# Realization of IntDCT with Arbitrary Block Size Using Relation between DCT-IV and Parallel Block System of DCT-II

Taizo Suzuki, Hideaki Hayano and Masaaki Ikehara

Department of Electronics and Electrical Engineering, Keio University
3-14-1 Hiyoshi, Kohoku, Yokohama, Kanagawa 223-8522, Japan
Phone: +81-45-566-1530, E-mail: {suzuki, hayano, ikehara}@tkhm.elec.keio.ac.jp

**Abstract**

Many integer discrete cosine transforms (IntDCTs), which are constructed by cascading lifting structures, have been proposed for lossless-to-lossy image coding. However, the conventional methods are mostly only suitable for a block size of 8. This size may not be the best one for various applications including image coding. This paper proposes a novel IntDCT using the relation between the discrete cosine transform type IV (DCT-IV) and the parallel block system of the discrete cosine transform type II (DCT-II). Our IntDCT can be used with an arbitrary block size $M$ ($M = 2^n$, $n \in \mathbb{N}$), which results in better coding performance than a block size of 8 in lossless-to-lossy image coding.

## 1. Introduction

In JPEG [1], which is an international standard for image coding, the discrete cosine transform (DCT) [2] and differential pulse code modulation (DPCM) are separately applied to lossy and lossless image coding, respectively. DCT has several types (type I to -VIII), excellent energy compaction capability and numerous fast implementations, and myriad devices have been developed using it. DCT type II (DCT-II) and type III (DCT-III) are commonly called DCT and inverse DCT (IDCT), respectively. Consequently, DCT is especially promising since it is compatible with lossy JPEG.

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 required to obtain higher quality and a higher compression ratio. Such lossless-to-lossy image coding has already been achieved by JPEG2000 [3] and HD Photo (JPEG-XR) [4]. However, existing state-of-the-art technology such as JPEG2000 and JPEG-XR cannot be used to replace JPEG because JPEG is mostly used and its compressed data is extremely widely used worldwide. This means that a transform for lossless-to-lossy image coding must be compatible with JPEG. Additionally, an image encoded by such a transform must also be decodable even by the existing JPEG decoder if lossless data is not required.

This work was supported in part by a Grant-in-Aid for the Japan Society for the Promotion of Science (JSPS) Fellows and a Grant-in-Aid for the Global Center of Excellence (G-COE) 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), which are constructed by cascading lifting structures [5], have been proposed for lossless-to-lossy image coding [6, 7]. However, the conventional methods are mostly only suitable for a block size of 8. This size may not be the best one for various applications including image coding.

This paper proposes a novel IntDCT using the relation between DCT-IV and the parallel block system of DCT-II. Our IntDCT can be used with an arbitrary block size $M$ ($M = 2^n$, $n \in \mathbb{N}$), which results in better coding performance than a block size of 8 in lossless-to-lossy image coding.

*Notation*: $\mathbf{I}$, $\mathbf{J}$, $\mathbf{D}$, $diag\{.\}$, $\mathbf{M}^T$ and $\mathbf{M}^{[M]}$ are the identity matrix, an invertible matrix, a diagonal matrix, a block diagonal matrix, the transpose of a matrix and an $M \times M$ matrix, respectively. In this paper, $\left[\mathbf{D}^{[\mathbf{M}]}\right]_{ii} = (-1)^i$ $(0 \le i \le M-1)$.

## 2. Review

### 2.1. Lifting structure

The lifting structure [5] for various applications has been researched widely. Since the structure achieves an integer-to-integer transform, lossless-to-lossy image coding can be implemented. In Figure 1, which shows lifting structures, the analysis input signals $x_i$ and $x_j$, the analysis output and synthesis input signals $y_i$ and $y_j$, the synthesis output signals $z_i$ and $z_j$ and the lifting coefficients $T_0$ and $T_1$ are represented by

$$
\begin{aligned}
y_j &= x_j + round[T_0 x_i] \\
y_i &= x_i + round[T_1 y_j] \\
z_i &= y_i - round[T_1 y_j] = x_i \\
z_j &= y_j - round[T_0 z_i] = x_j
\end{aligned}
$$

