REALIZATION OF LOSSLESS-TO-LOSSY IMAGE CODING COMPATIBLE WITH JPEG STANDARD BY DIRECT-LIFTING OF DCT-IDCT

Taizo Suzuki and Masaaki Ikehara

Department of Electronics and Electrical Engineering, Keio University
Yokohama, Kanagawa 223-8522, Japan
Email: {suzuki, ikehara}@tkhm.elec.keio.ac.jp

ABSTRACT

A discrete cosine transform (DCT) can be easily implemented in software and hardware for the JPEG and MPEG formats. Recently, some integer DCTs (IntDCTs) for lossless-to-lossy image coding have been proposed, but they do not satisfy enough compatibility with JPEG standard. This paper proposes a realization of lossless-to-lossy image coding which has higher compatibility with it than the conventional IntDCTs. Our IntDCT is implemented by direct-lifting of DCT and inverse DCT (IDCT) and has high performance in lossless image coding compared with any IntDCT while keeping high compatibility with JPEG standard. Finally, our method is validated by its application to lossless-to-lossy image coding.

***Index Terms*—** integer discrete cosine transform (IntDCT), direct-lifting, lossless-to-lossy image coding, JPEG standard

1. INTRODUCTION

JPEG [1] is an international standard in image coding. Discrete cosine transform (DCT) [2] and differential pulse code modulation (DPCM) are separately applied to lossy and lossless image coding, respectively. DCT can be classified into types I to VIII and has excellent energy compaction capability, numerous fast implementations and numerous applications on JPEG devices. DCT type-II (DCT-II) and -III (DCT-III) are commonly called DCT and inverse DCT (IDCT). Since DCT and IDCT are compatible with the lossy JPEG, they are especially promising.

Recently, with the rapid development of hardware such as PCs and mobile phones, and the continual expansion of broadband, lossless-to-lossy image coding, which is the unification of lossy and lossless image coding, is demanded to obtain higher quality and compression ratio. Such lossless-to-lossy image coding has already been achieved by JPEG2000 [3] and HD Photo (JPEG-XR) [4]. These next-generation standards are based on discrete wavelet transform (DWT) and the hierarchical lapped biorthogonal transform (HLBT), respectively. However, the existing state-of-the-art technologies such as JPEG2000 and JPEG-XR cannot be substituted for the JPEG standard because JPEG is most frequently used and its compressed data is spread far over the world. This means that the transform for lossless-to-lossy image coding should have compatibility with JPEG. Additionally, an image encoded by such transform can also be decoded by the existing JPEG decoder if lossless data is not demanded.

This work is supported in part by a Grant-in-Aid for the Global Center of Excellence for High-Level Global Cooperation for Leading-Edge Platform on Access Spaces from the Ministry of Education, Culture, Sport, Science, and Technology in Japan.

Up to now, many integer DCTs (IntDCTs) [5–8], which is constructed by integer coefficients, lifting structures [9] and rounding operations, have been proposed for lossless-to-lossy image coding. In this regard, however, high compatibility with JPEG standard is desired, but the conventional IntDCTs do not satisfy enough compatibility with it.

This paper proposes a simple lifting structure, which is called direct-lifting, using only DCT and IDCT matrices for every lifting block. Also, two-dimensional block transform (2DBT) applied to every lifting block achieves better coding performance because the total number of rounding operations are reduced. As a result, our IntDCT has high performance in lossless image coding compared with any IntDCT while keeping high compatibility with JPEG standard. Although our method requires a small side information block (SIB), it is validated by its application to lossless-to-lossy image coding.

Notations: $\mathbf{I}$, $\mathbf{M}^T$ and $diag\{.\}$ are the identity matrix, transpose of a matrix $\mathbf{M}$ and block diagonal matrix, respectively. For simplicity, we omit matrix sizes when they are obvious.

