THEORETICAL ANALYSIS ON OPTIMUM WORD LENGTH ASSIGNMENT FOR INTEGER DCT

Masahiro IWAHASHI, Kohtaro NAKAGAWA, Somchart CHOKCHAITAM†, Yoshihide TONOMURA

Nagaoka University of Technology, Nagaoka, Niigata, 940-2188, JAPAN
†Faculty of Engineering, Thammasat University, Bangkok, THAILAND

ABSTRACT

Recently JPEG 2000 has adopted the integer wavelet to attain "lossless" coding of images. However existing DCT based "lossy" coding still prevails all over the world. This paper focuses on the integer DCT which can perform "lossless" coding maintaining compatibility to the "lossy" DCT coding. So far, we have optimized word length of each multiplier in the integer DCT considering colored spectrum of input image signal. However it is remained unknown that 1) Is the assigned word length robust to change of input signal? 2) How does it depend on structure of the transform? This report is intended as an investigation of theoretical analysis to answer such questions.

1. INTRODUCTION

Over the past few years, numerous attempts have been made to demonstrate advantages of the integer transforms, which includes rounding operations in the lifting structure. As a result, the next international standard JPEG 2000 adopts the integer wavelet for "lossless" coding of images [1].

On the other hand, the existing JPEG international standard based on the discrete cosine transform (DCT) prevails all over the world [1]. Therefore the integer DCT is becoming an attractive tool to implement "lossless" coding maintaining compatibility to the existing DCT based "lossy" coding [2–5].

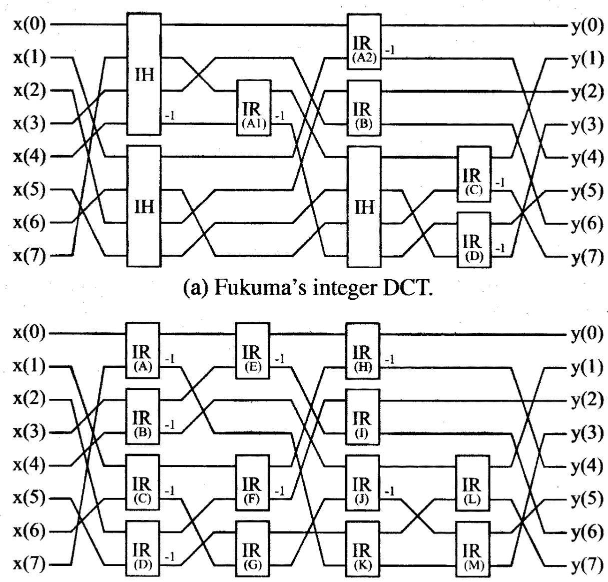
So far, non-separable 2D structuring [3] and the integer Hadamard transform [4] reduced the number of rounding operations and multipliers respectively. We have also optimized word length of each multiplier in the integer DCT considering colored spectrum of input image signal [5]. It is helpful in designing a gate level LSI circuit for low power consumption.

This report is intended as an investigation of theoretical analysis that was lacked in our previous report [5]. To give an actual example, following questions have been unknown up to now. 1) How is the assigned word length robust to various input signals? 2) How does it depend on various algorithms?

2. WORD LENGTH ASSIGNMENT

2.1. Integer DCT

The integer DCT converts integer input vector $x(n)$, n=1,2,$\cdots$,8, into uncorrelated output vector $y(n)$ also expressed with integer. It is possible to achieve effective "lossless" coding applying an entropy coder directly to the output vector $y(n)$. "Lossy" coding is also implemented by inserting quantization between the transform and the entropy coder. Fig.1 (a) illustrates Fukuma's algorithm [4]



(a) Fukuma's integer DCT.

(b) Chen's DCT algorithm.

Fig. 1. Integer DCT algorithms(forward transform).

that is composed of integer Hadamard transform (IH) and integer rotation transform (IR). The IR produces two integer outputs $y(0)$ and $y(1)$ from two integer inputs $x(0)$ and $x(1)$ as

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

where "$m_{1(i)}, m_{2(i)}, m_{3(i)}$" indicate multiplier coefficients in the i-th IR for i $\in \{A1, A2, B, C, D\}$. Note that this processing is actually performed with the rounding operation in the lifting structure [3,4]. As indicated in figure 1 (a), the integer DCT has five IRs and each IR has three multipliers. Therefore the DCT has fifteen multipliers in total. Table 1 summarizes their values. Hereinafter, the coefficients are referred to as h_k for $k = \{0, 1, \cdots, K-1\}$ with $K = 15$.