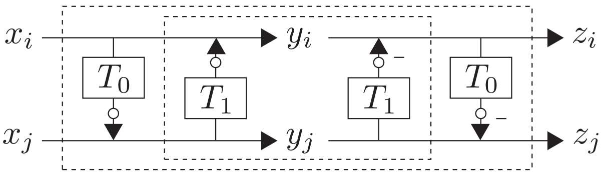


Figure 1: Lifting structures (white circles: rounding operations)

in the analysis and synthesis parts, respectively, where $round[.]$ is the rounding of (.).

In this case, the lifting matrices and their inverse matrices are as follows:

$$\begin{bmatrix} 1 & 0 \\ T_0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ T_0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -T_0 & 1 \end{bmatrix},$$
$$\begin{bmatrix} 1 & T_1 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & T_1 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & -T_1 \\ 0 & 1 \end{bmatrix}$$

### 2.2. DCT

DCT-II, -III and -IV are often used for multimedia coding as image, video and audio coding. The $m$-column and $n$-row element of $M$-channel DCT-II, -III and -IV matrices $\mathbf{C}_{II}^{[M]}$, $\mathbf{C}_{III}^{[M]}$ and $\mathbf{C}_{IV}^{[M]}$ are defined as

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

where $0 \leq m, n \leq M-1$ and $c_i = 1/\sqrt{2}$ ($i = 0$) or 1 ($i \neq 0$). For simplicity, let us define that $M = 2^n$ ($n \in \mathbb{N}$).

## 3. IntDCT using Relation between DCT-IV and Parallel Block System of DCT-II

### 3.1. Relation between DCT-IV and parallel block system of DCT-II

We consider how to process two individual signals by DCT-II as shown on the left side of Figure 2. This system is called the parallel block system of DCT-II and is expressed by

$$\begin{bmatrix} \mathbf{y}_i \\ \mathbf{y}_j \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{II}^{[M]} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{II}^{[M]} \end{bmatrix} \begin{bmatrix} \mathbf{x}_i \\ \mathbf{x}_j \end{bmatrix} \tag{1}$$

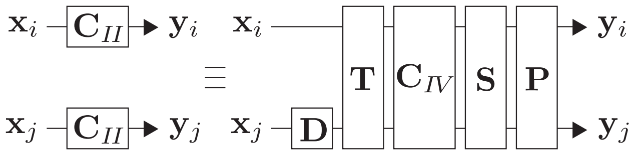


Figure 2: Relation between DCT-IV and parallel block system of DCT-II

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are $M \times M$ input signals, and $\mathbf{y}_i$ and $\mathbf{y}_j$ are $M \times M$ output signals, respectively.

On the other hand, Wang has found that DCT-IV can be factorized into DCT-III and Givens rotation matrices for fast implementation as [8]

$$\mathbf{C}_{IV}^{[2M]} = \mathbf{T}^{[2M]} \begin{bmatrix} \mathbf{C}_{III}^{[M]} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{[M]}\mathbf{C}_{III}^{[M]} \end{bmatrix} \mathbf{P}^{[2M]}\mathbf{S}^{[2M]} \tag{2}$$

where $0 \leq i \leq 2M-1$,

$$\mathbf{T}^{[2M]} = \begin{bmatrix} \mathbf{T}_{cos}^{[M]} & \mathbf{T}_{sin}^{[M]}\mathbf{J} \\ \mathbf{J}\mathbf{T}_{sin}^{[M]} & -\mathbf{J}\mathbf{T}_{cos}^{[M]}\mathbf{J} \end{bmatrix}$$
$$\text{where} \begin{cases} \left[\mathbf{T}_{cos}^{[M]}\right]_{ii} = \cos\left(\frac{(2i+1)\pi}{8M}\right) \\ \left[\mathbf{T}_{sin}^{[M]}\right]_{ii} = \sin\left(\frac{(2i+1)\pi}{8M}\right) \end{cases}$$

