A SIMPLIFIED REALIZATION OF NORMALIZED INTEGER WHT FOR MULTIPLIERLESS INTEGER DCT

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

Walsh-Hadamard transform (WHT) based multiplierless integer discrete cosine transform (IntDCT) has structural regularity even in short word length lifting coefficients. It, however, cannot apply to image coding without the quantization part because WHT was implemented by only ± 1 adder operations without the normalization scaling factors. Although we have already presented a normalized integer WHT (IntWHT) as its solution, it also has many adder operations. In this paper, using a two-dimensional (2-D) separable transform of one-dimensional (1-D) normalized WHT is applied to each lifting coefficient, we present a more simplified realization of normalized IntWHT with structural regularity for short word length lifting coefficients. Finally, in lossless-to-lossy image coding, IntDCT based on the proposed IntWHT is validated.

Index Terms— Integer discrete cosine transform (IntDCT), integer Walsh-Hadamard transform (IntWHT), lossless-to-lossy image coding

1. INTRODUCTION

JPEG [1] based on discrete cosine transform (DCT) is the first standard for still image coding and is limited to only lossy image coding. Although lossless JPEG [2] based on differential pulse code modulation (DPCM) has been proposed for lossless image coding, each algorithm is required to obtain both lossy and lossless data due to incompatibility between DCT and DPCM, respectively. On the other hand, JPEG2000 [3] and JPEG XR [4] achieve the unification of lossy and lossless image coding, which is called lossless-to-lossy image coding, by a discrete wavelet transform (DWT) and a hierarchical lapped biorthogonal transform (HLBT) constructed by lifting structures [5] and rounding operations. Nevertheless, these next generation standard cannot decode the existing JPEG data because they do not have the compatibility with JPEG standard.

Then, an integer DCT (IntDCT) compatible with DCT in JPEG is required and several IntDCTs have been already proposed [6–9]. However, the conventional IntDCTs generate checker-board artifacts in short word length lifting coefficients, $k/2^n$ (k, $n \in \mathbb{N}$), and low bitrate coding. Such coefficients are extremely desired for a realization of software/hardware of IntDCT.

In this paper, we present a more simplified realization of normalized IntWHT with structural regularity for short word length lifting coefficients. Our previous work [10] has already achieved by using a transposition of normalization factors and considering a two-dimensional (2-D) separable transform of one-dimensional (1-D) normalized WHT. This paper improves coding performance by direct application of a 2-D separable transform of 1-D normalized WHT to each lifting structure in IntWHT. As result, our IntDCT is validated in lossless-to-lossy image coding.

Notations: $\mathbf{I}$ is an identity matrix, $\mathbf{M}^T$ is a transpose of matrix $\mathbf{M}$ and $\mathbf{M}^{[N]}$ is an $N \times N$ square matrix, respectively. Also, for simplicity, $M = 2^m$ ($m \in \mathbb{N}$).

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.

2. REVIEW

2.1. Multiplierless Lifting Structure

The lifting structure [5], also known as the ladder structure, is a special type of lattice structure, a cascading construction using only elementary matrices - identity matrices with single nonzero off-diagonal element.

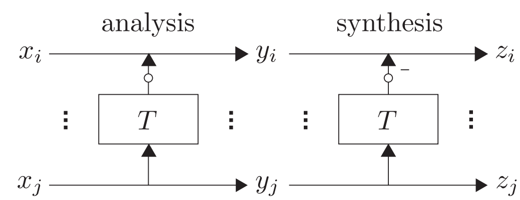


Fig. 1. A lifting structure (white circles: rounding operations).

Fig. 1 shows a basic lifting structure. In this case, the lifting matrix and its inverse matrix are as follows:

$$\begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -T \\ 0 & 1 \end{bmatrix}$$

where T is lifting coefficient. Then, they are represented by

$$\begin{aligned} y_i &= x_i + round[Tx_j] \\ y_j &= x_j \end{aligned} \quad \rightarrow \quad \begin{aligned} z_i &= y_i - round[Ty_j] = x_i \\ z_j &= y_j = x_j \end{aligned}$$

where $round[.]$ is a rounding operation. Thus the lifting structure with rounding operation can achieve lossless-to-lossy coding.

