A Blur of Reconstructed Images Caused by Orthogonal Transforms : DCT & DWT and an Evaluation of Image Fidelity

M. IIZUKA* M. KITAMURA** Y. NAKASHIMA***

The simplified representation techniques in quasi-color are proposed in order to intuitively estimate the visual appearances and differences between similar images, i. e., the blur of reconstructed images caused by orthogonal transforms : DCT and DWT. One of the available role of diffemence images is discussed together with the statistical measures : RMSE and/or PSNR for the evaluation of image fidelity.

1. Introduction

A discrete cosine transform (DCT) algorithm has been adopted for the digital compression of still images, and special chip sets are being produced for their implementation in JPEG standard [1].

Wavelets, filter banks, and multiresolution signal analysis have been used independently in the fields of applied mathematics and/or digital signal and image processing. Recently, modern technology converged to form a unified single theory. A discrete wavelet transform (DWT) is a king of the orthogonal transform techniques such as the discrete Fourier transform (DFT) and the DCT. Many papers and comprehensive reviews have been published concerning a wavelet based image processing and its applications [2]~[9].

In the case of DCT, an idea of adaptive spaptive spatial filtering is used for the quantization of DCT transform coefficients in the spatial frequency domain. In the case of DWT, two types of D_2 and D_4 filter coefficients for carrying out wavelet transforms are treated, and the spatial filtering influence between the two is compared in terms of decomposition and/or reconstruction of sub-image data.

In order to intuitively estimate the visual appearance and distinguishable difference between similar images, the simplified representation techniques in quasi color are used for a subtle demonstration of reconstructed images and difference images together with the statistical measures : RMSE and/or PSNR.

2. DCT and JPEG Technology for Still Image Data Compression

JPEG is an acronym for "Joint Photographic Experts Group". This group was established for the purpose of developing a standard technology for color image compression. Continuous tone color displays with high quality are expected for most personal computing systems.

The main purpose of image compression is to represent the reconstructed images by means of

*Professor, Tokyo Institute of Polytechnics
**Assistant, Research & Education Center of Software Engineering, Tokyo Institute of Polytechnics
***Associate Professor, Toyama University
Received Aug. 30, 1996

less coding data in order to save storage costs or transmission time and costs. The most effective compression techniques are achieved by approximating an original image rather than reproducing it exactly.

The discrete cosine transform (DCT) was first applied to image compression in Ahmed, Natarajan, and Rao's pioneering work [10], in which they pointed out that this particular transform was very close to the Karhunen-Loēve transform (KLT), i. e., a kind of transform that produces uncorrelated transformed coefficients [11]. Decorrelation of the coefficients is very important for image data compression, because each coefficient can then be treated independently without the loss of compression efficiency. Another important aspect of DCT is the ability to quantize the DCT coefficients using an idea of spatial filtering processing or visually-weighted quantization values which are recommended as the luminance quantization table and the chrominance quantization table in JPEG.

The 2-D FDCT and IDCT can be constructed from products of the terms of a horizontal 1-D DCT and a vertical 1-D DCT.

FDCT:

$$F(u, v) = (2/N) \cdot C(u) \cdot C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \{\cos[(2x+1) \cdot u\pi/2N] \cdot \cos[(2y+1) \cdot v\pi/2N]\} \quad (1)$$

IDCT:

$$f(x, y) = (2/N) \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u) \cdot C(v) \cdot F(u, v) \{\cos[(2x+1) \cdot u\pi/2N] \cdot \cos[(2y+1) \cdot v\pi/2N]\} \quad (2)$$

where, x, y : sampling point in x-y spatial coordinates

u, v : spatial frequency

f(x, y): 2-D sample or image data

F(u, v): 2-D DCT transform coefficients

$C(p) = 1/\sqrt{2}$ for $p=0$

$= 1$ for $p>0$

N : number of sampling point in sub block at x or y direction

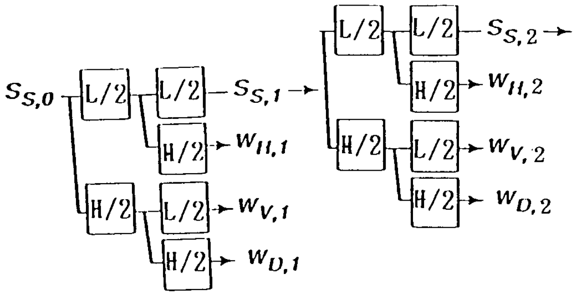
Note that a sub block division such as 8×8 and 16×16 is used in order to decrease the computation time in the case of 2-D DCT. Each DCT transform coefficient may be computed as the numerical value of real number.

