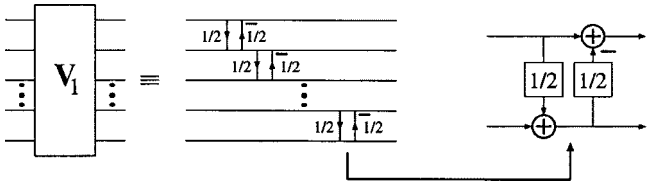

Fig. 3. V_1 matrix of the general M-channel LiftLT.

LT lattice consists of $(KM(M-2)/2)$ rotations θ_i and KM diagonal scaling factors α_i.

One disadvantage of the aforementioned SVD-based parameterization is the redundancy in computational complexity. Since each rotation angle takes four multiplications and two additions (some manipulations yield three multiplications and three additions), the actual cost of performing a matrix multiplication in the SVD structure actually surpasses that of direct multiplication. In this paper, we propose to characterize these invertible matrices using shears (also known as the lifting steps or the ladder structures). It is not too difficult to prove that any $N \times N$ invertible matrix can be completely characterized by $N(N-1)$ shears, N diagonal scaling factors, and possibly permutation matrices (this simply follows from the Gauss–Jordan elimination process). The final ladder-based parameterization is illustrated in Fig. 2(b). Under this parameterization, the computational complexity is, at most, equal to that of direct multiplication. In most cases, the computational complexity can be reduced significantly by setting the diagonal scaling factors to unity and/or choosing the lifting coefficients to be dyadic, as described in the next section. The most general LT lattice now consists of $(KM(M-2)/2)$ lifting steps p_i, u_i and KM diagonal scaling factors α_i.

IV. Fast LiftLT

This section is devoted to the design of a high-performance yet low-complexity lapped transform based on fast-lifting steps called LiftLT that can hopefully replace the DCT in the near future. To minimize the transform's complexity, we choose a small overlapping factor $K = 2$ and set the initial stage $\boldsymbol{E}_0$ to be the $M \times M$ DCT itself, where M is even. Following the excellent observation in [6], we apply a scaling to the first DCT's antisymmetric basis to generate synthesis LT-basis functions whose end values decay smoothly to zero—a crucial requirement in blocking artifacts elimination. However, instead of scaling the analysis by $\sqrt{2}$ and the synthesis by $(1/\sqrt{2})$, we opt for $(25/16)$ and its inverse $(16/25)$ since they allow the implementation of both the analysis and synthesis bank in integer arithmetic. Another nice value that also works almost as well as $(25/16)$ is $(5/4)$.

After two series of $\pm$ 1 butterflies $\boldsymbol{W}$ and the delay chain $\boldsymbol{\Lambda}(z)$, the LT symmetric basis functions already have good attenuation, especially at DC ($\omega = 0$). Hence, we can comfortably set $\boldsymbol{U}_1 = \boldsymbol{I}_{M/2}$. Now, there is only the parameterization of $\boldsymbol{V}_1$ left. There are $(M/2)((M/2)-1)$ free lifting steps and $(M/2)$ free diagonal scaling factors here. However, we propose to construct $\boldsymbol{V}_1$ by cascading $2((M/2)-1)$ lifting steps. Each can be implemented using only one simple bit shift and one addition, as shown in Fig. 3.

The final LiftLT's lattice structure is presented in Fig. 4. The frequency and impulse responses of the 8 $\times$ 16 LiftLT's basis functions are depicted in Fig. 5. The LiftLT should be sufficiently fast for many applications, especially in hardware, since most of the additional computation comes from the two butterflies and the six shift-and-add lifting steps. It is even faster than the type-I fast LOT [2]. Besides its low complexity, the LiftLT also possesses many desirable characteristics such as high energy compaction, low attenuation near DC, and smoothly-decaying synthesis basis functions to eliminate blocking artifacts completely. For the AR(1) image model with $\rho = 0.95$, the eight-channel LiftLT in Fig. 4 achieves a coding gain of 9.54 dB. The comparison of complexity between the LiftLT and other popular transforms is tabulated in Table I. Notice that the LiftLT's performance is already very close to that of the optimal GLBT (9.63 dB coding gain) [3], [4], whereas its complexity is the lowest among the transforms in comparison, excluding the DCT's. The 12 $\times$ 24 and the 16 $\times$

