Extended Block-Lifting-Based Lapped Transforms

Suzuki Taizo, Kudo Hiroyuki

IEEE signal processing letters

volume 22

number 10

page range 1657-1660

year 2015

(C) 2015, IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising/promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

URL http://hdl.handle.net/2241/00124676

doi: 10.1109/LSP.2015.2422837
Abstract—We extend an original lapped transform (LT) and use block-lifting factorization to get an extended block-lifting-based LT (XBL-LT). The block-lifting structure maps integer input signals to integer output signals and results in a reversible transform that reduces rounding errors by merging many rounding operations. Although other such block-lifting-based LTs (BL-LTs) have been proposed, they are forcibly constrained by the use of discrete cosine transform (DCT) matrices. In contrast, XBL-LT is DCT-unconstrained and hence also embodies the DCT-constrained form. Furthermore, it has fewer rounding operations by merging the scaling factor with block-lifting coefficients. The both DCT-constrained and unconstrained XBL-LTs perform well at lossy-to-lossless image coding which has scalability from lossless data to lossy data.

Index Terms—Block-lifting structure, lapped transform (LT), lossy-to-lossless image coding

I. INTRODUCTION

Lapped transforms (LTs) [1] are popular subband transforms that can be used as substitutes for the discrete cosine transform (DCT) [2] used in image/video coding (image coding). Although almost all of the JPEG and H.26x series [3–5], image coding standards, use DCTs for their good energy compaction, DCT-based image coding generates unpleasant artifacts, i.e., blocking artifacts, at low bit rates due to their ignoring the continuity of the blocks. LT-based image coding solves that problem by using a processing that works over the blocks.

The lifting structure [6] is a very important technology to achieve a lossless mode in subband transform-based image coding. It maps integer input signals to integer output signals; i.e., it is an integer-to-integer transform. The 4 × 8 lifting-based LT (L-LT) [7] in JPEG XR [8], the newest image coding standard, is a time-domain LT (TDLT) [9] with simple scaling factors and lifting structures. In spite of its simple structure, it performs well at lossy-to-lossless image coding, which has scalability from lossless to lossy data. The block-lifting structure [10] is a class of lifting structures and results in a reversible transform that reduces rounding errors by merging many rounding operations. Inspired by the L-LT in JPEG XR, we have proposed block-lifting-based LTs (BL-LTs) [11], [12] that perform well with larger block size than those of the L-LT in JPEG XR. However, they are forcibly DCT-constrained and degrade coding performance at high bit rates.

Here, we extend an original LT and use block-lifting factorization to get an extended BL-LT (XBL-LT). The XBL-LT is DCT-unconstrained, unlike the BL-LTs presented in [11], [12], and hence also embodies the DCT-constrained form. Furthermore, more rounding operations than the methods described in our previous studies are removed by merging the scaling factor with block-lifting coefficients. As a result, the both DCT-constrained and unconstrained XBL-LTs perform well at lossy-to-lossless image coding.

Notation: The italic letter $M$ ($M = 2^n$, $n \in \mathbb{N}$) denotes the block size. Boldface letters $I_m$, $J_m$, $0$, and $D_m$ denote an $m \times m$ identity matrix, an $m \times m$ reversal matrix, a null matrix, and an $m \times m$ diagonal matrix with alternating ±1 entries (i.e., diag{1, −1, 1, −1, · · · }). The superscripts $T$ and −1 respectively mean the transpose and inverse of a matrix.

