Optimum Word Length Allocation for Multipliers of Integer DCT

Masahiro IWAHASHI, Osamu NISHIDA, Somchart CHOKCHAITAM[+] and Noriyoshi KAMBAYASHI
Nagaoka University of Technology, Nagaoka, Niigata, 940-2188, JAPAN
[+] Faculty of Engineering, Thammasat University, Bangkok, THAILAND
iwahashi@nagaokaut.ac.jp, http://tech.nagaokaut.ac.jp/

ABSTRACT

Recently, the integer DCT (Int-DCT), which has the rounding operations in the lifting structure, is attracting many researchers' attention as an effective method for DCT based lossy / lossless unified coding. So far, focuses of the previous reports relevant to the Int-DCT have been limited to a few topics such as how to reduce the number of multipliers with the four point lossless Hadamard transform and the non-separable two dimensional LDCT. What seems to be lacking, however, is how to express multipliers' word length as short as possible for reduction of hardware complexity.

This report defines a new "SNR sensitivity" as an indicator of how the word length truncation of multiplier coefficients affects quality of the decoded image, and also proposes a new word length allocation method based on the sensitivity. As a result, two [bit] in average shorter word length is attained under equivalent quality of the decoded image.

1. INTRODUCTION

The JPEG international standard algorithm based on the discrete cosine transform (DCT) [1] is widely used as a lossy coding in the field of image communications and storages. Recently, the integer transform [2], which includes rounding operations in the lifting structure [3], is becoming popular as a key technique to lossless and lossy unified waveform coding [4]. Especially the integer DCT [5-7] is attractive as the unified coding with compatibility to the conventional DCT based algorithms.

So far, relevant to the integer DCT, previous reports focused on reducing the rounding operations with the non-separable 2D structuring [5] and reducing multipliers with the integer Hadamard transform [7]. Optimization of the basis function of the orthogonal transform (integer KLT) is also reported [8]. What seems to be lacking, however, is how to express multipliers' word length as short as possible for reduction of hardware complexity.

This report proposes an optimum word length allocation for multipliers of the integer DCT. Overview of the integer DCT is summarized in **2.**. The "SNR sensitivity" is defined and evaluated for typical input image data and the sensitivity is applied to optimum word length allocation using the least square method in **3.**. Effectiveness of the proposed method is confirmed in **4.**.

2. THE INTEGER DCT

2.1 Integer DCT (Int-DCT)

Fig.1 illustrates algorithm of the integer DCT (Int-DCT). This transform, which is composed of the integer Hadanard transform (IH) describes in **2.2** and integer rotation transform (IR) in **2.3**, transforms integer input vector x(n), (n=1,2・・・,8) into output vector expressed with integer. Therefore it is possible to achieve effective lossless coding applying an entropy coding directly to the output vector. Lossy coding compatible to the conventional DCT based algorithm is also possible with inserting the quantization procedure.

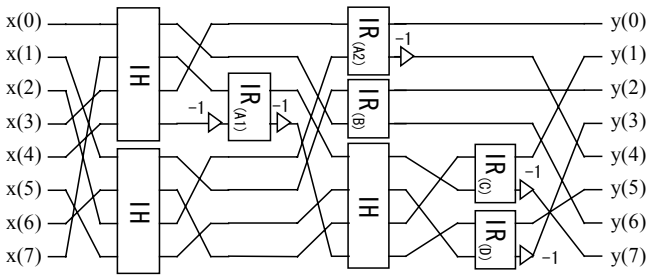


Fig.1 Integer DCT (Forward transform).

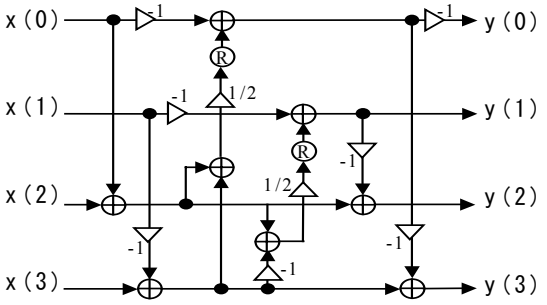


Fig.2 Integer Hadamard transform (IH).

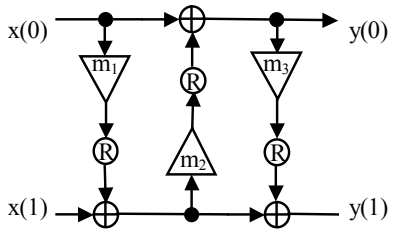


Fig.3 Integer rotation transform (IR).