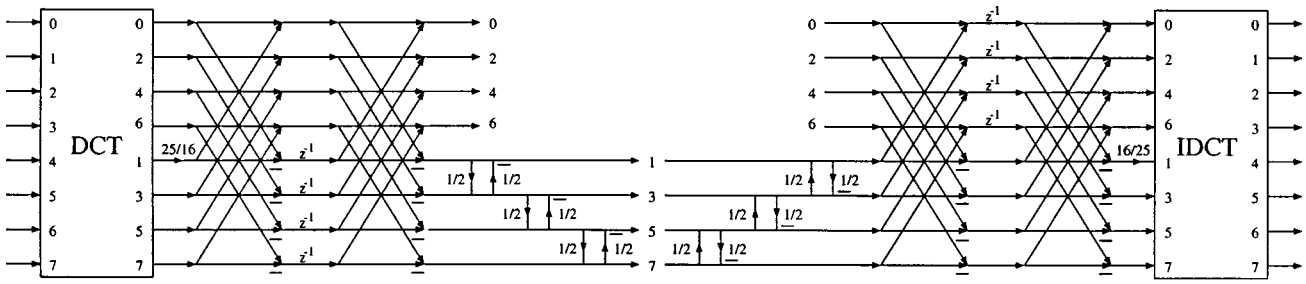

Fig. 4. Complete analysis/synthesis unnormalized LiftLT lattice (drawn for $M = 8$).

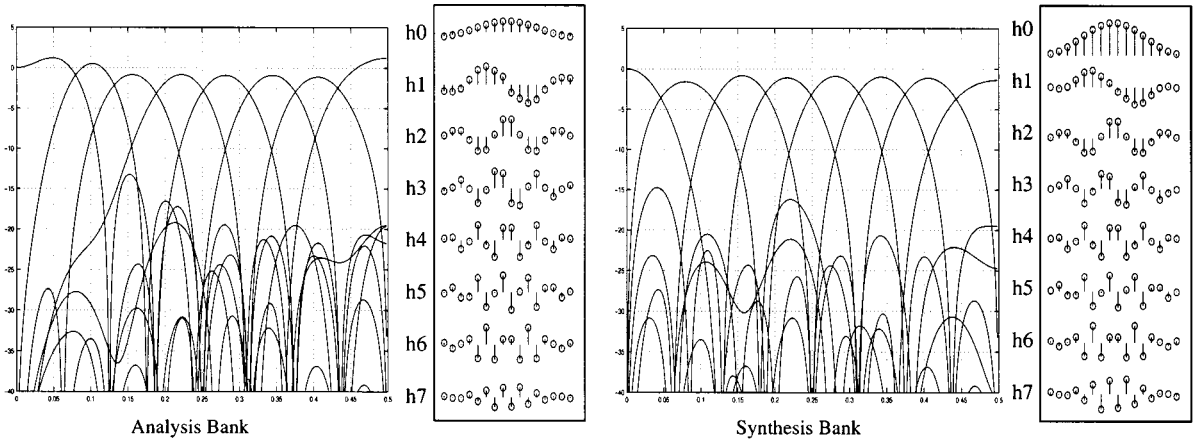

Fig. 5. Frequency and time responses of the 8×16 LiftLT. Left: analysis bank. Right: synthesis bank.

TABLE I
COMPARISON OF TRANSFORM COMPLEXITY: NUMBER OF OPERATIONS NEEDED PER EIGHT TRANSFORM COEFFICIENTS

Transform	*No. of Multiplications*	*No. of Additions*	*No. of Shifts*
8×8 DCT	13	29	0
8×16 Type-I Fast LOT	22	54	0
9/7 Wavelet, 3-level, lifting implementation	42	56	0
8×16 Fast LiftLT	14	51	6

32 fast LiftLT's improve the coding gain to 9.75 dB and 9.83 dB, respectively.