$$\mathbf{P}^{[2M]} = \begin{bmatrix} \mathbf{P}_{top}^{[\frac{N}{2}\times N]} \\ \mathbf{P}_{bottom}^{[\frac{N}{2}\times N]} \end{bmatrix} \mathbf{P}_{pm}^{[2M]}$$
$$\text{where} \begin{cases} \left[\mathbf{P}_{top}^{[\frac{N}{2}\times N]}\right]_{ij} = \begin{cases} 1 & (j = 2i) \\ 0 & (\text{otherwise}) \end{cases} \\ \left[\mathbf{P}_{bottom}^{[\frac{N}{2}\times N]}\right]_{ij} = \begin{cases} 1 & (j = N-2i-1) \\ 0 & (\text{otherwise}) \end{cases} \\ \left[\mathbf{P}_{pm}^{[2M]}\right]_{ii} = \begin{cases} -1 & (i = 4k-1,\ k \in \mathbb{N}) \\ 0 & (\text{otherwise}) \end{cases} \end{cases}$$

and

$$\mathbf{S}^{[2M]} = \begin{bmatrix} 1 & & & & \\ & \mathbf{R}_{\frac{\pi}{4}} & & \mathbf{0} & \\ & & \ddots & & \\ & \mathbf{0} & & \mathbf{R}_{\frac{\pi}{4}} & \\ & & & & 1 \end{bmatrix}$$
$$\text{where } \mathbf{R}_{\frac{\pi}{4}} = \begin{bmatrix} \cos(\frac{\pi}{4}) & -\sin(\frac{\pi}{4}) \\ \sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{bmatrix}$$

Here, Equation (2) can be rewritten as

$$\begin{bmatrix} \mathbf{C}_{II}^{[M]} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{II}^{[M]} \end{bmatrix} = \mathbf{P}^{[2M]}\mathbf{S}^{[2M]}\mathbf{C}_{IV}^{[2M]}\mathbf{T}^{[2M]} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{[M]} \end{bmatrix} \tag{3}$$

where $\mathbf{C}_{III}^{T} = \mathbf{C}_{II}$ and $\mathbf{C}_{IV}^{T} = \mathbf{C}_{IV}$. Equation (3) is the relation between DCT-IV and the parallel block system of DCT-II shown in Figure 2.

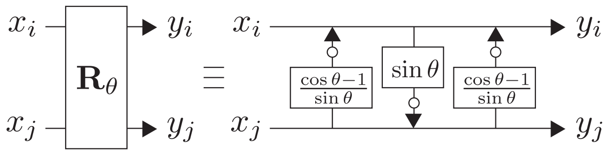


Figure 3: Lifting factorization of Givens rotation matrix (white circles: rounding operations)

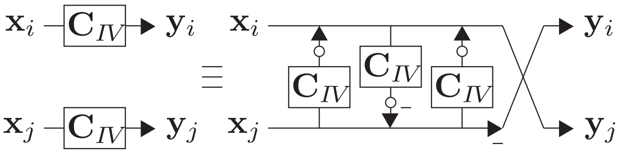


Figure 4: Lifting factorization of DCT-IV (white circles: rounding operations)

### 3.2. Lifting factorization of $M$-channel DCT-IV and rotation matrices

To achieve lossless-to-lossy image coding, $\mathbf{S}^{[2M]}$, $\mathbf{T}^{[2M]}$ and $\mathbf{C}_{IV}^{[2M]}$ in Equation (3) must be factorized into lifting structures. First, since $\mathbf{S}^{[2M]}$ and $\mathbf{T}^{[2M]}$ can be regarded as cascading Givens rotation matrices,