Fig 1 (b) indicates Chen's DCT algorithm. There are some other algorithms for integer DCT such as [2,3,6,7], however, we focus on these two in figure 1 as examples to demonstrate our theoretical analysis below.

Table 1. Coefficients of IR in the Fukuma's algorithm.

$m_{1(A1)}$	h_0	$1-\sqrt{2}$	$m_{1(C)}$	h_9	$\frac{1-\cos(3\pi/16)}{\sin(3\pi/16)}$
$m_{2(A1)}$	h_1	$\frac{1}{\sqrt{2}}$	$m_{2(C)}$	h_{10}	$-\sin(3\pi/16)$
$m_{3(A1)}$	h_2	$1-\sqrt{2}$	$m_{3(C)}$	h_{11}	$\frac{1-\cos(3\pi/16)}{\sin(3\pi/16)}$
$m_{1(A2)}$	h_3	$1-\sqrt{2}$	$m_{1(D)}$	h_{12}	$\frac{\cos(\pi/16)-1}{\sin(\pi/16)}$
$m_{2(A2)}$	h_4	$\frac{1}{\sqrt{2}}$	$m_{2(D)}$	h_{13}	$\sin(\pi/16)$
$m_{1(A2)}$	h_5	$1-\sqrt{2}$	$m_{2(D)}$	h_{14}	$\frac{\cos(\pi/16)-1}{\sin(\pi/16)}$
$m_{1(B)}$	h_6	$\frac{\sin(\pi/8)-1}{\cos(\pi/8)}$			
$m_{2(B)}$	h_7	$\cos(\pi/8)$			
$m_{3(B)}$	h_8	$\frac{\sin(3\pi/8)-1}{\cos(\pi/8)}$			

2.2. SNR Sensitivity for Colored Signal

Expressing word length of multipliers as short as possible is one of feasible ways to reduce circuit scale at gate level for power conservation of an electric device. For such as ASIC design of LSI, assuming that the total circuit scale is almost proportional to word length of multipliers, it is our concern how to assign word length to each multipliers of the integer DCT under a given averaged word length. It is known by intuition that a more sensitive multiplier needs longer word length. It is also mathematically described as follows.

Firstly, we focus on the situation that the original signal vector $\boldsymbol{X} = [x(0) \cdots x(7)]^T$ in each block is transformed into $\boldsymbol{Y} = [y(0) \cdots y(7)]^T$ by the existing forward DCT with almost infinite word length coefficients, and then $\boldsymbol{Y}$ is transformed by the integer backward transform to produce decoded signal vector $\hat{\boldsymbol{X}} = [\hat{x}(0) \cdots \hat{x}(7)]^T$. Here, only one coefficient $h_k \in \{0, 1, \cdots, K-1\}$ in the **backward** transform is truncated into W_k bit binary expression. Note that the rounding operations are excluded from this analysis since it is assumed to be negligible. In this case, the decoded signal contains the error due to finite word length expression of one coefficient h_k. Defining the error variance on the decoded signal by

$$\Delta\sigma_k^2 = \frac{1}{8} E\left[\left(\hat{\boldsymbol{X}} - \boldsymbol{X}\right)^T \left(\hat{\boldsymbol{X}} - \boldsymbol{X}\right)\right] \tag{2}$$

where $E[\,]$ denotes ensemble average over all the blocks, the SNR sensitivity [5] is defined by

$$S_k = \frac{\Delta\sigma_k}{\Delta h_k} \tag{3}$$

where Δh_k is a difference between original coefficient value h_k and its truncated version $\hat{h}_k$.

