2010 International Workshop on Nonlinear Circuits,
Communication and Signal Processing
NCSP'10, Waikiki, Hawaii, March 3-5, 2010

M-channel IntDCT using Relation between DCT-IV and Block Parallel 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
Email: {suzuki, hayano, ikehara}@tkhm.elec.keio.ac.jp

Abstract

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

Index Terms — integer discrete cosine transform (IntDCT), lossless-to-lossy image coding, arbitrary block size, block parallel system.

1. Introduction

In JPEG [1] which 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 has several types: type-I to -VIII, excellent energy compaction capability, numerous fast implementations and myriad developments of its device. DCT type-II (DCT-II) and -III (DCT-III) are commonly called DCT and inverse DCT (IDCT). Consequently, DCT is especially promising, since it is compatible with the lossy JPEG.

Recently, with the rapid development of hardware as PC and mobile phone etc. 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 been already achieved by JPEG2000 [3] and HD Photo (JPEG-XR) [4]. However, the existing state-of-the-art technology as JPEG2000 and JPEG-XR can not be the substitute standard of JPEG because it is mostly used and its compressed data is spreads too much over the world. It means that the transform for lossless-to-lossy image coding must have the compatibility with JPEG. Additionally, the encoded image by such transform can be also decoded even 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), 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 described about only block size 8. Its size may not be the best one in various applications including image coding.

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

Notations: $\mathbf{I}$, $\mathbf{J}$, $\mathbf{D}$, $diag\{.\}$, $\mathbf{M}^T$ and $\mathbf{M}^{[M]}$ are identity matrix, invertible matrix, a diagonal matrix, a block diagonal matrix, a transpose of matrix and an $M \times M$ matrix. In this paper, let us define as $\left[\mathbf{D}^{[\mathbf{M}]}\right]_{ii} = (-1)^i$ ($0 \leq i \leq M-1$).

2. Review

2.1. Lifting Structure

Lifting structure [5] is researched widely for various applications. Since the structure achieves integer-to-integer transform, lossless-to-lossy image coding can be implemented. In Fig. 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 lifting coefficients T_0 and T_1 are represented by

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

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

In this case, the lifting matrices and their inverse matrices

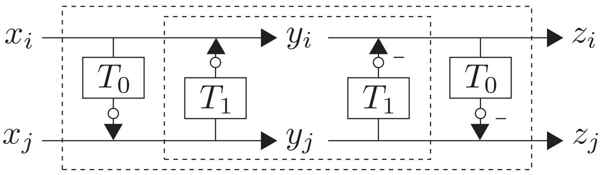

Figure 1: Lifting structures (white circles: rounding operations).

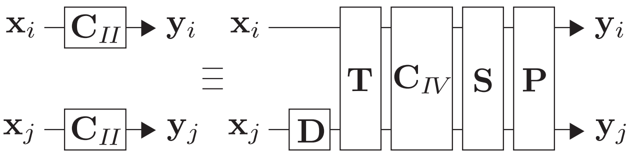

Figure 2: Relation between DCT-IV and block parallel system of DCT-II.

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 $0 \le m, n \le M-1$ 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 Block Parallel System of DCT-II

3.1. Relation between DCT-IV and Block Parallel System of DCT-II

We consider to process two individual signals by DCT-II as shown in the left part of Fig. 2. It is called block parallel system of DCT-II, and expressed by

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

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 $0 \le i \le 2M-1$,

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

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

and

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

Here, (2) can be represented by

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

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

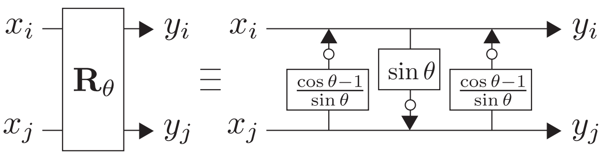

Figure 3: Lifting factorization of Givens rotation matrix (white circles: rounding operations).

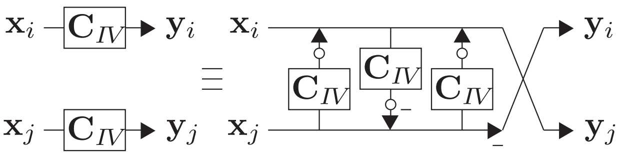

Figure 4: Lifting factorization of DCT-IV (white circles: rounding operations).

3.2. Lifting Factorization of M-channel DCT-IV and Rotation Matrices

For achievement to lossless-to-lossy image coding, $\mathbf{S}^{[2M]}$, $\mathbf{T}^{[2M]}$ and $\mathbf{C}_{IV}^{[2M]}$ in (3) must be factorized to 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 lifting factorization of $\mathbf{S}^{[2M]}$ and $\mathbf{T}^{[2M]}$. Fig. 3 shows lifting factorization of Givens rotation matrix. Next, we can consider to process two individual signals by DCT-IV like DCT-II in (1) as $diag\{\mathbf{C}_{IV}^{[M]}, \mathbf{C}_{IV}^{[M]}\}$. It is known that block parallels system of DCT-IV can be factorized to lifting structures which just have 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 by Fig. 4. Then, because (4) does not exist in (3), let us consider that double block parallel system of DCT-II in (3) as

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

It is clear that $diag\{\mathbf{C}_{IV}^{[M]}, \mathbf{C}_{IV}^{[M]}\}$ in (5) can factorized into lifting structures by (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 proposed IntDCTs are 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. The set partitioning in hierarchical trees (SPIHT) progressive image transmission algorithm [10] was used to encode the transformed images. We chose [6, 7] as the conventional IntDCTs and ten 512×512 grayscale test images, for example, "Goldhill" and "Lena".

4.1. Lossless Image Coding

First, the proposed IntDCTs are applied to lossless image coding. The comparison of

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

in lossless image coding is shown in Table 1. The coding performance of the proposed IntDCTs is better than the performance of all of the conventional methods.

4.2. Lossy Image Coding

If lossy compressed data is required, it 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

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

where MSE is the mean squared error, with other lossy image coding methods is shown in Table 2 and Fig. 5. In Table 2 and Fig. 5, we see that our IntDCTs denote better results on the PSNR against the conventional ones and the high frequency region in the reconstructed image using our IntDCTs is well-approximated.

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

Filter	[6]	[7]	Prop. IntDCT		
M	8	8	8	16	32
bitrate: 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**
bitrate: 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**
bitrate: 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

Figure 5: Enlarged images of “Goldhill” in lossy image coding: (top-left) original, (to-right) Komatsu’s [6], (middle-left) Liang’s [7], (middle-right) 8-channel prop., (bottom-left) 16-channel prop., (bottom-right) 32-channel prop.

5. Conclusion

This paper proposed M-channel integer discrete cosine transform (IntDCT) using relation between DCT-IV and block parallel system of DCT-II. Our proposals were validated in lossless-to-lossy image coding. As a result, 8-channel proposal was comparable to the conventional IntDCTs, and 16 and 32-channel proposals showed better coding performance. Since our proposal achieves easy extension to larger size, it has a perspective for various applications.

References

[1] W. Pennebaker and J. Mitchell, *JPEG, Still Image Data Compression Standard,* NY: 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, Sep. 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, Nov. 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,” in *Proc. of ICASSP’98,* Seattle, WA, 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, Dec. 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, Apr. 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 and Language Process.*, vol. 14, no. 6, pp. 2201–2211, Nov. 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.