2.2 Integer Hadamard Transform (IH)

The integer Hadamard transform illustrated in figure 2 contains rounding operations "R" in the lifting structure. Relation between input and output is

$$\begin{pmatrix} -y(0) \\ y(1) \\ y(2) \\ y(3) \end{pmatrix} = \begin{pmatrix} \mathbf{I}_2 & \mathbf{O}_2 \\ -\mathbf{J}_2 & \mathbf{I}_2 \end{pmatrix} \begin{pmatrix} \mathbf{I}_2 & \mathbf{W}_2 \\ \mathbf{O}_2 & \mathbf{I}_2 \end{pmatrix} \begin{pmatrix} -\mathbf{I}_2 & \mathbf{O}_2 \\ \mathbf{P}_2 & \mathbf{I}_2 \end{pmatrix} \begin{pmatrix} x(0) \\ x(1) \\ x(2) \\ x(3) \end{pmatrix} \quad (1)$$

where

$$\mathbf{P}_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \mathbf{J}_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \mathbf{W}_2 = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \mathbf{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \mathbf{O}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

0-7803-7762-1/03/$17.00 ©2003 IEEE

2.3 Integer Rotation Transform (IR)

The integer rotation transform in figure 3 has the in-and-out relation defined by

$$\begin{pmatrix} y(0) \\ y(1) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ m_{3(i)} & 1 \end{pmatrix} \begin{pmatrix} 1 & m_{2(i)} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ m_{1(i)} & 1 \end{pmatrix} \begin{pmatrix} x(0) \\ x(1) \end{pmatrix} \tag{2}$$

where "$m_{1(i)}$, $m_{2(i)}$, $m_{3(i)}$" indicate multiplier coefficients in the i^{th} IR (i= A1, A2, B, C, D). As indicated in figure 1, the Int-DCT has five IRs and each IR has three multipliers. Therefore the Int-DCT has fifteen multipliers in total. Their values are given by

$$\begin{aligned} M_{(A1)} &= M_{(A2)} = \left[1-2^{1/2} \quad 2^{-1/2} \quad 1-2^{1/2}\right] \\ M_{(B)} &= \left[\frac{\sin(\pi/8)-1}{\cos(\pi/8)} \quad \cos(\pi/8) \quad \frac{\cos(3\pi/8)-1}{\cos(\pi/8)}\right] \\ M_{(C)} &= \left[\frac{1-\cos(3\pi/16)}{\sin(3\pi/16)} \quad -\sin(3\pi/16) \quad \frac{1-\cos(3\pi/16)}{\sin(3\pi/16)}\right] \\ M_{(D)} &= \left[\frac{\cos(\pi/16)-1}{\sin(\pi/16)} \quad \sin(\pi/16) \quad \frac{\cos(\pi/16)-1}{\sin(\pi/16)}\right] \end{aligned} \tag{3}$$

where

$$M_{(i)} = [m_{1(i)}\ m_{2(i)}\ m_{3(i)}]$$
(i= A1, A2, B, C, D)

In a LSI implementation, each of these real numbers is approximated to a binary number with finite word length. Purpose of this report is how to allocate the optimum word length to these multiplier coefficients considering the "SNR sensitivity" in **3.2**.

3. WORD LENGTH ALLOCATION

3.1 Word Length Truncation

The multiplier coefficient $m_{j(i)}$, (i= A1, A2, B, C, D, j=1,2,3), is expressed as h_k, (k=0,1,⋯,14), by

$$h_k = (-1)^{B_0} \cdot \sum_{j=1}^{\infty} B_j 2^{-j}, \qquad k = 0,1,\cdots,14 \tag{4}$$

where B_j, (j=0,1,・・・) is 0 or 1. Under the finite word length expression in this report, h_k is truncated into W_k [bit] binary value h_k'. Namely,

$$h_k' = (-1)^{B_0} \cdot \sum_{j=1}^{W_k} B_j' 2^{-j}, \qquad k = 0,1,\cdots,14 \tag{5}$$

Fig.1 indicates relation between

$$\Delta h_k = h_k - h_k' \tag{6}$$

and

$$PSNR = 20\log_{10}\left(\frac{255}{\Delta\sigma_k}\right) \quad [dB] \tag{7}$$

where

$$\Delta\sigma_k = \sqrt{\frac{1}{N}\sum_{n=0}^{N-1} \{y(n)-x(n)\}^2}$$

for k=5, 8, 11 for AR(1) model with ρ =0.95. In the above equations, x(n) is transformed into y(n) by forward Int-DCT with $W_k = \infty$ [bit] and backward Int-DCT with $W_k \neq \infty$ [bit]. These lines in the figure can be expressed by

$$PSNR = c_0 + c_1 \cdot \log_{10}(\Delta h_k) \tag{8}$$

Table 1 indicates c_1 and c_0 for all the coefficients.