2.3. Optimum Word Length Assignment

Since $\Delta h_k \propto 2^{-W_k}$, optimum word length assignment is performed by minimizing total error energy $\sigma_{total}^2 = \sum_{k=0}^{K-1} \sigma_k^2$ under a given average word length $\overline{W} = \sum_{k=0}^{K-1} W_k/K$. This problem is solved by linear equations with the Lagrange's method [5] as

$$W_k - \overline{W} = \log_2 S_k - \log_2 \overline{S} \tag{4}$$

Table 2. Assigned word length W_k for AR(1), $\rho = 0.95, \overline{W} = 3$.

k	W_k	k	W_k
0	3	9	3
1	3	10	2
2	2	11	4
3	5	12	2
4	4	13	2
5	5	14	2
6	2		
7	3		
8	3		

where

$$\overline{W} = \frac{1}{K}\sum_{k=0}^{K-1} W_k, \quad \overline{S} = \prod_{k=0}^{K-1} \sqrt[K]{S_k}. \tag{5}$$

For AR(1) input model signal with $\rho = 0.95$ for example, the assignment results are given by table 2.

3. THEORETICAL ANALYSIS

3.1. Propagation of the Error Signal

For the purpose of theoretical analysis in this report, firstly, we mathematically describe how the error signal Δh_k is propagated trough the integer DCT, for example in figure 1 (a). Denoting the forward DCT as $\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X}$ and backward DCT as $\boldsymbol{X} = \boldsymbol{G}\boldsymbol{Y}$, according to fig.1(a), matrices are described by

$$\boldsymbol{H} = \boldsymbol{P}_4 \begin{bmatrix} \boldsymbol{I}_4 & \boldsymbol{O}_4 \\ \boldsymbol{O}_4 & \begin{bmatrix} \boldsymbol{H}_C & \boldsymbol{O}_2 \\ \boldsymbol{O}_4 & \boldsymbol{H}_D \end{bmatrix} \end{bmatrix} \boldsymbol{P}_3 \begin{bmatrix} \begin{bmatrix} \boldsymbol{H}_{A2} & \boldsymbol{O}_2 \\ \boldsymbol{O}_4 & \boldsymbol{H}_B \end{bmatrix} & \boldsymbol{O}_4 \\ \boldsymbol{O}_4 & \boldsymbol{H}_{IH} \end{bmatrix} \cdot \boldsymbol{P}_2 \begin{bmatrix} \begin{bmatrix} \boldsymbol{I}_2 & \boldsymbol{O}_2 \\ \boldsymbol{O}_2 & \boldsymbol{H}_{A1} \end{bmatrix} & \boldsymbol{O}_4 \\ \boldsymbol{O}_4 & \boldsymbol{I}_4 \end{bmatrix} \boldsymbol{P}_1 \begin{bmatrix} \boldsymbol{H}_{IH} & \boldsymbol{O}_2 \\ \boldsymbol{O}_2 & \boldsymbol{H}_{IH} \end{bmatrix} \boldsymbol{P}_0, \tag{6}$$

$$\boldsymbol{G} = \boldsymbol{P}_0^{-1} \begin{bmatrix} \boldsymbol{G}_{IH} & \boldsymbol{O}_2 \\ \boldsymbol{O}_2 & \boldsymbol{G}_{IH} \end{bmatrix} \boldsymbol{P}_1^{-1} \begin{bmatrix} \begin{bmatrix} \boldsymbol{I}_2 & \boldsymbol{O}_2 \\ \boldsymbol{O}_2 & \boldsymbol{G}_{A1} \end{bmatrix} & \boldsymbol{O}_4 \\ \boldsymbol{O}_4 & \boldsymbol{I}_4 \end{bmatrix} \boldsymbol{P}_2^{-1} \cdot \begin{bmatrix} \begin{bmatrix} \boldsymbol{G}_{A2} & \boldsymbol{O}_2 \\ \boldsymbol{O}_2 & \boldsymbol{G}_B \end{bmatrix} & \boldsymbol{O}_4 \\ \boldsymbol{O}_4 & \boldsymbol{I}_4 \end{bmatrix} \boldsymbol{P}_3^{-1} \begin{bmatrix} \boldsymbol{I}_4 & \boldsymbol{O}_4 \\ \boldsymbol{O}_4 & \begin{bmatrix} \boldsymbol{G}_C & \boldsymbol{O}_2 \\ \boldsymbol{O}_4 & \boldsymbol{G}_D \end{bmatrix} \end{bmatrix} \boldsymbol{P}_4^{-1}. \tag{7}$$

$\boldsymbol{H}_i$ for $i \in \{\text{A1,A2,B,C,D}\}$ and $\boldsymbol{H}_{IH}$ indicate the integer rotation transform (IR) and the integer Hadamard transform (IH) respectively. $\boldsymbol{I}_i$ and $\boldsymbol{O}_i$ for $i \in \{2, 4\}$ denote unity square matrix and zero square matrix of size i respectively. $\boldsymbol{P}_0 \sim \boldsymbol{P}_4$ are permutation matrices. Note that $\boldsymbol{G}_i = \boldsymbol{H}_i^{-1}$ for $i \in \{A1, A2, B, C, D, IH\}$.