3. Discrete Wavelet Transform for Signal and Image Processing

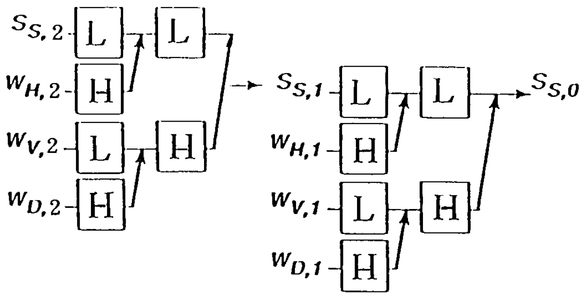
Daubechies [12] constructed compactly supported orthonormal wavelets based on iterations of discrete filters. Newland [13] introduced a wavelet program, i. e., an interactive software package for scientific and engineering numerical computation which uses matrix operation.

Fig. 1 shows a successive process of wavelet decomposition and reconstruction of images. The boldface letters such as 「L」 and 「H」 stand for the use of the low pass and high pass filters, respectively. The first subscript letter means the specified domains of smoothing, horizontal, vertical and diagonal sub-images, respectively. The second subscript means the specified resolution level for wavelet transforms, on the assumption that zero means the start resolution level of an original image (corresponding to the smoothing image). A better approximation of image data may be obtained by adding the details corresponding to Wj to the approximate signal Sj at the specified resolution level j. The approximate signal Sj stands for the smoothing image, and the details contain the three different sub-images: horizontal, vertical, and diagonal images.

An expression for the decomposition and its



(a) Wavelet decomposition from level : 0 to 1, and from level : 1 to 2



(b) Wavelet reconstruction from level : 2 to 1, and from level : 1 to 0

Fig. 1 Decomposition and its reconstruction of digital images by means of spatial low/high pass filters

reconstruction techniques of digital images using the matrix operations may be formulated in connection with Fig. 1. The four different sub-images at the resolution level : $j+1$ may be transformed from the smoothing image at the resolution level : j.

Note that the constant parameter 1/2 in the spatial low pass and high pass filters, i.e., L and H is contained in the case of Fig. 1 (a), because two types of spatial filters relevant to the united square matrix must be normalized in the matrix operation for the discrete wavelet transform [14] .

On the other hand, only one smoothing image, i. e., reconstructed image at the resolution level $j-1$ may be synthesized from the sum of four different sub-images at the resolution level j.

(a) Decomposition from level j to $j+1$

$$\begin{aligned} S_{s,j+1} &= (1/4) \cdot L_j \cdot S_{s,j} \cdot (L_j)^t \\ W_{H,j+1} &= (1/4) \cdot L_j \cdot S_{s,j} \cdot (H_j)^t \\ W_{V,j+1} &= (1/4) \cdot H_j \cdot S_{s,j} \cdot (L_j)^t \\ W_{D,j+1} &= (1/4) \cdot H_j \cdot S_{s,j} \cdot (H_j)^t \end{aligned} \qquad (3)$$

(b) Reconstruction from level j to $j-1$

$$\begin{aligned} S_{s,j-1} = {} & (L_j)^t \cdot S_{s,j} \cdot L_j + (H_j)^t \cdot W_{H,j} \cdot L_j \\ & + (L_j)^t \cdot W_{V,j} \cdot H_j + (H_j)^t \cdot W_{D,j} \cdot H_j \end{aligned} \qquad (4)$$

where, $S_{s,j}$: smoothing (or original) image at resolution level j

$W_{H,j}$; $W_{V,j}$; $W_{D,j}$: horizontal, vertical, and diagonal image at resolution level j, respectively

L_j ; H_j : spatial low-pass and high-pass filter composed of $2^{n-1} \times 2^n$ components, respectively

n : integer ($n \geqq 1$)

t : transposed symbol for matrix

A united square matrix may be composed of two types of filter coefficient, i. e., a combination of the spatial low pass and high pass filter coefficients. The wavelet transforms as well as Fourier transform and cosine transform called the orthogonal transforms may be recursively carried out using an idea of the united square matrix[14].

A pair of two-dimensional discrete wavelet transform and its inverse transform may be performed directly by means of the matrix operations, on the assumption that an original image data has the nth power of 2.

Image decomposition with four small sub-images in matrix expression

$$\begin{aligned} W &= \{(1/2) \times T \times F\} \cdot \{(1/2) \times T^t\} \\ &= 0.25 \times \{T \times F \times T^t\} \end{aligned} \qquad (5)$$

Inage reconstruction or synthesis in matrix expression

$$F = T^t \times W \times T \qquad (6)$$

where, F : original image

W : wavelet image (a set of four small sub-images)

T : united square matrix composed of a combination of low-pass and high-pass filters

Note that the two same constant parameters : 1/2 are necessary for the normalization of the united square matrix : T in the case of image decomposition as described in Eq.(5).

4. Reconstructed Similar Images and Fidelity Evaluation of Image Quality

Fig. 2 shows a simple framework for comparing the difference between an original image data and a lossy modified and reconstructed image data. An important role of image transformation or decomposition is to reduce an extra-component or a dynamic range of the signal (or data) in order to eliminate redundant information that can be coded or compressed more effectively.