For high-speed implementation, lifting coefficients are required to approximate floating-point to hardware-friendly dyadic values such as $k/2^n$ (k, $n \in \mathbb{N}$) which can be implemented by only bit shift and addition operations. It performs fast implementation in a

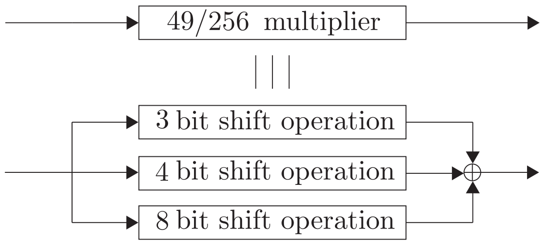


Fig. 2. An approximation from floating multiplication to bit shift and addition operations.

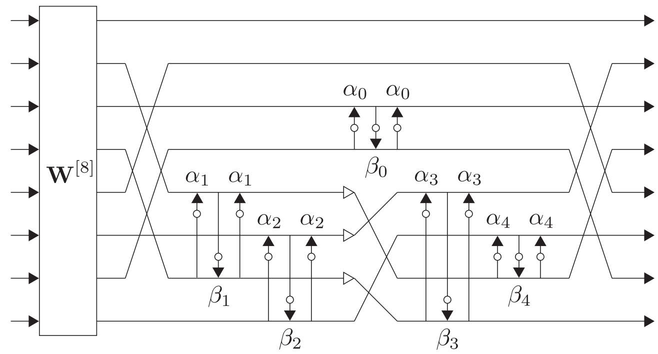


Fig. 3. 8-channel IntDCT based on WHT (white circles and white triangles: rounding operations and -1 operations, respectively).

real time software encoder and reduces the circuit size. The multiplications are expressed by $k/2^n$ which is n-bit word length. For example, a coefficient $49/2^8 = 49/256$ can be operated as

$$\frac{49}{256} = \frac{32}{256} + \frac{16}{256} + \frac{1}{256} = \frac{1}{2^3} + \frac{1}{2^4} + \frac{1}{2^8}. \tag{1}$$

From (1), we can replace the multiplier by $49/256$ to sum results from 3-bit shift, 4-bit shift and 8-bit shift operations as illustrated in Fig. 2. We can find that the perfect reconstruction in lifting structure is always kept even if lifting coefficients are approximated.

2.2. Regularity

In filter bank (FB) included DCT theory, the regularity is an important property for image compression [11]. An M-channel DCT has one regularity condition if DCT matrix $\mathbf{C}^{[M]}$ satisfies

$$\mathbf{C}^{[M]} \begin{bmatrix} 1 \\ z^{-1} \\ \vdots \\ z^{-(M-1)} \end{bmatrix} \Bigg|_{z=1} = \mathbf{C}^{[M]} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} \sqrt{M} \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{2}$$

The conventional IntDCTs generate checker-board artifacts in short word length lifting coefficients, $k/2^n$, and low bitrate coding due to losing regularity condition. However, IntDCT by Chen [8] generates no artifacts even in such case because it has structural regularity.

3. SIMPLIFIED MULTIPLIERLESS INTDCT BASED ON NORMALIZED INTWHT

3.1. Multiplierless IntDCT based on WHT

As well known, WHT has interesting relationships with other discrete transforms such as DCT. It is expressed by

$$\mathbf{W}^{[1]} = [1], \quad \mathbf{W}^{[2]} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$
$$\mathbf{W}^{[M]} = \mathbf{W}^{[2]} \otimes \mathbf{W}^{[\frac{M}{2}]} = \begin{bmatrix} \mathbf{W}^{[\frac{M}{2}]} & \mathbf{W}^{[\frac{M}{2}]} \\ \mathbf{W}^{[\frac{M}{2}]} & -\mathbf{W}^{[\frac{M}{2}]} \end{bmatrix} \tag{3}$$