Notating the inverse DCT with a truncated coefficient $\hat{h}_k = h_k + \Delta h_k$, $k \in \{0, 1, \cdots, K-1\}$ by $\hat{\boldsymbol{G}}$, error signal in the decoded signal $\hat{\boldsymbol{X}}$ becomes

$$\Delta \boldsymbol{X}_k = \hat{\boldsymbol{X}} - \boldsymbol{X} = \left(\hat{\boldsymbol{G}} - \boldsymbol{G}\right) \boldsymbol{H}\boldsymbol{X} \tag{8}$$

with an approximation

$$\hat{\boldsymbol{G}} - \boldsymbol{G} \approx \frac{\partial \boldsymbol{G}}{\partial h_k} \Delta h_k. \tag{9}$$

Table 3. Constant value c_{ki} for the Fukuma's integer DCT.

k	c_{k0}	c_{k1}	c_{k2}	c_{k3}	c_{k4}	c_{k5}	c_{k6}	c_{k7}
0	0.125	−0.031	−0.000	0.047	−0.031	−0.078	−0.000	−0.031
1	0.172	0.063	−0.030	−0.119	−0.000	0.003	−0.030	−0.058
2	0.125	−0.141	0.031	−0.031	0.031	0.031	0.031	−0.016
3	0.125	0.094	0.000	0.093	0.125	0.031	0.000	0.031
4	0.172	0.007	−0.182	0.068	0.172	−0.018	−0.061	0.043
5	0.125	0.219	0.188	0.156	0.125	0.094	0.063	0.031
6	0.125	−0.141	0.031	0.047	−0.125	0.078	−0.031	0.015
7	0.262	−0.189	0.035	0.023	−0.262	0.118	−0.035	0.048
8	0.125	0.123	−0.044	−0.142	−0.125	−0.035	0.044	0.053
9	0.125	0.046	0.110	−0.053	−0.031	−0.089	−0.066	−0.042
10	0.149	−0.238	0.215	−0.149	0.081	−0.054	0.005	−0.009
11	0.125	0.171	0.075	−0.016	−0.082	−0.111	−0.102	−0.060
12	0.125	−0.109	−0.110	0.146	−0.031	−0.067	0.066	−0.020
13	0.127	0.052	−0.159	−0.131	0.034	0.100	0.021	−0.044
14	0.125	−0.115	−0.109	0.157	−0.034	−0.074	0.068	−0.019

3.2. Robustness to input signals

Next, we analyze robustness of the optimally assigned word length W_k in equation (4) against variety of input signals. Concretely, assuming AR(1) model as an input signal, we decsribe W_k in term of correlation ρ of the model given by

$$x(n) = w(n) + \rho \cdot x(n-1) \tag{10}$$

where $w(n)$ denotes a white noise. In this case, covariance matrix $\boldsymbol{C}_{xx}$ of the input signal $\boldsymbol{X}$ is expressed with ρ by

$$\boldsymbol{C}_{xx} = \boldsymbol{E}\left[\boldsymbol{X}\boldsymbol{X}^T\right] = \sigma_x^2 \boldsymbol{R}_{xx}(\rho) \tag{11}$$

where

$$\boldsymbol{R}_{xx}(\rho) = \begin{bmatrix} 1 & \rho & \cdots & \rho^7 \\ \rho & 1 & \cdots & \rho^6 \\ \vdots & \vdots & \ddots & \vdots \\ \rho^7 & \rho^6 & \cdots & 1 \end{bmatrix}. \tag{12}$$