Many statistical and quantitative measures are proposed in order to evaluate the image fidelity between the two similar images [15]. It is important to note that a lower RMSE (root mean square error) or equivalently a higher PSNR (peak signal-to-noise ratio) does not necessarily imply a higher sabjective reconstructed inage quality. The error metrics defined as RMSE and PSNR do not always

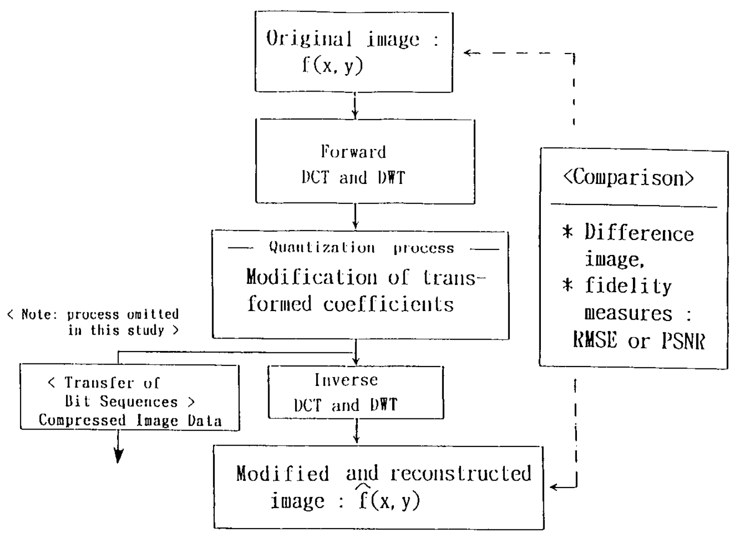


Fig. 2 A simple framework for comparing the difference between original image data and modified & reconstructed image data

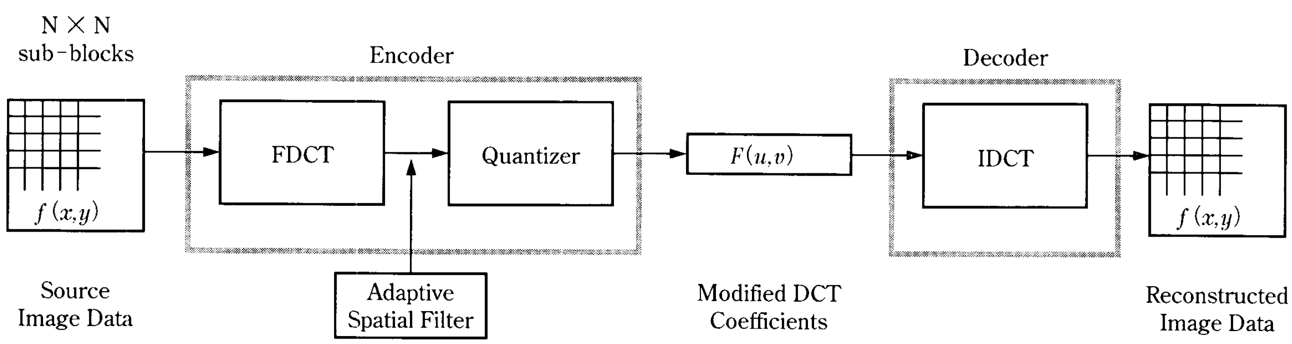


Fig. 3 Simplified FDCT and IDCT process by means of adaptive spatial filter

correlate well with perceived image quality. Denoting the original N×N image by f(x, y) and the compressed-reconstructed image by $\hat{f}$ (x, y), RMSE is given by the following expression.

$$RMSE = (1/N)\sqrt{\sum_x \sum_y [f(x,y) - \hat{f}(x,y)]^2} \quad (7)$$

The related measure of PSNR in dB is computed for a 4-bit (0−15 gray levels) image as follows.

$$PSNR = 20 \times \log_{10}(15/RMSE) \text{ [dB]} \quad (8)$$

Fig. 3 shows a simplified computation process for FDCT and IDCT by using an adaptive spatial filter. The spatial filtering process is used in place of the quantization table which is recommended by JPEG. Entropy coding techniques for preparing compressed code data are omitted in this study.

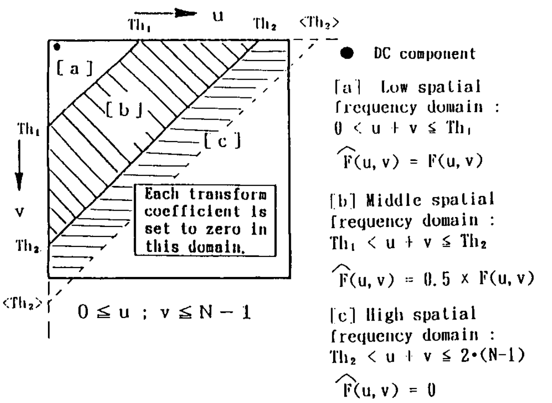


Fig. 4 Spatial frequency domain and modified DCT coefficients

Fig. 4 shows three kinds of spatial frequency domains and the relation between DCT coefficients: F(u, v) and modified DCT coefficients: $\hat{F}$(u, v). In the high frequency domain, the modified DCT coefficients are approximately replaced by the zero value in order to delete the image data for the sake of convenience and in order to discuss the visual appearance effect of reconstructed images.