where $\otimes$ is Kronecker product. Note that (3) is still not normalized by the scaling factor $1/\sqrt{M}$. Let us normalize $\mathbf{W}$ such that $\hat{\mathbf{W}} = (1/\sqrt{M})\mathbf{W}$ is an orthonormal matrix: $\hat{\mathbf{W}}^{-1} = \hat{\mathbf{W}}^T$. Then one can prove that [8]

$$\mathbf{C}^{[M]} = \hat{\mathbf{Q}}^{[M]}\hat{\mathbf{W}}^{[M]} = \mathbf{P}^{[M]^T}\mathbf{Q}^{[M]}\mathbf{P}^{[M]}\hat{\mathbf{W}}^{[M]} \tag{4}$$

Table 1. Lifting coefficients of 8-channel IntDCT based on WHT.

	floating-point	$k/2^3$	$k/2^4$	$k/2^5$	$k/2^6$
α_0	0.1989123674…	1/4	3/16	3/16	13/64
β_0	-0.3826834324…	-3/8	-3/8	-3/8	-3/8
α_1	-0.6681786379…	-5/8	-11/16	-21/32	-43/64
β_1	0.9238795326…	7/8	15/16	15/16	59/64
α_2	-0.6681786379…	-5/8	-11/16	-21/32	-43/64
β_2	0.9238795326…	7/8	15/16	15/16	59/64
α_3	-0.8206787908…	-7/8	-13/16	-13/16	-53/64
β_3	0.9807852804…	1	1	31/32	63/64
α_4	-0.3033466836…	-1/4	-5/16	-5/16	-19/64
β_4	0.5555702330…	1/2	9/16	9/16	9/16

where $\mathbf{P}$ is a permutation matrix and $\mathbf{Q}$ is an orthogonal matrix. Using this factorization, Chen has proposed an IntDCT [8]. Fig. 3 and Table 1 show 8-channel IntDCT based on WHT by Chen and its lifting coefficients, respectively. In this regard, however, this IntDCT by Chen is ill-suited for lossless image coding because it needs a quantization part due to use of unnormalized WHT. On the other hand, we find that (2) is represented by

$$\mathbf{C}^{[M]} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \hat{\mathbf{W}}^{[M]} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} \sqrt{M} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

where $\hat{\mathbf{W}}^{[M]}$ is the pre-processing part in DCT. This means that WHT satisfies all regularity condition in DCT. Hence, similar to our previous work [10], this paper solves the normalization problem of WHT: $\hat{\mathbf{W}}^{[8]} = (1/\sqrt{8})\mathbf{W}^{[8]}$, in IntDCT by Chen while keeping the other part as it is.

3.2. A Realization of Normalized IntWHT

We have presented a multiplierless IntDCT based on normalized IntWHT [10]. First, a 2-D separable transform of 1-D DCT $\mathbf{C}$ is considered. When we apply $\mathbf{C}$ into a 2-D input signal $\mathbf{x}$ in column- and row-wise separately, the 2-D separable transform is expressed by

$$\mathbf{y} = (\mathbf{C}(\mathbf{C}\mathbf{x})^T)^T = \mathbf{C}\mathbf{x}\mathbf{C}^T \tag{5}$$

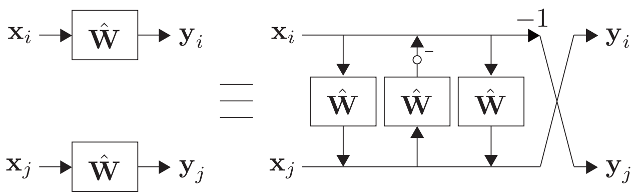


Fig. 4. 1-D transform of normalized IntWHT (white circles: rounding operations).

where $\mathbf{y}$ is the output signal. Since the transform $\mathbf{C}$ can be factorized as $\mathbf{C} = \hat{\mathbf{Q}}\hat{\mathbf{W}}$ in (4), (5) is represented by

$$\mathbf{y} = \hat{\mathbf{Q}}\hat{\mathbf{W}}\mathbf{x}(\hat{\mathbf{W}}\hat{\mathbf{Q}})^T = \hat{\mathbf{Q}}\hat{\mathbf{W}}\mathbf{x}\hat{\mathbf{W}}^T\hat{\mathbf{Q}}^T.$$

This equation means that 2-D separable transform by $\hat{\mathbf{Q}}$ is applied after 2-D separable transform by $\hat{\mathbf{W}}$.