II. REVIEW AND DEFINITION

A. Lapped Transform (LT)

In accordance with [11], [12], let $E(z)$ be a polyphase matrix of an $M \times 2M$ LT with a scaling factor $s$ derived from the L-LT in JPEG XR [7]:

$$E(z) = P \begin{bmatrix} I_N & 0 \\ 0 & S_N^T C_N^{III} \end{bmatrix} \Gamma(z) \begin{bmatrix} C_N^{II} & 0 \\ 0 & C_N^{IV} J_N \end{bmatrix} S W J M,$$  

(1)

where

$$\Gamma(z) = WA(z)W,$$

$$W = \frac{1}{\sqrt{2}} \begin{bmatrix} I_N & I_N \\ -I_N & I_N \end{bmatrix},$$

$$S = \text{diag}\{s I_N, s^{-1} I_N\}.$$

$C_M^{IV}$ and $S_M^{IV}$ are type-X DCT (DCT-X) and type-X discrete sine transform (DST-X) matrices, and the $(i, j)$-elements of $C_M^{II}$ and $C_M^{IV}$ are

$$[C_M^{II}]_{i,j} = \frac{2}{M} c_i \cos \left( \frac{i (j + 1/2) \pi}{M} \right),$$

$$[C_M^{IV}]_{i,j} = \frac{2}{M} \cos \left( \frac{i + 1/2}{j + 1/2} \pi \right).$$

where $c_k = 1/\sqrt{2}$ ($k = 0$) or 1 ($k \neq 0$), respectively. The following relationships between matrices can be established:

$$C_M^{III} = (C_M^{II})^{-1} = (C_M^{IV})^T, S_M^{IV} = D_M C_M^{IV} J_M.$$  

(2)

Here, $P$ is an $M \times M$ permutation matrix. The optimal scaling $s$ in $S$ is empirically determined, e.g., $s = 0.8981$ if $M = 8$ and 0.9360 if $M = 16$. Since the LT in Eq. (1) with $s = 1$ is completely equivalent to a lapped orthogonal transform (LOT), we will use the LT in Eq. (1) as a representative expression of LT.

T. Suzuki and H. Kudo are with the Faculty of Engineering, Information and Systems, University of Tsukuba, Tsukuba, Ibaraki, 305-8573 Japan (e-mail: {taizo, kudo}@cs.tsukuba.ac.jp).
where

\[ \mathbf{y} = \mathbf{D}_N \mathbf{y} \]

and

\[ \mathbf{z} = \mathbf{D}_N^{-1} \mathbf{z} \]

(1) to a DCT-unconstrained LT as follows:

\[ \mathbf{y}_i = \begin{bmatrix} 1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 1 \end{bmatrix} \mathbf{x}_i, \quad \mathbf{y}_j = \begin{bmatrix} 1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 1 \end{bmatrix} \mathbf{x}_j \]

\[ \mathbf{z}_i = \mathbf{y}_i + \text{round}(\mathbf{B}_0 \mathbf{x}_i), \quad \mathbf{z}_j = \mathbf{y}_j - \text{round}(\mathbf{B}_1 \mathbf{y}_j) \]

where \( \mathbf{x}_i, \mathbf{y}_j \), and \( \mathbf{z}_x \) are \( N \times 1 \) input/output vector signals, \( \text{round}(\cdot) \) is a rounding operation, and the block-lifting coefficients \( \mathbf{B}_0 \) and \( \mathbf{B}_1 \) are \( N \times N \) arbitrary matrices. In this case, the matrices and their inverses are expressed as lower and upper block-lifting matrices as follows:

\[ \mathbf{L}[\mathbf{B}_0] = \begin{bmatrix} \mathbf{I}_N & 0 \\ \mathbf{B}_0 & \mathbf{I}_N \end{bmatrix}, \quad \mathbf{L}[\mathbf{B}_0]^{-1} = \mathbf{L}[-\mathbf{B}_0] \]

\[ \mathbf{L}[\mathbf{B}_1] = \begin{bmatrix} \mathbf{I}_N & \mathbf{B}_1 \\ 0 & \mathbf{I}_N \end{bmatrix}, \quad \mathbf{L}[\mathbf{B}_1]^{-1} = \mathbf{L}[-\mathbf{B}_1]. \]