Fig. 5 shows an oiginal image in quasi color, the power distribution of DCT tranform coefficients, and reconstructed images by using 256 sub-blocks. Each sub block is composed of 8×8 pixels. Each pattern of DCT coefficients in sub blocks is different from each other. The upper left corner in each sub-block stands for the DC component, and the numerical value of DC component is much larger in contrast with that of many AC components. Note that the negative value of DCT transform coefficients is replaced by the zero value when demonstrating visually the aspect of power distribution. The visual aspect of reconstructed images is almost similar to the original image, although the statistical values of image fidelity: RMSE and PSNR are a little different each other.

Fig. 6 shows the reconstructed images in the use of D_2 and D_4 filter coefficients. It is difficult to discern the visual difference between two similar reconstructed images at the same resolution level. But the visual appearance caused by the resolution level: 1 and the resolution level: 2 or 3 is very clear each other at a glance. As a result, the role of D_2 filter coefficients is equivalent to that of D_4 filter coefficients in the case of the wavelet based digital image processing. The display results in Fig. 6(a) are the almost same to those of Fig. 5 (c) and (d) intuitively. Note that the idea of the adaptive spatial filter is used with a view to deleting a part of the high frequency component of DCT transform coefficients in Fig. 5.

Fig. 7 demostrates the results of difference image: ⟨original image - reconstructed image⟩ displayed in quasi colors in the case of two types of D_2 filter coefficients. The spatial high pass filter is composed of [1, −1] and [−1, 1], on condition that the spatial low-pass filter is the same each other. When the following condi-