Next, we consider to process two individual signals $\mathbf{x}_i$ and $\mathbf{x}_j$ by normalized WHT $\hat{\mathbf{W}}$ as shown in the left part of Fig. 4. The output signals are transformed by

$$\begin{bmatrix} \mathbf{y}_i \\ \mathbf{y}_j \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{W}} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{W}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_i \\ \mathbf{x}_j \end{bmatrix}. \tag{6}$$

$\mathrm{diag}\{\hat{\mathbf{W}}, \hat{\mathbf{W}}\}$ in (6) can be factorized into the complete lifting structures such as

$$\begin{bmatrix} \hat{\mathbf{W}} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{W}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \hat{\mathbf{W}} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\hat{\mathbf{W}} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \hat{\mathbf{W}} & \mathbf{I} \end{bmatrix} \tag{7}$$

where $\hat{\mathbf{W}}\hat{\mathbf{W}} = \mathbf{I}$.

Thus the parallel block system of normalized WHTs can be implemented as shown in the right part of Fig. 4. However, if $\hat{\mathbf{W}}$ is a normalized WHT with nonideal coefficients, the system in (7) cannot be consisted because of $\hat{\mathbf{W}}\hat{\mathbf{W}} \neq \mathbf{I}$. It means short word length allocation cannot be applied to lifting coefficients. In [10], using a transposition of normalization factors and considering a 2-D separable transform of 1-D normalized WHT as shown in Fig. 5(a), the problem has been solved.

3.3. A More Simplified Realization of Normalized IntWHT

In this paper, we consider that $\hat{\mathbf{W}}$ is applied to an $M \times M$ input signal $\mathbf{x}$ in column- and row-wise separately in each lifting structure. If $\mathcal{W}(\mathbf{x})$ is defined as

$$\mathcal{W}(\mathbf{x}) \triangleq \hat{\mathbf{W}}\mathbf{x}\hat{\mathbf{W}}^T = \frac{1}{M}\mathbf{W}\mathbf{x}\mathbf{W}^T,$$

(7) is represented by

$$\begin{bmatrix} \mathcal{W}(\mathbf{x}) & \mathbf{0} \\ \mathbf{0} & \mathcal{W}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathcal{W}(\mathbf{x}) & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathcal{W}(\mathbf{x}) \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathcal{W}(\mathbf{x}) & \mathbf{I} \end{bmatrix}.$$

As result, the normalized IntWHT can be constructed by only $\log_2 M$-bit shifts and ± 1 adder operations without losing structural regularity. Also, since lifting structures are reduced by half in case of our previous work [10] as shown in Fig. 5(b), it is obvious that $3M^2$ adder operations are eliminated.

Table 2. Comparison of coding gain (C_G [dB]).

Lifting coefficients	$k/2^3$	$k/2^4$	$k/2^5$	$k/2^6$	float
Tran's [6]	6.14	6.21	6.23	6.22	6.23
Hao's [7]	7.98	8.41	8.75	8.75	**8.83**
Chokchaitam's [9]	8.08	8.61	8.72	8.75	8.76
Previous work [10]	**8.73**	**8.82**	**8.82**	**8.82**	**8.83**
Prop. IntDCT	**8.73**	**8.82**	**8.82**	**8.82**	**8.83**

3.4. Expansion from Normalized IntWHT to IntDCT

Finally, the normalized IntWHT is applied to the pre-processing part in IntDCT. In that time, $\hat{\mathbf{Q}}^{[M]}$ as post-processing is implemented as it is. This IntDCT has structural regularity because the normalized IntWHT has structural regularity. In this paper, M is 8 and all transform coefficients exclusive of ones in WHT are used ones in [10].