2. REVIEW

2.1. Block-Lifting Structure

Lifting, also known as a ladder structure, is an implementation method of wavelet transforms originally proposed by Sweldens [9]. It is a special type of lattice structure: a cascading construction using only elementary matrices, that is, identity matrices with a single nonzero off-diagonal element. Also, the block-lifting structure [10], shown in Fig. 1, is known as an efficient lifting structure for lossless-to-lossy coding. It achieves a higher compression ratio because it can merge many rounding operations. The left part in Fig. 1 shows a block-lifting structure before rounding operations are merged, and the right part in Fig. 1 shows the structure after rounding operations are merged. It is clear that the number of rounding operations is reduced from N^2 to N when the size of the lifting block $\mathbf{T}$ is $N \times N$. In the right part in Fig. 1, the analysis input signal vectors $\mathbf{x}_i$ and $\mathbf{x}_j$, the analysis output and synthesis input signal vectors $\mathbf{y}_i$ and $\mathbf{y}_j$, the synthesis output signal vectors $\mathbf{z}_i$ and $\mathbf{z}_j$, and lifting block $\mathbf{T}$ are presented by

$$\begin{aligned}
&\mathbf{y}_i = \mathbf{x}_i + round[\mathbf{T}\mathbf{x}_j],\\
&\mathbf{y}_j = \mathbf{x}_j,\\
&\mathbf{z}_i = \mathbf{y}_i - round[\mathbf{T}\mathbf{y}_j] = \mathbf{x}_i, \quad \text{and}\\
&\mathbf{z}_j = \mathbf{y}_j = \mathbf{x}_j.
\end{aligned}$$

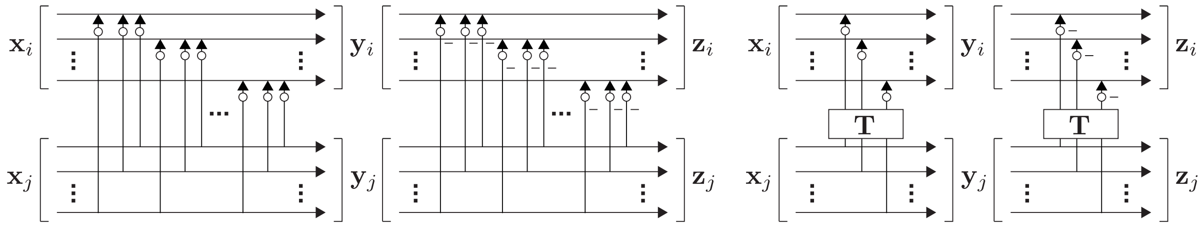

Fig. 1. Block-lifting structure (white circles: rounding operations): (left) before rounding operations are merged, (right) after they are merged.

In this case, the block-lifting matrix and its inverse matrix are as follows:

$$\begin{bmatrix}\mathbf{I} & \mathbf{T}\\ \mathbf{0} & \mathbf{I}\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}\mathbf{I} & \mathbf{T}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}^{-1} = \begin{bmatrix}\mathbf{I} & -\mathbf{T}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}.$$

2.2. Discrete Cosine Transform (DCT) and Integer Discrete Cosine Transform (IntDCT)

This paper describes only types DCT-II and -III which are commonly called DCT and IDCT [2], respectively. These are often used for image and video coding such as JPEG and MPEG. The m-column and n-row elements of M-channel DCT and IDCT matrices $\mathbf{C}$ and $\mathbf{D}$ are defined as

$$[\mathbf{C}]_{m,n} = \sqrt{\frac{2}{M}} c_m \cos\left(\frac{m\,(n+1/2)\,\pi}{M}\right)$$
$$[\mathbf{D}]_{m,n} = \sqrt{\frac{2}{M}} c_n \cos\left(\frac{(m+1/2)\,n\pi}{M}\right)$$

where $\mathbf{D} = \mathbf{C}^{-1} = \mathbf{C}^T$, $0 \le m, n \le M-1$,

$$c_m = \begin{cases} \frac{1}{\sqrt{2}} & (m=0) \\ 1 & (m \neq 0) \end{cases} \quad \text{and} \quad c_n = \begin{cases} \frac{1}{\sqrt{2}} & (n=0) \\ 1 & (n \neq 0) \end{cases},$$

respectively. For simplicity, let us define $M = 2^n$ ($n \in \mathbb{N}$). Also, many fast implementations of DCT and IDCT have been widely researched. Note that the coding system by DCT is limited to operation in only lossy image coding because distortion of the decoded image is unavoidable with these lossy algorithms.