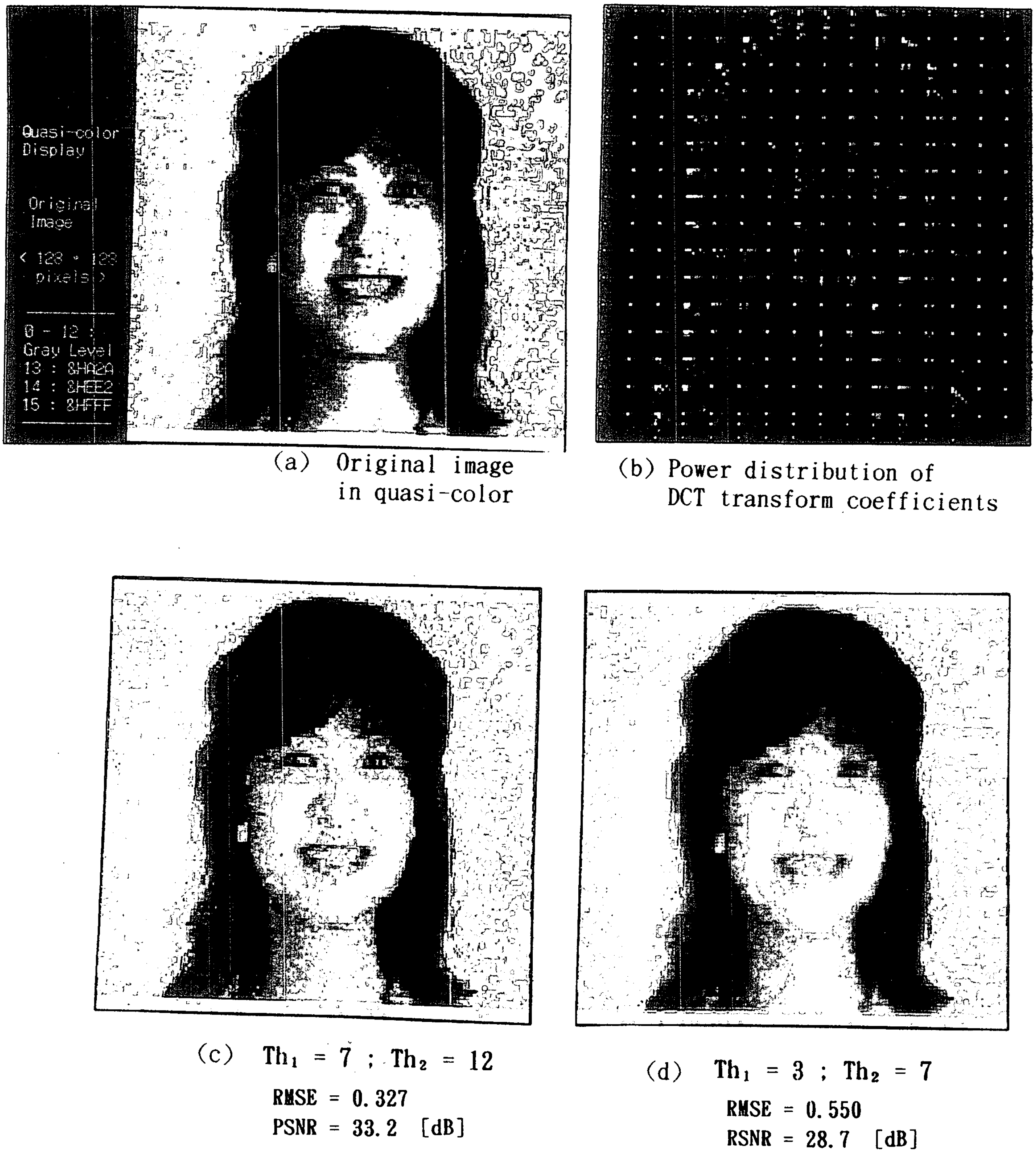


Fig. 5 Original image, DCT transform coefficients, and reconstructed images by using 16×16 sub-blocks in the case of DCT

tion: $-1 \leqq$ D. I. $=$ O. I. $-$ R. I. $\leqq 1$ holds, yellow color is demonstrated as the background color in order to discriminate two similar images intuitively. The conditions: D. I. < -1 and D. I. > 1 hold, red and blue colors are used, respectively. As a result. a kind of the outline of the original image is stressed, because there exists a subtle difference between the two images.

Fig. 8 shows the estimation of image fidelity: RMSE and PNSR calculated using two types of D_2 and D_4 filter coefficients in the case of 128×128 pixels. As the resolution level of wavelet transforms increases, the estimation of image fidelity: RMSE and PSNR becomes