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

is adopted for the lifting factorization of $\mathbf{S}^{[2M]}$ and $\mathbf{T}^{[2M]}$. Figure 3 shows the lifting factorization of a Givens rotation matrix. Next, we consider how to process two individual signals by DCT-IV similar to that by DCT-II in Equation (1) as $diag\{\mathbf{C}_{IV}^{[M]}, \mathbf{C}_{IV}^{[M]}\}$. It is known that the parallel block system of DCT-IV can be factorized into lifting structures with only DCT-IV in each coefficient [9] as follows:

$$\begin{bmatrix}\mathbf{C}_{IV}^{[M]} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{IV}^{[M]}\end{bmatrix} = \begin{bmatrix}\mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{C}_{IV}^{[M]} \\ \mathbf{0} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{0} \\ -\mathbf{C}_{IV}^{[M]} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{C}_{IV}^{[M]} \\ \mathbf{0} & \mathbf{I}\end{bmatrix} \tag{4}$$

and this factorization is shown in Figure 4. Then, because none of the terms in Equation (4) exist in Equation (3), let us consider the double parallel block system of DCT-II in Equation (3) as

$$\begin{aligned} &diag\left\{\begin{bmatrix}\mathbf{C}_{II}^{[M]} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{II}^{[M]}\end{bmatrix}, \begin{bmatrix}\mathbf{C}_{II}^{[M]} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{II}^{[M]}\end{bmatrix}\right\} = \\ &\begin{bmatrix}\mathbf{P}^{[2M]} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}^{[2M]}\end{bmatrix}\begin{bmatrix}\mathbf{S}^{[2M]} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}^{[2M]}\end{bmatrix}\begin{bmatrix}\mathbf{C}_{IV}^{[2M]} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{IV}^{[2M]}\end{bmatrix} \\ &\times \begin{bmatrix}\mathbf{T}^{[2M]} & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{[2M]}\end{bmatrix} \times diag\left\{\begin{bmatrix}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{[M]}\end{bmatrix}, \begin{bmatrix}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{[M]}\end{bmatrix}\right\} \end{aligned} \tag{5}$$

It is clear that $diag\{\mathbf{C}_{IV}^{[M]}, \mathbf{C}_{IV}^{[M]}\}$ in Equation (5) can be factorized into lifting structures using Equation (4).

Table 1: Lossless image coding results (LBR [bpp])

| Filter | [6] | [7] | Prop. IntDCT | | |
|---|---|---|---|---|---|
| $M$ | 8 | 8 | 8 | 16 | 32 |
| Baboon | 6.33 | 6.93 | 6.32 | 6.22 | **6.20** |
| Barbara | 5.07 | 5.54 | 5.05 | 4.85 | **4.80** |
| Boat | 5.27 | 5.77 | 5.26 | 5.14 | **5.13** |
| Elaine | 5.30 | 5.95 | 5.30 | 5.20 | **5.08** |
| Finger1 | 6.11 | 6.41 | 6.11 | 5.84 | **5.72** |
| Finger2 | 5.91 | 6.17 | 5.90 | 5.59 | **5.45** |
| Goldhill | 5.23 | 5.74 | 5.22 | 5.11 | **5.09** |
| Grass | 6.22 | 6.70 | 6.22 | 6.12 | **6.07** |
| Lena | 4.74 | 5.26 | 4.72 | **4.61** | **4.61** |
| Pepper | 5.02 | 5.55 | 5.02 | **4.94** | 4.97 |
| Avg. | 5.52 | 6.00 | 5.51 | 5.36 | **5.31** |

## 4. Simulation Results

In this section, the use of the proposed IntDCT for lossless-to-lossy image coding is validated by considering the lossless bit rate (LBR) [bpp] for lossless image coding and the peak signal-to-noise ratio (PSNR) [dB] for lossy image coding. The set partitioning in hierarchical trees (SPIHT) progressive image transmission algorithm [10] was used to encode the transformed images. We used the results in [6, 7] for comparison and ten 512 $\times$ 512 gray scale test images including “Barbara”, “Goldhill” and “Lena”.

### 4.1. Lossless image coding

First, the proposed IntDCT is applied to lossless image coding. The comparison of

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

for lossless image coding is shown in Table 1. The coding performance of the proposed IntDCT is better than that of the conventional methods.