4. RESULTS

4.1. Coding Gain and Frequency Response

Coding gain is one of the most important factors to be considered in compression applications. A transform with higher coding gain compacts more energy into a fewer number of coefficients. As a result, higher objective performances such as PSNR would be achieved after quantization. Since coding gain of the DCT approximates the optimal KLT closely, it is desired that the our IntDCT has similar coding gain to that of the original DCT. The biorthogonal coding gain C_G is defined as [11]

$$C_G \text{ [dB]} = 10\log_{10} \frac{\sigma_x^2}{\prod_{k=0}^{M-1} \sigma_{x_k}^2 \parallel f_k \parallel^2}$$

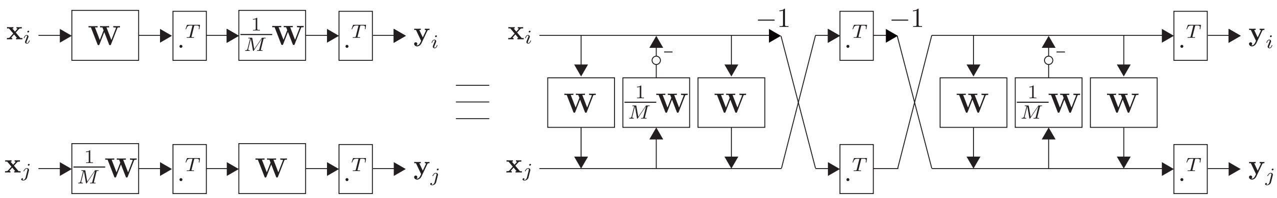
where σ_x^2 is the variance of the input signal, $\sigma_{x_k}^2$ is the variance of the k-th subbands and $\parallel f_k \parallel^2$ is the norm of the k-th synthesis filter. Table 2 shows the comparison of coding gain of the proposed IntDCT and the conventional ones. It is clear that coding gains of the conventional IntDCTs are not kept in short word length lifting coefficients, but the proposed IntDCT almost kept it even in 4-bit word length lifting coefficients.

Frequency responses of the proposed and conventional IntDCTs are shown in Fig. 6. It is clear that the conventional IntDCTs have DC leakage which may generate a checker-board artifact in low bitrate coding. The proposed IntDCT does not have DC leakage because regularity in the proposed normalized IntWHT as preprocessing in IntDCT can be always kept in short word length lifting coefficients. Also, although Tran's BinDCT [6] without the scaling factors do not have DC leakage, it shows worse coding gain than others. Thus, it is deselected from a target for comparison in lossless-to-lossy image coding.

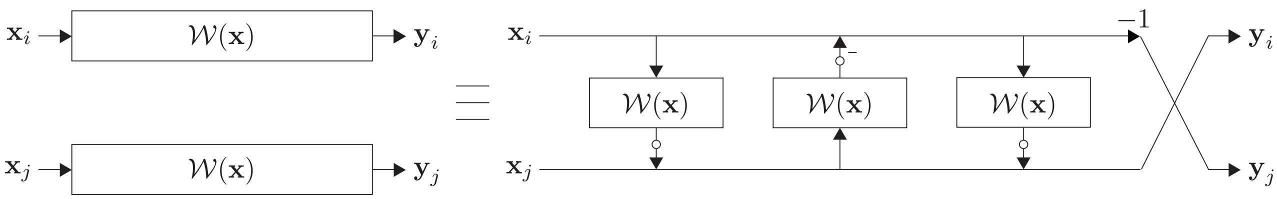
4.2. Lossless-to-Lossy Image Coding

The proposed IntDCT is applied to lossless-to-lossy image coding. twenty test images, which have 512×512 size, were used. The set partitioning in hierarchical trees (SPIHT) progressive image transmission algorithm [12] was used to encode the transformed images. The comparison of lossless bitrate (LBR)

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



(a)



(b)

Fig. 5. 2-D separable transform of normalized IntWHTs (white circles and $.^T$: rounding operations and a transpose of .): (a) our previous work [10], (b) the proposed IntWHT.