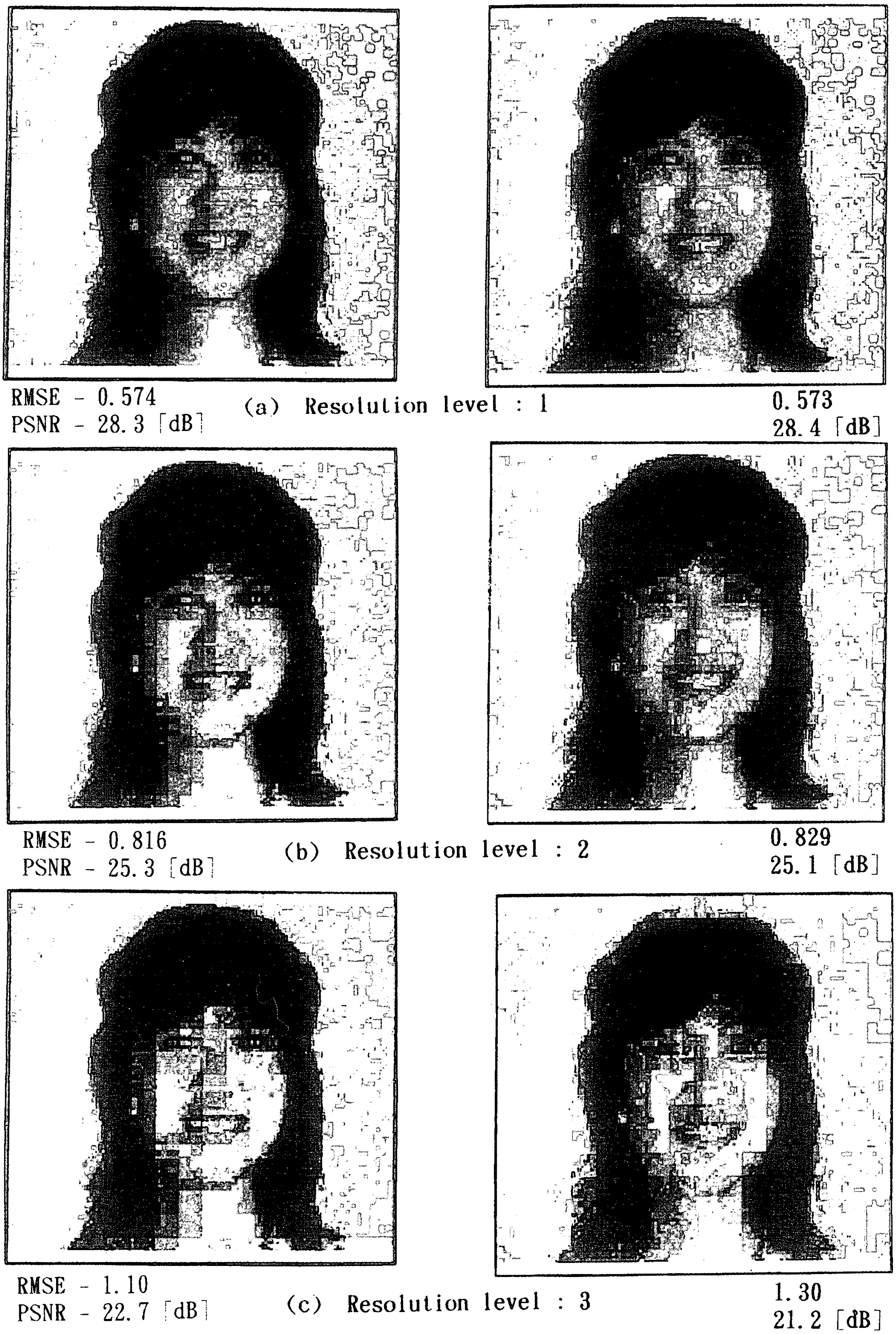


Fig. 6 Reconstructed images in the case of DWT
Left : D_2 wavelet filter coefficient
Right : D_4 wavelet filter coefficient

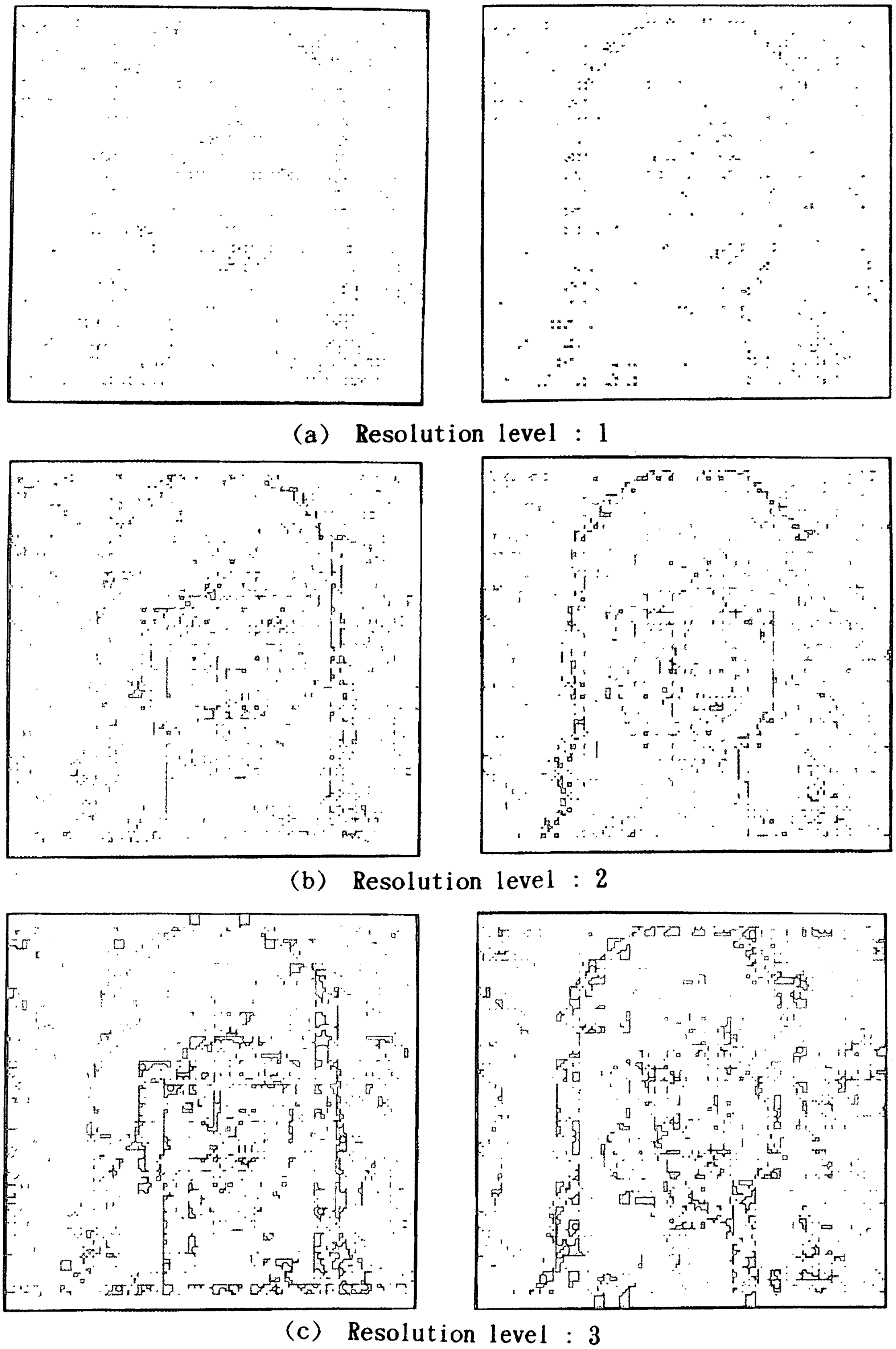


Fig. 7 Difference image in quasi-color in the case of D_2 filter coefficients,
Left: Filter Cos. [1, 1] & [1, −1]; Right: Filter Cos. [1, 1] & [−1, 1]