3.2 SNR Sensitivity

This report defines the "SNR sensitivity" as effect of the finite word length expression on quality of the decoded image by

$$S_k = \frac{\Delta\sigma_k}{\Delta h_k}, \qquad k = 0,1,\cdots,14 \tag{9}$$

and also defines the "relative SNR sensitivity" by

$$SR_k = \frac{S_k}{\prod_{p=0}^{14} \sqrt[15]{S_p}}, \qquad k = 0,1,\cdots,14 \tag{10}$$

These sensitivities are related to c_0 and c_1 in table 1 as follows. Substituting eq.(9) into eq.(7), PSNR is

$$\begin{aligned} PSNR &= 20\log_{10}\frac{255}{S_k \Delta h_k} \\ &= 20\log_{10}\frac{255}{S_k} - 20\log_{10}(\Delta h_k) \end{aligned} \tag{11}$$

Comparing eq.(8) and eq.(11), we get

$$c_0 = 20\log_{10}\frac{255}{S_k}, \quad c_1 = -20 \tag{12}$$

Therefore relation between the sensitivity S_k and c_0 is

$$S_k = 255 \cdot 10^{-c_0/20} \tag{13}$$

The sensitivities calculated from c_0 in table 1 are also summarized in the same table. Please refer to [9] for theoretical analysis.

Table 1 Sensitivity of the coefficients for AR(1) with ρ =0.95.

k	i	j	C1	C0	Sk	SRk
0	A1	1	−20	33.46	5.42	0.72
1		2	−20	31.57	6.73	0.89
2		3	−20	39.78	2.62	0.35
3	A2	1	−20	**8.85**	**92.01**	**12.19**
4		2	−20	**12.83**	**58.19**	**7.71**
5		3	−20	**5.85**	**129.96**	**17.22**
6	B	1	−20	39.32	2.76	0.37
7		2	−20	35.44	4.31	0.57
8		3	−20	34.50	4.81	0.64
9	C	1	−20	30.67	7.46	0.99
10		2	−20	37.20	3.52	0.47
11		3	−20	29.09	8.96	1.19
12	D	1	−20	41.28	2.20	0.29
13		2	−20	37.50	3.40	0.45
14		3	−20	41.28	2.20	0.29

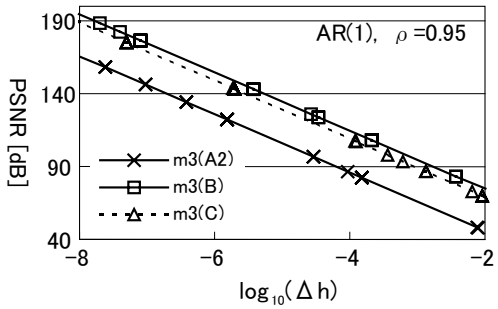


Fig.4 Relation between finite word length errors and image quality.

3.2 Optimum Allocation Method

According to eq.(9), effect of all the truncations on the decoded image is

$$\sigma_{total} = \sum_{k=0}^{14} \Delta\sigma_k = \sum_{k=0}^{14} S_k \cdot \Delta h_k \tag{14}$$

From eq.(4), (5), (6),

$$\Delta h_k = 2^{-Wk}\left(a_k + b_k\right) \tag{15}$$

where

$$a_k = \sum_{j=1}^{\infty} B_{(j+W_k)} \cdot 2^{-j}$$

$$b_k = 2^{W_k}\left(\sum_{j=1}^{W_k} B_j \cdot 2^{-j} - \sum_{j=1}^{W_k} B_j' \cdot 2^{-j}\right)$$

Therefore, substituting eq.(15) into eq.(14),

$$\sigma_{total} = \sum_{k=0}^{14} S_k \cdot 2^{-Wk} \tag{16}$$

Purpose of this report is now summarized as follows.

$$\boxed{\begin{array}{ll} \text{minimize} & \sigma_{total} = \sum_{k=0}^{14} S_k \cdot 2^{-Wk} \\ \text{subject to} & \sum_{k=0}^{14} W_k = 15 \cdot \overline{W} \end{array}} \tag{17}$$