V. APPLICATION IN IMAGE CODING

To be fair, the same set partitioning in hierarchical trees (SPIHT) algorithm's quantizer and entropy coder [7] is utilized to encode the coefficients of every transform. The transforms in comparison are the 8×8 DCT, the new 8×16 LiftLT, and the popular 9/7-tap biorthogonal wavelet with a six-level decomposition. In the two block-transform cases, we use the modified zerotree structure in [8], where each block of transform coefficients are treated analogously to a full wavelet tree, and three more levels of decomposition are employed to decorrelate the DC subband further. The objective coding results (PSNR in dB) for standard 512×512 Lena, Goldhill, and Barbara test images are tabulated in Table II.

The LiftLT outperforms its block transform relatives for all test images at all bit rates. The visual quality of its reconstructed images is also superior, as demonstrated in Fig. 6. Blocking is completely avoided, whereas ringing is reasonably contained. Compared to the wavelet transform, the LiftLT is quite competitive on smooth images (about 0.2 dB below on Lena). However, for more complex images such as Goldhill or Barbara, the LiftLT consistently surpasses the 9/7-tap wavelet. The PSNR improvement can reach as high as 1.5 dB.

TABLE II
OBJECTIVE CODING RESULT COMPARISON (PSNR IN dB)

	Lena			*Goldhill*			*Barbara*		
Comp. Ratio	9/7 WL SPIHT	8×8 DCT	8×16 LiftLT	9/7 WL SPIHT	8×8 DCT	8×16 LiftLT	9/7 WL SPIHT	8×8 DCT	8×16 LiftLT
1:8	40.41	39.91	40.21	36.55	36.25	36.56	36.41	36.31	37.57
1:16	37.21	36.38	37.11	33.13	32.76	33.22	31.40	31.11	32.82
1:32	34.11	32.90	34.00	30.56	30.07	30.63	27.58	27.28	28.93
1:64	31.10	29.67	30.90	28.48	27.93	28.54	24.86	24.58	25.93
1:100	29.35	27.80	29.03	27.38	26.65	27.28	23.76	23.42	24.50
1:128	28.38	26.91	28.12	26.73	26.01	26.70	23.35	22.68	23.47

VI. CONCLUSIONS

We have presented in this paper the theory, design, and implementation of the LiftLT. The LiftLT is based on a fast, efficient, robust, and modular lattice structure. With only one more multiplication (which can also be implemented with shift-and-add operations), 22 more additions, and four more delay elements comparing to the DCT, our novel transform offers a fast, low-cost, VLSI-friendly implementation while

Fig. 6. Quarter portions of reconstructed Barbara images at 1 : 32 compression ratio. From left to right: 8 × 8 DCT, 27.28 dB; 8 × 16 LOT, 28.71 dB; 9/7-tap wavelet, 27.58 dB; and 8 × 16 LiftLT, 28.93 dB.

providing high quality reconstructed images, both objectively and subjectively. The LiftLT even surpasses the 9/7-tap biorthogonal wavelet with irrational coefficients.

REFERENCES

[1] W. B. Pennebaker and J. L. Mitchell, *JPEG: Still Image Compression Standard*. New York: V. N. Reinhold, 1993.

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

[3] T. D. Tran, R. de Queiroz, and T. Q. Nguyen, "The generalized lapped biorthogonal transform," in *Proc. ICASSP*, Seattle, WA, May 1998, pp. 1441–1444.

[4] ——, "Linear phase perfect reconstruction filter bank: Lattice structure, design, and application in image coding," *IEEE Trans. Signal Processing*, Jan. 2000.

[5] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *J. Fourier Anal. Applicat.*, vol. 4, pp. 247–269, 1998.

[6] H. S. Malvar, "Biorthogonal and nonuniform lapped transforms for transform coding with reduced blocking and ringing artifacts," *IEEE Trans. Signal Processing*, vol. 46, pp. 1043–1053, Apr. 1998.

[7] 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.

[8] T. D. Tran and T. Q. Nguyen, "A progressive transmission image coder using linear phase filter banks as block transforms," *IEEE Trans. Image Processing*, vol. 8, pp. 1493–1507, Nov. 1999.