Rounding errors generated by the rounding operation in each lifting step degrade coding performance. The block-lifting structure reduces such rounding errors by merging many rounding operations. A special class of block-lifting structure is expressed as [13]

\[ \begin{bmatrix} \mathbf{M} & 0 \\ 0 & \mathbf{M}^{-1} \end{bmatrix} = \mathbf{L}[\mathbf{M}^{-1}] \mathbf{L}[-\mathbf{M}] \mathbf{L}[\mathbf{M}^{-1}] \mathbf{J}_M \]

(3)

\[ = \mathbf{J}_M \mathbf{L}[-\mathbf{M}] \mathbf{L}[\mathbf{M}^{-1}] \mathbf{L}[-\mathbf{M}], \]

(4)

where

\[ \mathbf{J}_M = \begin{bmatrix} 0 & \mathbf{I}_N \\ -\mathbf{I}_N & 0 \end{bmatrix}, \quad \mathbf{J}_M = \begin{bmatrix} 0 & -\mathbf{I}_N \\ \mathbf{I}_N & 0 \end{bmatrix}, \]

and \( \mathbf{M} \) is an \( N \times N \) arbitrary nonsingular matrix.

III. EXTENDED BLOCK-LIFTING-BASED LAPPED TRANSFORM (XBL-LT)

Theorem: We can extend the DCT-constrained LT in Eq. (1) to a DCT-unconstrained LT as follows:

\[ \mathbf{E} = \begin{bmatrix} \mathbf{I}_N & 0 \\ 0 & \mathbf{D}_N \mathbf{V} \end{bmatrix} \mathbf{J}_N \mathbf{J}_N^{-1} \mathbf{U} \mathbf{U}^{-1} \mathbf{J}_N \mathbf{J}_N^{-1} \mathbf{V} \mathbf{V}^{-1} \mathbf{J}_N \mathbf{J}_N^{-1} \]

(5)

where

\[ \mathbf{U} = \sqrt{2s} \mathbf{U}, \quad \mathbf{V} = \begin{bmatrix} \frac{1}{\sqrt{2s}} \mathbf{V} \\ \mathbf{W} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \mathbf{I}_N & \frac{1}{2} \mathbf{J}_N \\ -\mathbf{J}_N \end{bmatrix}, \]

by using Eq. (2) and skipping the sign inversion matrix \( \mathbf{D}_N \).

The scaling matrix \( \mathbf{S} \) is being embedded in \( \mathbf{U} \) and \( \mathbf{V} \). When \( \mathbf{U} = \mathbf{C}_N^{\mathbf{U}} \) and \( \mathbf{V} = \mathbf{C}_N^{\mathbf{V}} \), the LT in Eq. (5) is completely equivalent to the LT in Eq. (1) except that the signs are different. Then, we factorize the LT in Eq. (5) into block-lifting structures as

\[ \mathbf{E}(z) = \mathbf{P} \mathbf{L} \left[ \mathbf{B}_1 \right] \mathbf{U} \left[ \mathbf{B}_3 \right] \mathbf{L}(z) \mathbf{U} \left[ \mathbf{B}_2 \right] \mathbf{L} \left[ \mathbf{B}_1 \right] \mathbf{L} \left[ \mathbf{B}_0 \right] \mathbf{U} \left[ \mathbf{B}_4 \right], \]

(6)

wherein each matrix is defined as

\[ \mathbf{B}_0 = -\mathbf{N}^{-1}, \quad \mathbf{B}_1 = \mathbf{V}, \quad \mathbf{B}_2 = \mathbf{B}_0 + \mathbf{B}_4 \]

\[ \mathbf{B}_3 = -\frac{1}{2} \mathbf{U} \mathbf{J}_N \mathbf{V}, \quad \mathbf{B}_4 = \mathbf{V}^{-1} \mathbf{J}_N \mathbf{U}^{-1}. \]

Actually, the block-lifting matrices \( \mathbf{U} \left[ \mathbf{B}_3 \right] \) on both sides of the delay matrix \( \mathbf{L}(z) \) are collectively implemented, as shown in Fig. 2.