The problem is to fine the optimum value of W_k for each coefficient so that effect of the truncations is minimized under a given average word length. The solution to this problem is

$$\frac{2^{Wk}}{2^{W0}} = \frac{S_k}{S_0} \quad , \qquad k = 0,1,\cdots,14 \tag{18}$$

Namely,

$$W_k = \log_2 \frac{S_k}{S_0} + W_0 \quad , \qquad k = 0,1,\cdots,14 \tag{19}$$

In other way,

$$W_k = \log_2 S_k - \log_2 \overline{S} + \overline{W}, \qquad k = 0,1,\cdots,14 \tag{20}$$

where

$$\overline{W} = \tfrac{1}{15}\sum_{k=0}^{14} W_k,$$

$$\overline{S} = \prod_{k=0}^{14} \sqrt[15]{S_k}$$

Therefore the optimum word length allocation is given by the relative SNR sensitivity SR_k as follows.

$$\begin{aligned} \Delta W_k &= W_k - \overline{W} \\ &= \log_2 \frac{S_k}{\overline{S}} \\ &= \log_2 SR_k \end{aligned} \qquad k = 0,1,\cdots,14 \tag{21}$$

4. EXPERIMENTAL RESULTS

4.1 Allocation Results and Effectiveness

Table 2 summarizes a result of the optimum word length allocation. In the figure, $[\Delta W_k]$ indicates the nearest integer of ΔW_k calculated from eq.(21) with the sensitivity SR_k in table 1. The table indicates that $LR_{(A2)}$ needs long word length but $LR_{(D)}$ does not. Figure 5 illustrates PSNR versus averaged word length. It compares allocation results (proposed) and non-allocation results (existing). It is confirmed that the proposed method saves word length by two [bit] in average under the same PSNR. Fig.6 also shows effectiveness of the proposed method for an image coding.

Table 2 Examples of the optimum word length allocation. Word length $[\Delta W_k]$ for AR(1) with ρ =0.95.

k	i	j	Ex.1	Ex.2	Ex.3	Ex.4	Ex.5	Ex.6
0	A1	1	3	4	5	6	7	8
1		2	3	4	5	6	7	8
2		3	1	2	3	4	5	6
3	A2	1	7	8	9	10	11	12
4		2	6	7	8	9	10	11
5		3	7	8	9	10	11	12
6	B	1	2	3	4	5	6	7
7		2	2	3	4	5	6	7
8		3	2	3	4	5	6	7
9	C	1	3	4	5	6	7	8
10		2	2	3	4	5	6	7
11		3	3	4	5	6	7	8
12	D	1	1	2	3	4	5	6
13		2	2	3	4	5	6	7
14		3	1	2	3	4	5	6

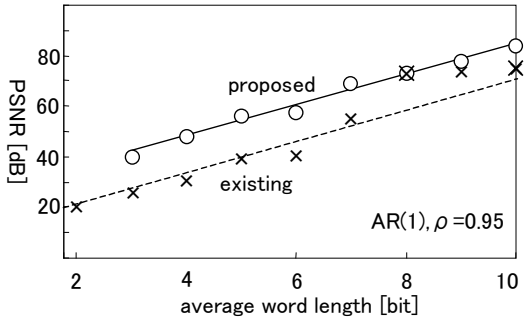
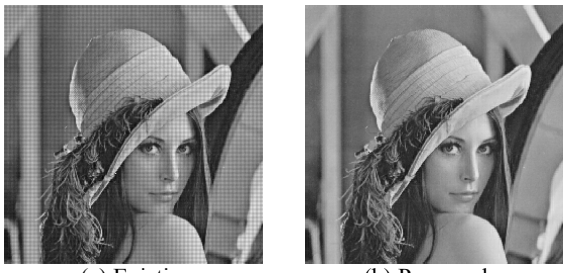


Fig.5 Effectiveness of the proposed method.



(a) Existing (b) Proposed

Fig.6 Comparison with decoded images. (Average word length = 3 bit)

4.2 Effectiveness of the Proposed Method

In the previous sections, the rounding operations, noted as "R" in figure 2 and 3, are neglected. In case of the rounding operations are included, as indicated in figure 7, PSNR saturates at 50 [dB] although the average word length increases. Note that quantization is not used. Therefore the maximum effective word length is 6 [bit] in this case. It is no use to increase word length longer than the maximum effective length summarized in table 8 for various input signals. Considering this maximum effective length, table 2 is renewed as table 9. This is peculiar to the Int-DCT.

