

Optimization of Size of Pixel Blocks for Orthogonal Transform in Optical Watermarking Technique

Yasunori Ishikawa, Kazutake Uehira, and Kazuhisa Yanaka

Abstract—We previously proposed a novel technology with which the images of real objects with no copyright protection could contain invisible digital watermarking, using spatially modulated illumination. In this “optical watermarking” technology, we used orthogonal transforms, such as a discrete cosine transform (DCT) or a Walsh-Hadamard transform (WHT), to produce watermarked images, where 1-b binary information was embedded into each pixel block. In this paper, we propose an optimal condition for a technique of robust optical watermarking that varies the size of pixel blocks by using a trade-off in the efficiency of embedded watermarking. We conducted experiments where 4×4 , 8×8 , and 16×16 pixels were used in one block. A detection accuracy of 100% was obtained by using a block with 16×16 pixels when embedded watermarking was extremely weak, although the accuracy did not necessarily reach 100% by using blocks with 4×4 or 8×8 pixels under the same embedding conditions. We also examined the effectiveness of using a Haar discrete wavelet transform (Haar DWT) as an orthogonal transform under the same experimental condition, and the results showed that the accuracy of detection was slightly inferior to DCT and WHT under very weak embedding conditions. The results from experiments revealed the effectiveness of our new proposal.

Index Terms—Digital watermarking, optical watermarking, orthogonal transform, protection against illegal photographing, spatially modulated illumination.

I. INTRODUCTION

AS DIGITAL IMAGE content is increasingly being distributed throughout various media, techniques of digital watermarking have been widely recognized as methods of protecting the copyrights of image content [1]–[8]. For example, digital watermarking is embedded in digital data before it is printed to prevent illegal use of images copied by digital cameras or scanners [9]–[12]. However, digital watermarking with this method has to be embedded before the image content itself is distributed. This cannot prevent photographs of valuable paintings in museums and galleries from being illegally taken with digital cameras.

We previously proposed a novel technology that could prevent the illegal use of images of objects that did not have watermarking [13]–[16]. This “optical watermarking” technique

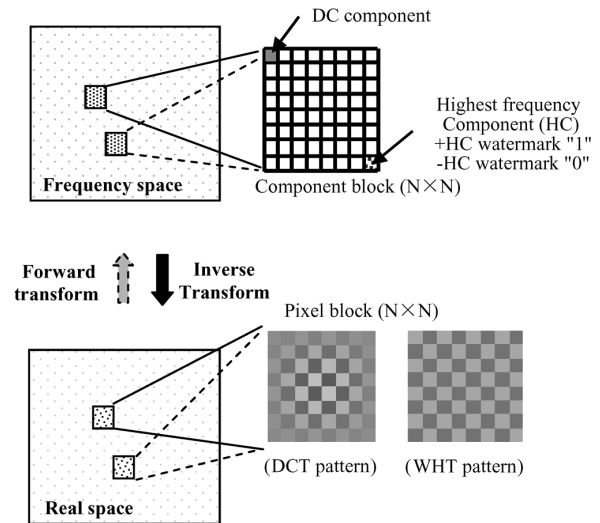


Fig. 1. Producing watermarks using DCT and WHT.

used illumination that invisibly contained the watermarking. An image of an object irradiated with such illumination also contained watermarking. The watermarking from an image taken with a camera could be extracted by image processing. We used orthogonal transforms, such as a discrete cosine transform (DCT) or a Walsh-Hadamard transform (WHT) as methods of embedding the watermarking. We also used a Haar discrete wavelet transform (Haar DWT) as an orthogonal transform [19]. Moreover, we also previously proposed techniques that were robust to various distortions due to the shooting and reflectance conditions of objects in practical cases [18].

We define the optimized pixel size in this paper to ensure 100% detection accuracy and to maximize the number of embedding bits per unit block size. We clarify the relation of the volume of data that can be embedded under unit pixel counts to the accuracy of detecting embedded watermarking data, and we propose an optimal condition for the technique of robust optical watermarking using an orthogonal transform.

II. PROCEDURE FOR PRODUCING OPTICAL WATERMARKS

Fig. 1 illustrates the procedure for watermarking using DCT or WHT. The watermarked area is divided into units of $N \times N$ pixel blocks, and each block has a DC component that gives an average brightness for the entire watermarked area, i.e., brightness of illumination. Every block also has the highest frequency component (HC) in both the x - and y -directions to express the 1-b binary information for watermarking. We used the phase of HC to express binary data i.e., “0” or “1.” When a 2D inverse

Manuscript received December 08, 2011; revised February 23, 2012, April 04, 2012, and May 07, 2012; accepted May 09, 2012. Date of publication June 19, 2012; date of current version July 27, 2012.

Y. Ishikawa was with the Faculty of Information Network and Communication, Kanagawa Institute of Technology, Atsugi-shi, Kanagawa 243-0292, Japan, and also with YIT Consulting Corporation, Yokohama, Kanagawa 231-0821, Japan (e-mail: yasu@yit-pe.com).

K. Uehira and K. Yanaka are with the Faculty of Information Network and Communication, Kanagawa Institute of Technology, Atsugi-shi, Kanagawa 243-0292, Japan.