On the other hand, many IntDCTs have been proposed for lossless-to-lossy image coding [5–8]. These IntDCTs are constructed by lifting structures [9] with rounding operations. Because fast implementation of DCT is achieved by cascading Givens rotation matrices

$$\mathbf{R}_\theta = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix},$$

most IntDCTs are redesigned by lifting factorization as

$$\mathbf{R}_\theta = \begin{bmatrix}1 & \frac{\cos\theta-1}{\sin\theta}\\ 0 & 1\end{bmatrix} \begin{bmatrix}1 & 0\\ \sin\theta & 1\end{bmatrix} \begin{bmatrix}1 & \frac{\cos\theta-1}{\sin\theta}\\ 0 & 1\end{bmatrix}.$$

Hence, note that most IntDCTs are constructed by many rounding operations, so that rounding error is generated much more.

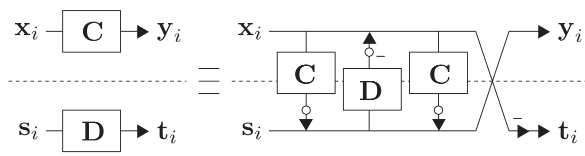

Fig. 2. Direct-lifting of DCT-IDCT (white circles: rounding operations).

3. EFFICIENT REALIZATION OF INTDCT

3.1. Direct-Lifting of DCT-IDCT

We process the two individual signals by DCT and IDCT, as shown on the left part in Fig. 2. The output signals are transformed by

$$\begin{bmatrix}\mathbf{y}_i\\ \mathbf{t}_i\end{bmatrix} = \begin{bmatrix}\mathbf{C} & \mathbf{0}\\ \mathbf{0} & \mathbf{D}\end{bmatrix} \begin{bmatrix}\mathbf{x}_i\\ \mathbf{s}_i\end{bmatrix}. \tag{1}$$

Here, $diag\{\mathbf{C}, \mathbf{D}\}$ in (1) can be factorized into complete block-lifting structures such as

$$\begin{bmatrix}\mathbf{C} & \mathbf{0}\\ \mathbf{0} & \mathbf{D}\end{bmatrix} = \begin{bmatrix}\mathbf{0} & \mathbf{I}\\ -\mathbf{I} & \mathbf{0}\end{bmatrix} \begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{C} & \mathbf{I}\end{bmatrix} \begin{bmatrix}\mathbf{I} & -\mathbf{D}\\ \mathbf{0} & \mathbf{I}\end{bmatrix} \begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{C} & \mathbf{I}\end{bmatrix}. \tag{2}$$

Also, its inverse transform is expressed by

$$\begin{bmatrix}\mathbf{C} & \mathbf{0}\\ \mathbf{0} & \mathbf{D}\end{bmatrix}^{-1} = \begin{bmatrix}\mathbf{D} & \mathbf{0}\\ \mathbf{0} & \mathbf{C}\end{bmatrix} = \begin{bmatrix}\mathbf{I} & \mathbf{0}\\ -\mathbf{C} & \mathbf{I}\end{bmatrix} \begin{bmatrix}\mathbf{I} & \mathbf{D}\\ \mathbf{0} & \mathbf{I}\end{bmatrix} \begin{bmatrix}\mathbf{I} & \mathbf{0}\\ -\mathbf{C} & \mathbf{I}\end{bmatrix} \begin{bmatrix}\mathbf{0} & -\mathbf{I}\\ \mathbf{I} & \mathbf{0}\end{bmatrix}. \tag{3}$$

Thus, the block parallel system of DCT-IDCT can be efficiently implemented by the block-lifting structure, as shown in the right part in Fig. 2. However, the systems in (2) and (3) have the problem that the transformed signals by IDCT are not suitable for image compression because IDCT includes a lot of DC leakages.