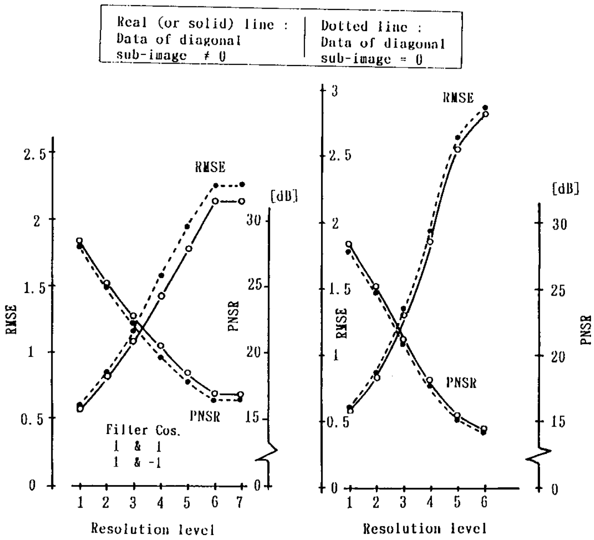


Fig. 8 Image fidelity : RMSE and PNSR calculated using D_2 and D_4 filtec coefficients in the case of 128×128 pixcls

clear. A certain blur and distortion of reconstructed images becomes distinguishable in total in the use of D_4 filter coefficients, because many neighboring image data, which do not directly relative to the main sampling points each other, are used for calculating the expansion coefficients for wavelet transform in conjunction with the united square matrix.

Fig. 9 shows the numerical values of statistical measures : RMSE and PSNR by deleling the wavelet expansion coefficients in three sub-images, i. e., diagonal, horizontal, and vertical image data in comparison with the use of the wavelet expansion coefficients in four sub-images at each resolution level in the case of two types of D_2 filter coefficients. As for the left results, the combination of D_2 filter coefficients : [1, 1] & [1, −1] is used. There is no difference between the two.

5. Conclusions

Digital image processing was carried out by applying a notion of united square matrix in the case of the two-dimensional discrete wavelet transforms. To sum up, the main conclusions are as follows.

(1) In contrast with the D_4 filter coefficients, the D_2 filter coefficients are very suitable for the simplified compression and/or reconstruction of digital images.

(2) The DWT may be effectively used in place of the DCT in the various digital image processing fields for the applications.

References

[1] W. B. Pennebaker and J. L. Mitchell : "JPEG Still Image Data Compression Standard", Van Nostrand Reinhold (1993).
G. K. Wallace : The JPEG still picture compression standard, Commun. ACM, Vol. 34 (1991) pp. 31-44.

[2] P. J. Burt and E. H. Adelson : The Laplacian pyramid as a compact image code, IEEE Trans. Commun., Vol. 31, No. 4 (1983) pp. 532-540.

[3] S. G. Mallat : A Theory for Multiresolution Signal Decomposition : The Wavelet Represetation, IEEE Trans on Pattern Analysis and Machine Intelligence, Vol. 11, No. 7 (1989) pp. 674-693.

[4] O. Rioul and M. Vetterli : Wavelest and Signal Processing, IEEE SP Magazine, Vol. 8, (1991) pp. 14-38.

[5] M. Vetterli and C. Herley : Wavelets and Filter Bank : Theory and Design, IEEE Trans. on Signal Processing, Vol. 40, No. 9 (1992) pp. 2207-2232.

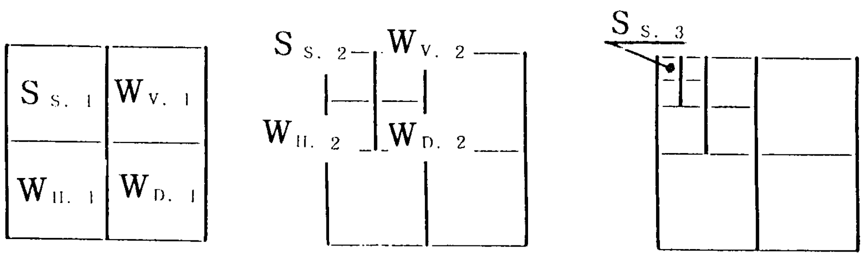
[6] M. Antonini, M. Barland, P. Marhieu, and I. Daubechies : Image Coding Using Wavelet Transform, IEEE Trans. on Image Processing, Vol. 1, No. 2 (1992) pp. 205-220.

[7] E. JT. Stollnitz, T. D. DeRose, and D. H. Salesin : Wavelets for Computer Graphics, A Primer-Part 1-, IEEE Computer Graphics and Applications, Vol. 15, No. 3 (1995) pp. 76-84 ; -Part 2-, Vol. 15, No. 4 (1955) pp. 75-85.