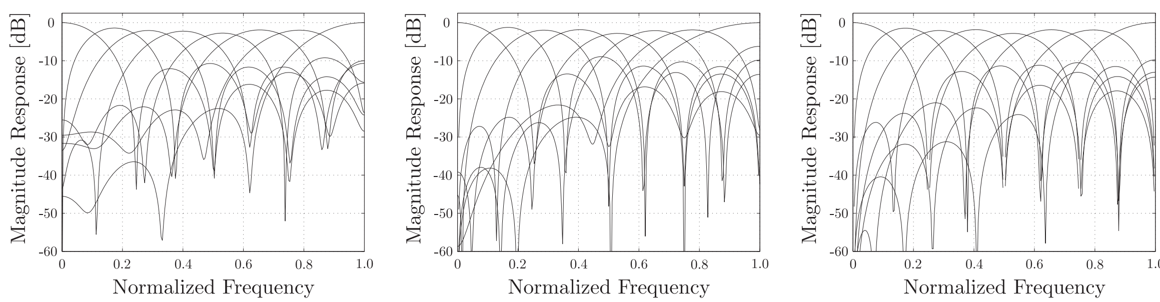


Fig. 6. Frequency responses in 4-bit word length lifting coefficients: (left) Hao's [7], (middle) Chokchaitam's [9], (right) the proposed IntDCT.

Table 3. Comparison of lossless image coding in 4-bit word length lifting coefficients (LBR [bpp]).

Images	[7]	[9]	[10]	Prop.
Baboon	6.84	6.35	6.28	**6.27**
Barbara	5.88	5.14	5.02	**4.97**
Boat	6.19	5.34	5.22	**5.19**
Elaine	6.31	5.42	5.27	**5.24**
Finger1	6.48	6.10	6.08	**6.07**
Finger2	6.45	5.92	5.87	**5.85**
Goldhill	5.89	5.28	5.18	**5.16**
Grass	6.52	6.22	6.18	**6.17**
Lena	5.89	4.87	4.68	**4.64**
Pepper	6.05	5.13	4.98	**4.95**
Avg.	6.23	5.54	5.43	**5.40**

in lossless image coding is shown in Table 3 where Avg. is average of LBRs in twenty images.

If lossy compressed data is required, it can be achieved by interrupting the obtained lossless bit stream. The comparison of peak signal-to-noise ratio (PSNR)

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

where MSE is the mean squared error in lossy image coding are shown in Table 4. Furthermore, Fig. 7 shows the enlarged images of "Lena" which are lossy compressed images by the proposed IntDCT and the conventional ones with 4-bit word length lifting coefficients when bitrate is 0.25[bpp].

In Table 3, 4 and Fig. 7, the proposed IntDCT presents better performance than the conventional ones in lossless-to-lossy image coding. Also, in Fig. 7, we can find that a checker-board artifact is not generated in the proposed IntDCT due to structural regularity.

5. CONCLUSION

This paper presents a more simplified realization of normalized integer Walsh-Hadamard transform (IntWHT) for multiplierless integer discrete cosine transform (IntDCT). It is obtained by considering that normalized WHT is applied to an input signal in column- and row-wise separately in each lifting structure. It has not only structural regularity even in short word length lifting coefficients, but also less adder operations and better performance than our previous work. It