3.2. Efficient Realization by Side Information Block (SIB)

We can easily come up with a simple realization for lossless-to-lossy image coding which is applied after an $N \times N$ image $\mathbf{X}$ is segmented to $M \times M$ block $\mathbf{x}_k$ ($0 \le k \le n-1$, $n = (N/M)^2$). The segmented block $\mathbf{x}_k$ is sequentially transformed by DCT as

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k \quad \text{for} \quad k = 1, 2, \cdots, n$$

where $\mathbf{y}_k$ is an output signal of $\mathbf{x}_k$. In parallel, $M \times M$ block $\mathbf{s}_k$ has to be transformed by IDCT, however, the transform by IDCT is undesirable for the image compression as described in the previous

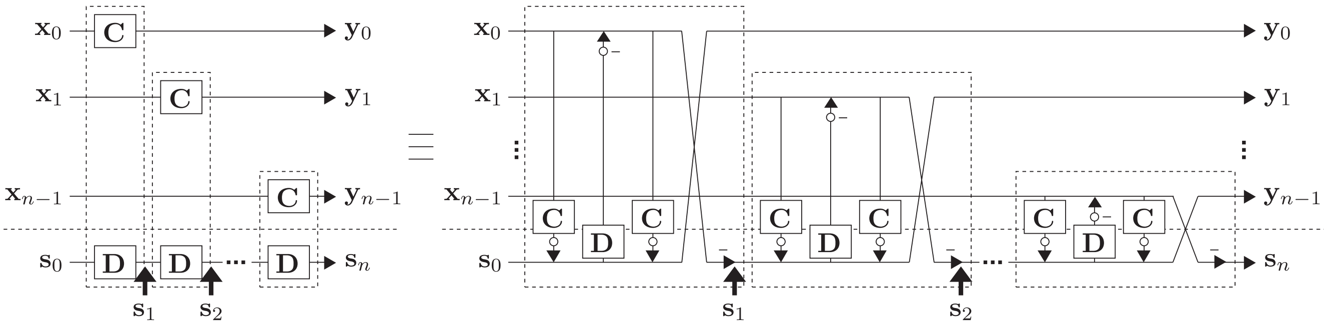

Fig. 3. IntDCT by direct-lifting of DCT-IDCT (white circles: rounding operations): (left) cascading of block parallel systems of DCT-IDCT, (right) lifting factorizations of it.

subsection. Then we prepare an $M \times M$ side information block (SIB) $\mathbf{s}_0$ as a null matrix and $\mathbf{s}_k$ is iteratively transformed from $\mathbf{s}_0$ by

$$\mathbf{s}_k = \mathbf{D}\mathbf{s}_{k-1} \quad \text{for} \quad k = 1, 2, \cdots, n.$$

As a result, the proposed realization for lossless-to-lossy image coding is presented by

$$\left[\begin{array}{c} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{n-1} \\ \hline \mathbf{s}_n \end{array}\right] = \left[\begin{array}{cccc|c} \mathbf{C} & & & & \\ & \mathbf{C} & & & \mathbf{0} \\ & & \ddots & & \\ & & & \mathbf{C} & \\ \hline & & \mathbf{0} & & \mathbf{D}^n \end{array}\right] \left[\begin{array}{c} \mathbf{x}_0 \\ \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{n-1} \\ \hline \mathbf{s}_0 \end{array}\right] \tag{4}$$

and $\mathbf{s}_n$ is also encoded with all of $\mathbf{y}_k$. The left part in Fig. 3 shows the system in (4). Note $\mathbf{s}_n \neq \mathbf{0}$ due to rounding error in each lifting structure. Moreover, it is clear that the combination of $\mathbf{C}$ and $\mathbf{D}$ can be factorized to lifting structures by (2), as shown in dashed line area in Fig. 3. As expected, the inverse transform of (4) is expressed by

$$\left[\begin{array}{c} \mathbf{x}_0 \\ \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{n-1} \\ \hline \mathbf{s}_0 \end{array}\right] = \left[\begin{array}{cccc|c} \mathbf{D} & & & & \\ & \mathbf{D} & & & \mathbf{0} \\ & & \ddots & & \\ & & & \mathbf{D} & \\ \hline & & \mathbf{0} & & \mathbf{C}^n \end{array}\right] \left[\begin{array}{c} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{n-1} \\ \hline \mathbf{s}_n \end{array}\right]$$

and the combination of $\mathbf{D}$ and $\mathbf{C}$ can be factorized to lifting structures by (3).