Proof: We can perform an easy matrix manipulation as follows:

\[ \begin{bmatrix} \mathbf{M}_0 & 0 \\ 0 & \mathbf{M}_1 \end{bmatrix} \begin{bmatrix} \mathbf{I}_N & \mathbf{N}_0 \\ \mathbf{N}_1 & \mathbf{I}_N \end{bmatrix} = \begin{bmatrix} \mathbf{I}_N & \mathbf{M}_0 \mathbf{N}_0 \mathbf{M}_1^{-1} \\ \mathbf{M}_1 \mathbf{N}_1 \mathbf{M}_0^{-1} & \mathbf{I}_N \end{bmatrix} \begin{bmatrix} \mathbf{M}_0 & 0 \\ 0 & \mathbf{M}_1 \end{bmatrix}, \]

where \( \mathbf{M}_x \) and \( \mathbf{N}_x \) are an \( N \times N \) arbitrary nonsingular matrix and \( N \times N \) arbitrary matrix, respectively. First, \( \mathbf{L}(z) \) in Eq. (5) can easily be represented by

\[ \mathbf{L}(z) = \mathbf{L} \left[ \mathbf{I}_N \right] \mathbf{U} \left[ -\frac{1}{2} \mathbf{I}_N \right] \mathbf{L}(z) \mathbf{U} \left[ -\frac{1}{2} \mathbf{I}_N \right] \mathbf{L} \left[ -\mathbf{I}_N \right]. \]

Next, \( \mathbf{U}^{-1} \) in Eq. (5) is moved to the right side of \( \mathbf{L}(z) \) and simplified as

\[ \mathbf{E}(z) = \mathbf{P} \mathbf{L} \left[ \mathbf{B}_4 \right] \mathbf{L}(z) \mathbf{U} \left[ -\frac{1}{2} \mathbf{I}_N \right] \mathbf{L}(z) \mathbf{U} \left[ -\frac{1}{2} \mathbf{I}_N \right] \mathbf{L} \left[ -\mathbf{I}_N \right], \]

(7)

because the block diagonal matrix \( \text{diag} \{ \mathbf{U}, \mathbf{U}^{-1} \} \) in \( \mathbf{E}(z) \) can be factorized into block-lifting structures as in Eq. (3). Then, \( \mathbf{V}^{-1} \mathbf{J}_N \) in Eq. (5) is moved to the right side of \( \mathbf{L}(z) \) and
simplified as

$$\Omega(z) \triangleq \begin{bmatrix} I_N & 0 \\ 0 & \Psi^{-1} J_N \end{bmatrix} \Psi(z) \begin{bmatrix} I_N & 0 \\ 0 & \Psi J_N \end{bmatrix} = \mathcal{L} \begin{bmatrix} W^{-1} \end{bmatrix} \begin{bmatrix} \frac{1}{2} \Psi \end{bmatrix} \Lambda(z) \begin{bmatrix} \frac{1}{2} \Psi \end{bmatrix} \mathcal{L} \begin{bmatrix} W \end{bmatrix} \begin{bmatrix} \frac{1}{2} \Lambda(z) \frac{1}{2} \Psi \end{bmatrix} \mathcal{L} \begin{bmatrix} W^{-1} \end{bmatrix} \begin{bmatrix} \frac{1}{2} \Psi \end{bmatrix} \begin{bmatrix} J_N & 0 \\ 0 & J_N \end{bmatrix},$$

where $$\mathcal{V} = U J_N \mathcal{V}$$, because the block diagonal matrix $$\text{diag}\{\Psi^{-1}, \mathcal{V}\}$$ in $$\Omega(z)$$ can also be factorized into block-lifting structures as in Eq. (4). Finally, the residual part $$\mathcal{W} J_M$$ in Eq. (5) and $$\text{diag}\{J_N, J_N\}$$ in $$\Omega(z)$$ are collectively factorized as

$$\begin{bmatrix} J_N & 0 \\ 0 & J_N \end{bmatrix} \mathcal{W} J_M = \mathcal{L} \begin{bmatrix} \frac{1}{2} J_N \end{bmatrix} \mathcal{L} \begin{bmatrix} J_N \end{bmatrix} \mathcal{J}_M.$$