### 4.2. Lossy image coding

If lossy compressed data is required, this can be achieved by interrupting the obtained lossless bit stream. Therefore, we can use the existing JPEG decoder to reconstruct an image. The comparison of

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

where MSE is the mean squared error, with other lossy image coding methods is shown in Table 2 and Figure 5. In Table 2 and Figure 5, we see that our IntDCT gives better results in terms of the PSNR than the conventional methods and that

Table 2: Lossy image coding results (PSNR [dB])

| Filter | [6] | [7] | Prop. IntDCT | | |
|---|---|---|---|---|---|
| $M$ | 8 | 8 | 8 | 16 | 32 |
| bit rate: 0.25 [bpp] | | | | | |
| Barbara | 25.91 | 23.33 | 25.91 | 27.95 | **28.71** |
| Goldhill | 28.73 | 27.34 | 28.73 | 29.70 | **29.97** |
| Lena | 30.74 | 29.08 | 30.75 | 32.65 | **33.04** |
| bit rate: 0.50 [bpp] | | | | | |
| Barbara | 29.83 | 27.24 | 29.83 | 31.79 | **32.44** |
| Goldhill | 31.58 | 28.77 | 31.59 | 32.32 | **32.48** |
| Lena | 34.77 | 32.75 | 34.77 | 36.09 | **36.18** |
| bit rate: 1.00 [bpp] | | | | | |
| Barbara | 35.52 | 31.58 | 35.51 | 36.89 | **37.20** |
| Goldhill | 34.99 | 31.72 | 35.01 | 35.52 | **35.60** |
| Lena | 38.27 | 35.91 | 38.34 | **39.08** | 39.02 |

the high-frequency region in the image reconstructed using our IntDCT is well approximated.

## 5. Conclusion

This paper proposed an $M$-channel integer discrete cosine transform (IntDCT) using the relation between DCT-IV and the parallel block system of DCT-II. Our proposed IntDCT was validated for lossless-to-lossy image coding. As a result, in the case of 8-channel our proposed method was comparable to conventional IntDCTs, and in the case of 16 and 32-channel it exhibited better coding performance. Since our proposed method can be easily extended to a larger block size, it is promising for various applications.

Figure 5: Enlarged images of “Goldhill” for lossy image coding: (top left) original, (top right) using the method of Komatsu $et\ al.$ [6], (middle left) using the method of Liang $et\ al.$ [7], (middle right) using the 8-channel proposed method, (bottom left) using the 16-channel proposed method, (bottom right) using the 32-channel proposed method

## References


[1] W. Pennebaker and J. Mitchell: JPEG, Still Image Data Compression Standard, Van Nostrand, 1993.

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

[3] A. Skodras, C. Christopoulis and T. Ebrahimi: The JPEG2000 still image compression standard, IEEE Trans. Signal Process. Mag., Vol. 18, No. 5, pp. 36-58, 2001.

[4] F. Dufaux, G. J. Sullivan and T. Ebrahimi: The JPEG XR image coding standard, IEEE Signal Process. Mag., Vol. 26, No. 6, pp. 195-199, 204, 2009.

[5] W. Sweldens: The lifting scheme: a custom-design construction of biorthogonal wavelets, Appl. Comput. Harmon. Anal., Vol. 3, No. 2, pp. 186-200, 1996.

[6] K. Komatsu and K. Sezaki: Reversible discrete cosine transform, Proc. ICASSP’98, May 1998.

[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, 2001.

[8] Z. Wang: On computing the discrete fourier and cosine transforms, IEEE Trans. Acoust. Speech Signal Process., Vol. 33, No. 4, pp. 1341-1344, 1985.

[9] G. Schuller, S. Oraintara, Y. Yokotani, R. Geiger and K. R. Rao: Lossless audio coding using the IntMDCT and rounding error shaping, IEEE Trans. Speech Audio Language Process., Vol. 14, No. 6, pp. 2201-2211, 2006.

[10] 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, 1996.