The proposed lossless-to-lossy image coding is summarized as follows:

1. Lossless bitstream with block $\mathbf{s}_n$ can be obtained by using (4) and an encoder.
2. In lossless mode, the image is reconstructed from all lossless bitstreams and $\mathbf{s}_n$. Each block $\mathbf{s}_k$ is successively inversely transformed by $\mathbf{s}_k = \mathbf{C}\mathbf{s}_{k+1}$ without any loss.
3. In lossy mode, the image is reconstructed by interrupting the lossless bitstream without using $\mathbf{s}_n$ and decoding it by JPEG standard.

Table 1. The comparisons of the number of rounding operations per an 1D transform of 8×1 signals.

K's [5]	H's [6]	L's [7]	C's [8]	Prop. IntDCT
29	9	15	8	**1.5**

3.3. More Efficient Improvement by Two-Dimensional Block Transform (2DBT)

DCT applied to $M \times M$ input signal $\mathbf{x}$ separately by column and row is expressed by

$$\mathbf{y} = \left(\mathbf{C}\left(\mathbf{C}\mathbf{x}\right)^T\right)^T = \mathbf{C}\mathbf{x}\mathbf{C}^T \triangleq \mathbf{C}_{2D}(\mathbf{x})$$

where $\mathbf{y}$ is an output signal of $\mathbf{x}$. Similar to $\mathbf{C}_{2D}(\mathbf{x})$, let us define $\mathbf{D}_{2D}(\mathbf{x}) \triangleq \mathbf{D}\mathbf{x}\mathbf{D}^T$. Then, we find that each lifting block is also defined as an $M \times M$ two-dimensional block transform (2DBT). Consequently, (2) is represented by

$$\begin{bmatrix} \mathbf{C}_{2D}(\mathbf{x}) & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{2D}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{C}_{2D}(\mathbf{x}) & \mathbf{I} \end{bmatrix} \times \begin{bmatrix} \mathbf{I} & -\mathbf{D}_{2D}(\mathbf{x}) \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{C}_{2D}(\mathbf{x}) & \mathbf{I} \end{bmatrix}. \tag{5}$$

Such lifting matrices by 2DBT in (5) are a more efficient for lossless-to-lossy image coding because the number of rounding operations per an one-dimensional (1D) transform of $M \times 1$ signals is reduced from $3M$ to $3M/16$ and rounding error is also reduced much more. The comparisons of the number of rounding operations when the block size is 8 are shown in Table 1.

4. RESULTS

In this section, the proposed 8-channel IntDCT is validated in lossless-to-lossy image coding by the lossless bitrate (LBR) [bpp] in lossless image coding and peak signal-to-noise ratio (PSNR) [dB] in lossy image coding. We chose [5–8] as the conventional IntDCTs and twenty 512×512 grayscale test images, for example, "Barbara", "Lena" and "Pepper". The set partitioning in hierarchical trees (SPIHT) progressive image transmission algorithm [11] was used to encode the transformed images.

Table 2. Lossless image coding results (LBR [bpp]).

Test image	DCT [2]	K's [5]	H's [6]	L's [7]	C's [8]	Prop. IntDCT
Airplane	-	4.51	4.41	4.90	4.40	**4.36**
Barbara	-	5.02	4.96	5.51	4.97	**4.95**
Boat	-	5.22	5.19	5.73	5.19	**5.18**
Goldhill	-	5.19	5.16	5.71	5.16	**5.15**
Lena	-	4.69	4.63	5.22	4.64	**4.62**
Pepper	-	4.98	4.95	5.52	4.96	**4.94**
Watch	-	4.35	4.15	4.55	4.18	**4.06**
Avg.	-	5.44	5.39	5.91	5.40	**5.38**

Table 3. Lossy image coding results (PSNR [dB]).

Test image	DCT [2]	K's [5]	H's [6]	L's [7]	C's [8]	Prop. IntDCT
bitrate: 0.25 [bpp]						
Barbara	**26.95**	26.93	26.94	23.70	26.73	**26.95**
Lena	**31.88**	31.83	31.84	29.03	31.85	31.86
Pepper	**31.43**	31.38	31.41	28.49	31.42	31.42
bitrate: 0.50 [bpp]						
Barbara	**30.70**	30.65	30.67	27.14	30.40	30.68
Lena	**35.60**	35.45	35.51	32.25	35.48	35.55
Pepper	**34.49**	34.38	34.43	31.52	34.33	34.46
bitrate: 1.00 [bpp]						
Barbara	**36.08**	35.88	35.96	31.18	35.84	36.03
Lena	**39.23**	38.70	38.97	35.52	38.93	39.12
Pepper	**37.04**	36.68	36.85	34.16	36.71	36.97