IV. EXPERIMENTAL RESULTS

A. Coding Gain, Frequency Response, and Number of Rounding Operations

We designed $$8 \times 16$$ and $$16 \times 32$$ XBL-LTs by optimizing the coding gain (CG) [14]

$$\text{CG [dB]} = 10 \log_{10} \frac{\sigma^2}{\prod_{k=0}^{M-1} \sigma_{x_k}^2 \| f_k \|^2},$$

where $$\sigma^2$$ is the variance of the input signal, $$\sigma_{x_k}^2$$ is the variance of the $$k$$-th subbands and $$\| f_k \|^2$$ is the norm of the $$k$$-th synthesis filter. To simplify the design in DCT-unconstrained case, we set $$\mathcal{U}$$ and $$\mathcal{V}$$ in Eq. (6) as $$N \times N$$ arbitrary unitary matrices, where $$\mathcal{U}$$ is designed such that it has structural one-degree regularity [10] to achieve good image coding. Table I compares of the CGs[dB] of the XBL-LTs in the DCT-constrained and unconstrained cases. In addition, Fig. 3 shows the frequency responses of the analysis and synthesis parts of the XBL-LTs for the same cases as in Table I. In Table I and Fig. 3, the DCT-unconstrained case had slightly better results than the DCT-constrained case. Table II shows the numbers of rounding operations of BL-LTs. It is clear that the XBL-LTs have fewer rounding operations than those of conventional BL-LTs.

B. Lossy-to-Lossless Image Coding

For convenience regarding the number of pages, lossy-to-lossless image coding was implemented in only the $$M = 8$$ case. We used L-LTs in [7], [9], [11], [12] as the conventional L-LTs. The L-LT in [7] is the $$4 \times 8$$ L-LT for JPEG XR. The L-LT in [9] is the $$8 \times 16$$ TDLT with the pre-filtering part indicated by Fig. 5 in Table V in [9] and the DCT part indicated by [15]. The L-LTs in [11], [12] are the DCT-constrained BL-LTs. After the images were transformed by the L-LTs and periodic extension at the boundaries, the transformed coefficients were rearranged from the subband mode to the multiresolution mode similar to the wavelet transform. They were encoded with a common wavelet-based zerotree

\begin{table}[h]
\centering
\caption{Coding gains of XBL-LTs.}
\begin{tabular}{|c|c|c|}
\hline
Block Size & DCT-Constrained & DCT-Unconstrained \\
\hline
8 & 9.4475 & 9.4538 \\
16 & 9.8455 & 9.8621 \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\caption{Numbers of rounding operations of BL-LTs.}
\begin{tabular}{|c|c|c|c|}
\hline
\hline
8 & 48 & 28 & 24 \\
16 & 96 & 56 & 48 \\
\hline
\end{tabular}
\end{table}

Fig. 2. XBL-LT (black and white circles mean adders and rounding operations, respectively).

Fig. 3. Frequency responses of analysis and synthesis parts of XBL-LTs (dashed and solid lines indicate DCT-constrained and unconstrained cases, respectively): (top) $$8 \times 16$$ XBL-LT, (bottom) $$16 \times 32$$ XBL-LT.
By extending an original LT and using block-lifting factorization, we developed an XBL-LT with fewer rounding operations. It is DCT-unconstrained and hence can be DCT-constrained as well; i.e., it can be considered to be a more general structure than other BL-LTs. Although we constrained the design by using unitary matrices in this paper, the DCT-unconstrained structures have the potential to achieve better coding.

**ACKNOWLEDGMENT**

This work was supported by JSPS Grant-in-Aid for Young Scientists (B), Grant Number 25820152.

**REFERENCES**