Table 8 The maximum effective word length [bit].

input	existing	proposed
AR(1), ρ =0.0	8	7
AR(1), ρ =0.1	8	7
AR(1), ρ =0.5	8	6
AR(1), ρ =0.8	8	6
AR(1), ρ =0.95	8	6
LENNA	8	7
BABARA	8	7
GIRL	7	5
AERIAL	8	7
CHURCH	8	7

Table 9 Examples of the optimum word length allocation under the maximum effective word length for AR(1), $\rho = 0.95$.

k	i	j	Ex.1	Ex.2	Ex.3	Ex.4	Ex.5	Ex.6
0	A1	1	3	4	5	6	6	6
1		2	3	4	5	6	6	6
2		3	1	2	3	4	4	4
3	A2	1	7	8	9	10	10	10
4		2	6	7	8	9	9	9
5		3	7	8	9	10	10	10
6	B	1	2	3	4	5	5	5
7		2	2	3	4	5	5	5
8		3	2	3	4	5	5	5
9	C	1	3	4	5	6	6	6
10		2	2	3	4	5	5	5
11		3	3	4	5	6	6	6
12	D	1	1	2	3	4	4	4
13		2	2	3	4	5	5	5
14		3	1	2	3	4	4	4

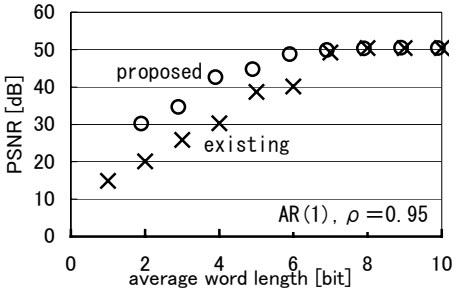


Fig.7 Influence of the rounding operations on PSNR.

5. CONCLUSION

A word length allocation method for multipliers of the integer DCT (Int-DCT) is proposed based on the "SNR sensitivity" which indicates how the word length truncation of each coefficient affects on quality of the decoded image. As a result, two [bit] in average shorter word length is attained under equivalent quality (PSNR) of decoded image. In other word, PSNR is improved by 15 dB. Effect of the rounding operations is also investigated as a matter peculiar to the Int-DCT. Please refer to [9] for theoretical analysis.

This research was supported in part by the Grants-in-Aid for Scientific Research in Japan No.14750284.

REFERENCES

[1] JPEG CD10918-1, Digital compression coding of continuous-tone still images, JPEG-9-R6, Jan.1991.

[2] M. D. Adams, F. Kossentini, "Reversible Integer-to-Integer Wavelet Transform for image Compression: Performance Evaluation and Analysis," IEEE Transactions on Image Processing, vol. 9, pp 1010 - 1024, no. 6, June 2000.

[3] W. Sweldens, "The lifting scheme: a construction of second generation wavelets," Tech. Rep. 1995:6, Industrial Math. Initiative, Dept. of Math., Univ. of South Carolina, 1995.

[4] D. S. Taubman, M. W. Marcellin, "JPEG 2000 - Image compression fundamentals, standards and practice", Kluwer Academic Publishers, 2002.

[5] K. Komatsu, K. Sezaki, 2D Lossless Discrete Cosine Transform, IEEE ICIP 2001, pp.466-469, 2001.

[6] S. Chokchaitam, M. Iwahashi, P. Zavarsky, N. Kambayashi, "A Bit-Rate Adaptive Coding System Based on Lossless DCT", IEICE Trans. on Fundamentals, Vol.E85-A, No.2, pp. 403 -413, Feb. 2002.

[7] S. Fukuma, K. Ohyama, M. Iwahashi, N. Kambayashi, "Lossless 8-point Fast Discrete Cosine Transform using Lossless Hadamard Transform," IEICE Technical Report, DSP99-103, pp.37-44, Oct. 1999.

[8] D. Takago, T. Takebe, "Multispectral Image Compression using Reversible WT and KLT," IEICE Trans. on Fundamentals, vol. J84-A, no.3, pp.298-308, March 2001.

[9] A.sugimori, M.Delgermaa, M.Iwahashi, "Optimum Word Length Allocation for Multipliers of Integer DCT ", IEICE Technical Report, DSP2002-181, pp.123-128, Jan. 2003. (available at : http://tech.nagaokaut.ac.jp/)