On the other hand, $\Delta\sigma_k^2$ in equation (2) is equivalent to

$$\Delta\sigma_k^2 = \frac{1}{8}E\left[tr\left[\left(\hat{\boldsymbol{X}} - \boldsymbol{X}\right)\left(\hat{\boldsymbol{X}} - \boldsymbol{X}\right)^T\right]\right] \tag{13}$$

where $tr\,[\,]$ denotes a trace, which is a sum of diagonal components, of a matrix. Therefore, substituting equations (8), (9) and (11), it becomes

$$\begin{aligned} \Delta\sigma_k^2 &= \frac{\Delta h_k^2}{8}\text{tr}\left[\frac{\partial \boldsymbol{G}}{\partial h_k}\boldsymbol{H}E\left[\boldsymbol{X}\boldsymbol{X}^T\right]\boldsymbol{H}^T\frac{\partial \boldsymbol{G}^T}{\partial h_k}\right] \\ &= \frac{\sigma_x^2 \Delta h_k^2}{8}\text{tr}\left[\frac{\partial \boldsymbol{G}}{\partial h_k}\boldsymbol{H}\boldsymbol{R}_{xx}(\rho)\boldsymbol{H}^T\frac{\partial \boldsymbol{G}^T}{\partial h_k}\right] \end{aligned} \tag{14}$$

As a result, the error variance is described as a function of ρ. More explicitly,

$$\Delta\sigma_k^2 = \sigma_x^2 \Delta h_k^2 \left(\sum_{i=0}^{7} c_{ki} \cdot \rho^i\right). \tag{15}$$

Coefficients c_{ki} are given by comparing equations (14) and (15). An example is given in table 3 for the Fukuma's integer DCT. Therefore, the sensitivity S_k and the optimum word length assignment W_k are theoretically evaluated with correlation ρ of the input signal, replacing equation (4) by

$$W_k - \overline{W} = \frac{1}{2}\log_2 \frac{\sum_{i=0}^{7} c_{ki} \cdot \rho^i}{\Pi_{l=0}^{K-1}\left(\sum_{j=0}^{7} c_{lj} \cdot \rho^j\right)^{1/K}}. \tag{16}$$

This analysis is similarly applicable to any other DCT algorithms such as Chen's algorithm in figure 1(b).

3.3. Dependence on algorithms

Finally, we theoretically investigate dependence of the word length assignment on different algorithms. When all the K (=15 in this report) coefficients are truncted, total error variance is given as a summation of equation (2) by

$$\Delta\sigma_{total}^2 = \sum_{k=0}^{K-1} \Delta\sigma_k^2. \tag{17}$$

Substituting equation (3), it is described with the sensitivity as

$$\Delta\sigma_{total}^2 = \sum_{k=0}^{K-1} S_k^2 \Delta h_k^2. \tag{18}$$

Since Δh_k has a uniform histogram in the range of

$$0 \leq \Delta h_k \leq 2^{-(W_k+1)}, \tag{19}$$

its variance is considered to be $\frac{2^{-2W_k}}{12}$. Also from equation (4), we have

$$\Delta\sigma_{total}^2 = \frac{K\overline{S}^2}{12}2^{-2\overline{W}}. \tag{20}$$

As a result, we can estimate dependence of the assignment on the different algorithms by the equation below, especially in the term of $\overline{S}$ by

$$\begin{aligned} \text{PSNR} &= 10\log_{10}\frac{255^2}{\Delta\sigma_{total}^2} \\ &= -20\log_{10}\overline{S} - 20\log_{10}\sqrt{K} \\ &\quad + (20log_{10}2)\,\overline{W} + 20log_{10}\left(255\sqrt{12}\right). \end{aligned} \tag{21}$$

Table 4. Validity of the theoretical evaluation of the sensitivity. Theoretical values are indicated in ( ).

k	S_k	Error[%]	k	S_k	Error[%]
0	5.0(5.1)	1.1	9	6.8(6.7)	0.3
1	6.1(6.2)	1.6	10	3.2(3.2)	0.3
2	2.4(2.4)	0.0	11	8.2(8.1)	0.2
3	20.2(20.3)	0.1	12	2.2(2.2)	1.4
4	13.3(13.0)	1.3	13	3.1(3.1)	0.7
5	28.8(28.6)	0.8	14	2.1(2.1)	2.0
6	2.4(2.4)	0.1			
7	4.3(4.3)	1.5			
8	4.5(4.5)	0.7			