[8] R. J. Clarke : "Digital Compression of Still Images and Video", Academic Press (1995).

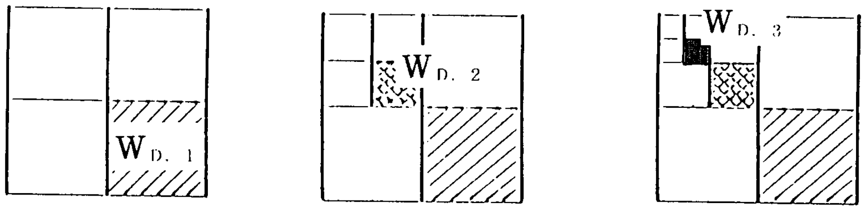
[9] G. W. Wornell : "Signal Processing with Fractals : A Wavelet-Based Approach", Prentice Hall PTR (1996)

[10] N. Ahmed, T. Natarajan, and K. R. Rao : Dis-



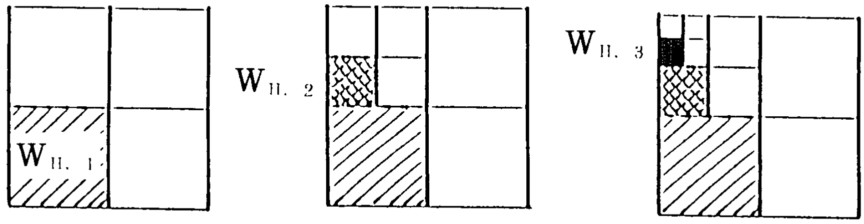
RMSE = 0.574 ; 0.573 0.816 ; 0.830 1.10 ; 1.11
PSNR = 28.3 ; 28.4 [dB] 25.3 ; 25.1 [dB] 22.7 ; 22.6 [dB]

(a) Use of all sub-images



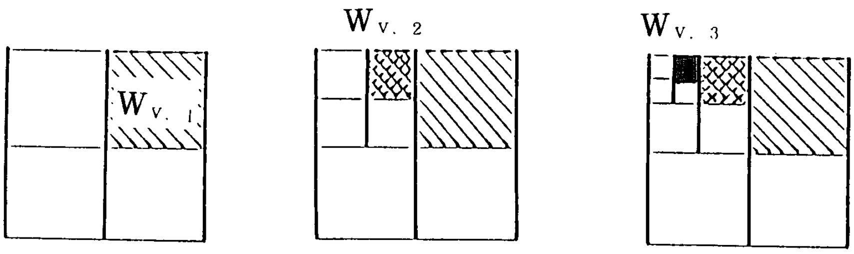
RMSE = 0.609 ; 0.608 0.861 ; 0.874 1.17 ; 1.18
PSNR = 27.8 ; 27.8 [dB] 24.8 ; 24.7 [dB] 22.2 ; 22.1 [dB]

(b) Deletion of all diagonal sub-images



RMSE = 0.705 ; 0.703 1.03 ; 1.03 1.81 ; 1.81
PSNR = 26.6 ; 26.6 [dB] 23.3 ; 23.2 [dB] 18.4 ; 18.4 [dB]

(c) Deletion of all horizontal sub-images



RMSE = 0.756 ; 0.758 1.22 ; 1.23 1.84 ; 1.86
PSNR = 26.0 ; 25.9 [dB] 21.8 ; 21.7 [dB] 18.2 ; 18.1 [dB]

(d) Deletion of all vertical sub-images

Fig. 9 RMSE and PSNR of reconstructed images caused by deleting expansion coefficients of sub-images at each resolution level in the case of D_2 filters: $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$

crete Cosine Transform, IEEE Trass. Computers, Vol. C-23, No. 1 (1974) pp. 90-93.

[11] R. J. Clarke : "Transform Cording of Images", Academic Press (1985) <The Karhunen-Loeve transform is introduced in Section 3.5, pp. 91-97 with specific transformd: DFT, WHT, HT, ST,

DCT, DTT, etc.〉

[12] I. Daubechies: Orthonormal bases of compactly supported wavelets, Commun. Pure Appl. Math., Vol. 41 (1988) pp. 909-996.

[13] D. E. Newland: "An Introduction to Random Vibrations, Spectral & Wavelet Analysis", Third edition, Longman Scientific & Technical (1993).

[14] M. Iizuka and M. Kitamura: Simplified compression and reconstruction techniques of digital images by means of discrete wavelet transforms, The Academic Reports, Vol. 18, No. 1 (1995) pp. 60-78.

[15] M. Iizuka: Visual appearance and its fidelity evaluation of digital image displayed by means of orthogonal transform and adaptive skatial filter, DISPLAYS, Vol. 15, No. 4 (1994) pp. 226-2240.