4.1. Application to Lossless Image Coding

First, the proposed IntDCTs are applied to lossless image coding. SIB $\mathbf{s}_n$ is encoded without compression because the data consist of very few bits:

$$\text{SIB bitrate [bpp]} = \frac{\lceil \log_2 (|\mathbf{s}_n|_{max}) + 1 \rceil M^2 [\text{bit}]}{\text{Total number of pixels [pixel]}}$$

where $\lceil . \rceil$ is the ceiling of (.). For example, the size is only 512 [bits] (0.01953125 [bpp]) in the 512×512 image "Barbara" transformed by the proposed IntDCT. The comparisons of

$$\text{LBR [bpp]} = \frac{\text{Total number of bits [bit]}}{\text{Total number of pixels [pixel]}} + \text{SIB bitrate}$$

in lossless image coding are shown in Table 2. The coding performance of the proposed IntDCTs is better than the performance of all of the conventional methods.

4.2. Application to Lossy Image Coding

If lossy compressed data is required, it can be achieved by interrupting the obtained lossless bitstream without SIB $\mathbf{s}_n$. Therefore, we can use the existing JPEG decoder to reconstruct an image. The comparisons of

$$\text{PSNR [dB]} = 10 \log_{10} \left(\frac{255^2}{\text{MSE}} \right),$$

where MSE is the mean squared error, with other lossy image coding methods are shown in Table 3. As lossless image coding, in Table 3, we see that the proposed IntDCT shows better coding performance than do the conventional methods and has performance comparable to the JPEG coder even in lossy image coding with low bitrates.

5. CONCLUSION

This paper presents a realization of lossless-to-lossy image coding which has higher compatibility with JPEG standard than the conventional integer discrete cosine transforms (IntDCTs). Our IntDCT is implemented by direct-lifting of DCT and inverse DCT (IDCT) and two-dimensional block transform (2DBT) is applied to every lifting block to reduce the total number of rounding operations and shows better coding performance. On the other hand, although a side information block (SIB) is required, its size is the same as that of DCT and the amount of information for coding is small. As a result, our IntDCT has high performance in lossless image coding compared with any IntDCT while keeping high compatibility with JPEG standard. Our IntDCT also shows better coding performance than the conventional methods in lossless-to-lossy image coding. This method can be applied into not only DCT but any orthogonal transform. Our future work will be deleting the side information and reducing the computational cost.

6. REFERENCES

[1] ISO/IEC 10918-1, *Information technology - Digital compression and coding of continuous-tone still images Requirements and Guidelines*.

[2] K. R. Rao and P. Yip, *Discrete Cosine Transform Algorithms*, Academic Press, 1990.

[3] ISO/IEC 15444-1, *Information technology - JPEG 2000 image coding system: Core coding system*.

[4] ISO/IEC 29199-2, *JPEG XR image coding system - Part 2: Image coding specification*.

[5] K. Komatsu and K. Sezaki, "Reversible discrete cosine transform," in *Proc. of ICASSP'98*, Seattle, WA, May 1998.

[6] P. Hao and Q. Shi, "Matrix factorizations for reversible integer mapping," *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2314–2324, Oct. 2001.

[7] J. Liang and T. D. Tran, "Fast multiplierless approximations of the DCT with the lifting scheme," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3032–3044, Dec. 2001.

[8] S. Chokchaitam, M. Iwahashi, and S. Jitapunkul, "A new unified lossless/lossy image compression based on a new Integer DCT," *IEICE Trans. Inf. Syst.*, vol. E88-D, no. 2, pp. 403–413, Feb. 2005.

[9] W. Sweldens, "The lifting scheme: A new philosophy in biorthogonal wavelet constructions," in *Proc. of SPIE 2569*, 1995.

[10] S. Iwamura, Y. Tanaka, and M. Ikehara, "An efficient lifting structure of biorthogonal filter banks for lossless image coding," in *Proc. of ICIP'07*, San Antonio, TX, Sep. 2007.

[11] 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, no. 3, pp. 243–250, June 1996.