Table 5. Applicable range of ρ of the input signal to the optimum word length desinged for AR(1) model with $\rho = 0.95$.

k	range	k	range
0	$0.000 < \rho < 0.998$	9	$0.000 < \rho < 1.000$
1	$0.000 < \rho < 1.000$	10	$0.682 < \rho < 1.000$
2	$0.506 < \rho < 0.995$	11	$0.574 < \rho < 0.985$
3	$0.900 < \rho < 0.979$	12	$0.487 < \rho < 0.982$
4	$0.800 < \rho < 0.963$	13	$0.848 < \rho < 1.000$
5	$0.803 < \rho < 0.954$	14	$0.769 < \rho < 0.978$
6	$0.544 < \rho < 0.995$		
7	$0.000 < \rho < 0.983$		
8	$0.000 < \rho < 0.990$		

4. EXPERIMENTAL RESULTS

In the previous section, we theoretically derived estimation formula for dependence of the assigned word length on "input signals" and "algorithms". These are base on expressing the sensitivity S_k with input signal's correlation ρ. Table 4 inspects validity of the theoretical evaluation with table 3. Its accuracy (difference) was confirmed to be less that 2.0 %.

4.1. Robustness to Input Signals

As an application of the theoretical analysis in **3.2**, we investigated applicable range of ρ of the input signal to the optimum word length designed for AR(1) model with $\rho = 0.95$ and 65536 samples. Figure 2 illustrates correlation of the input signal versus optimum word length. Table 5 summarizes applicable range of ρ of the input signal which was given by figure 2. The result of the experiment gave good agreement with the value that had been obtained by theoretical calculation.

4.2. Dependence on algorithms

As an application of the theoretical analysis in **3.3**, we also investigated dependence of the optimum word length designed for AR(1) model with $\rho = 0.95$ and 65536 samples on different algorithms. Figure 3 indicates experimental results as well as theoretical evaluation results by equation (21). The result of the experiment was also in beautiful agreement with the value that had been obtained by theoretical calculation.

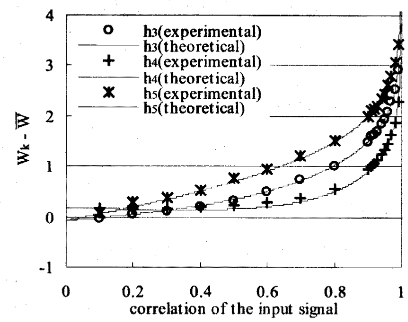


Fig. 2. Dependence of the assigned word length on input signals.

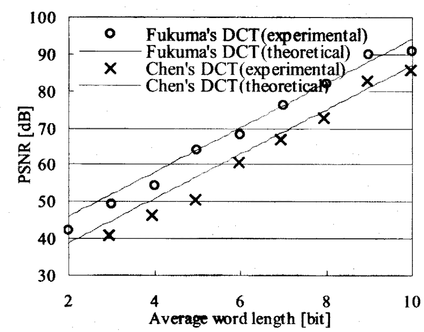


Fig. 3. Dependence of the assigned word length on input signals.

5. CONCLUSION

This report investigated theoretical analysis that was lacked in our previous report on optimum word length assignment for multipliers in the integer DCT. To answer to the following questions: 1) How is the assigned word length robust to various input signals? 2) How does it depend on various algorithms?, we derived theoretical estimation formula as explicit functions of correlation of the input image signal. Its validity was inspected and accuracy was confirmed.

6. REFERENCES

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

[2] Y. J. Chen, S. Oraintara, T. Nguyen, "Video compression using integer DCT", IEEE ICIP, vol.2, pp.844-847, 2000.

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

[4] 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.

[5] M.Iwahashi, O.Nishida, S.Chokchaitam, N.Kanbayashi, "Optimum Word Length Allocation for Multipliers of Integer DCT", IEEE ISCAS 2003, vol.II, pp.400-403, May 2003.

[6] W. Philips, "The lossless DCT for combined lossy / lossless image coding", IEEE ICIP, vol.3, pp.871-875, 1998.

[7] G. Charith, K. Abhayaratne, "N-Point Discrete Cosine Transforms that Map Integers To Integers for Lossless Image / Video Coding", Proc. Picture Coding Symposium (PCS), pp 417-422, April 2003.