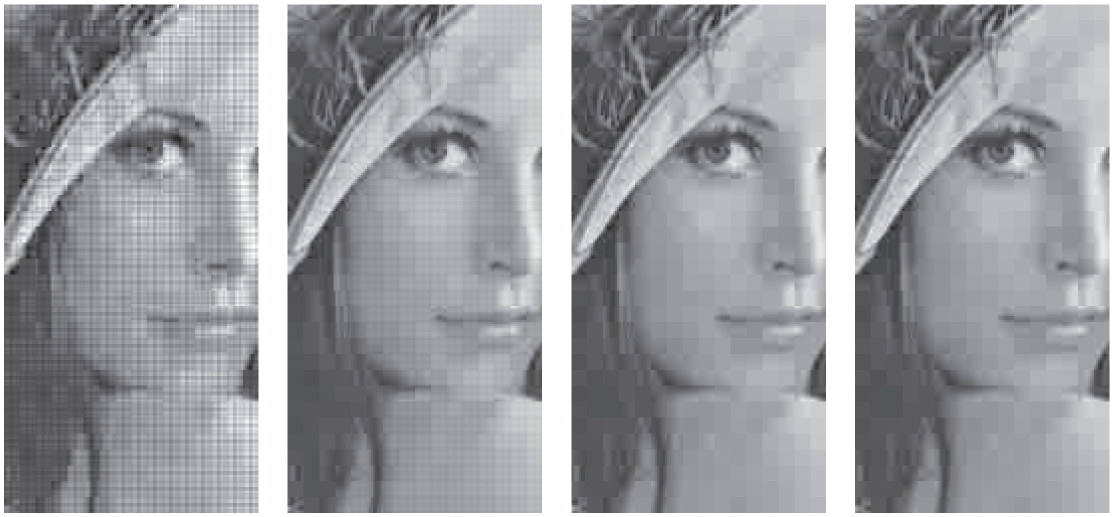


Fig. 7. Comparison of lossy reconstructed images "Lena" in 4-bit word length lifting coefficients (bitrate: 0.25 [bpp]): (left)-(right) Hao's [7], Chokchaitam's [9], our previous work [10] and the proposed IntDCT.

Table 4. Comparison of lossy image coding in 4-bit word length lifting coefficients (PSNR [dB]).

Images	bitrate [bpp]	[7]	[9]	[10]	Prop.
Barbara	0.25	23.78	26.49	26.92	**26.94**
	0.50	25.95	29.99	30.64	**30.66**
	1.00	29.71	35.04	35.77	**35.95**
Boat	0.25	23.21	28.14	28.61	**28.63**
	0.50	25.17	31.21	31.79	**31.85**
	1.00	28.63	34.54	35.26	**35.43**
Finger1	0.25	20.96	22.19	**22.53**	**22.53**
	0.50	22.90	25.28	25.53	**25.54**
	1.00	26.43	29.31	29.55	**29.60**
Goldhill	0.25	25.37	28.88	29.34	**29.37**
	0.50	27.20	31.42	31.87	**31.94**
	1.00	30.69	34.56	35.11	**35.29**
Lena	0.25	24.14	31.37	31.79	**31.84**
	0.50	26.31	34.22	35.45	**35.50**
	1.00	30.17	37.03	38.54	**38.98**
Pepper	0.25	24.22	31.05	31.35	**31.39**
	0.50	26.52	33.17	34.35	**34.39**
	1.00	30.53	35.68	36.59	**36.86**

was proved by the comparisons of coding gains, frequency responses and lossless-to-lossy image coding simulations.

6. REFERENCES

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

[2] ISO/IEC 14495-1, *Information technology - Lossless and near-lossless compression of continuous-tone still images: Baseline*.

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

[6] T. D. Tran, "The BinDCT: fast multiplierless approximation of the DCT," *IEEE Signal Process. Lett.*, vol. 7, no. 6, pp. 141–144, June 2000.

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

[8] Y. J. Chen, S. Oraintara, T. D. Tran, K. Amaratunga, and T. Q. Nguyen, "Multiplierless approximation of transforms with adder constraint," *IEEE Signal Process. Lett.*, vol. 9, no. 11, pp. 344–347, Nov. 2002.

[9] S. Chokchaitam and M. Iwahashi, "A new optimum-word-length-assignment (OWLA) multiplierless integer DCT for unified lossless/lossy image coding," in *Proc. of IWAIT'07*, Bangkok, Thailand, Jan. 2007.

[10] T. Suzuki and M. Ikehara, "Structurally regular integer discrete cosine transform for low-bit-word-length coefficients," in *Proc. of ISCAS'10*, Paris, France, June 2010.

[11] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1996.

[12] 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.