Digital Object Identifier 10.1109/JDT.2012.2201133

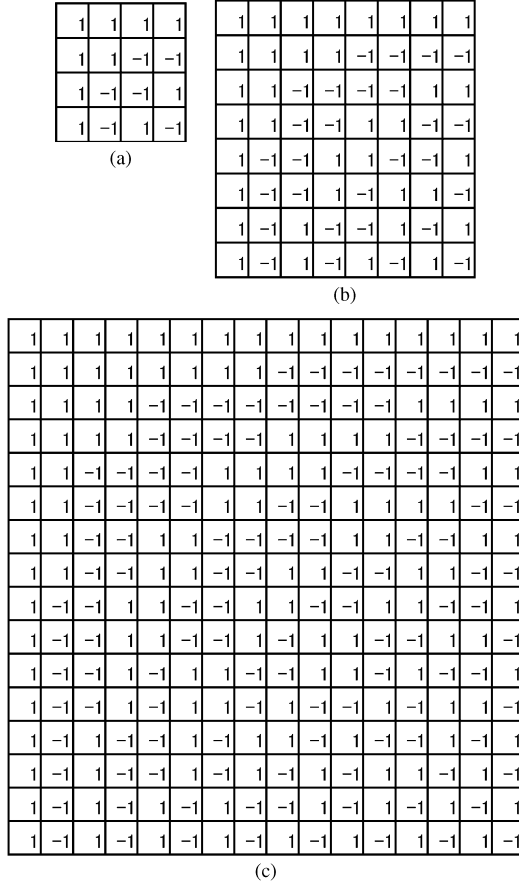


Fig. 2. Walsh-Hadamard matrix: (a) 4×4 matrix; (b) 8×8 matrix; and (c) 16×16 matrix.

DCT (i-DCT) is used to produce watermarked images, this is mathematically expressed by

$$f_{i,j}(x,y) = \sum_u^{N-1} \sum_v^{N-1} C(u)C(v)F_{i,j}(u,v) \times \cos\left\{\frac{(2x+1)u\pi}{2N}\right\} \cos\left\{\frac{(2y+1)v\pi}{2N}\right\} \quad (1)$$

where $f_{i,j}(x,y)$ are the watermarked image data for pixel (x,y) of block (i,j) in real space, $F_{i,j}(u,v)$ are the data for component (u,v) of block (i,j) in frequency space, and N is the number of pixels in the block in the x - and y -directions. Here, $C(u)$ and $C(v)$ are given as

$$C(u) = \begin{cases} 1 & (u=0) \\ \sqrt{2} & (u \neq 0) \end{cases}, \quad C(v) = \begin{cases} 1 & (v=0) \\ \sqrt{2} & (v \neq 0) \end{cases}.$$

When a 2D inverse WHT (i-WHT) is used, the equation is expressed by

$$f_{i,j}(x,y) = \frac{1}{N} \sum_u^{N-1} \sum_v^{N-1} F_{i,j}(u,v)wh(x,u)wh(y,v) \quad (2)$$

where $wh(i,j)$ denotes a component of the Walsh-Hadamard matrix in Fig. 2.

Fig. 3 illustrates where a Haar DWT is used to produce watermarked images, which means that a multi-resolution image is

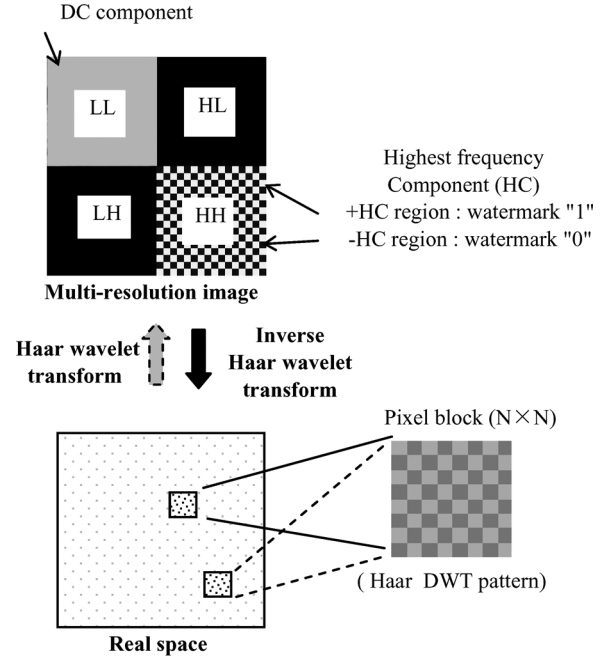
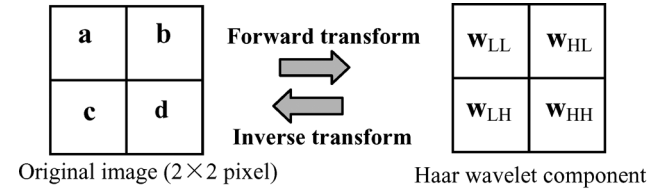


Fig. 3. Producing watermarks using Haar wavelet transform.



$$\begin{aligned} W_{LL} &= \frac{1}{4}(a+b+c+d) \\ W_{HL} &= \frac{1}{4}(a-b+c-d) \\ W_{LH} &= \frac{1}{4}(a+b-c-d) \\ W_{HH} &= \frac{1}{4}(a-b-c+d) \end{aligned} \quad \text{(Forward transform)}$$

$$\begin{aligned} a &= W_{LL} + W_{HL} + W_{LH} + W_{HH} \\ b &= W_{LL} - W_{HL} + W_{LH} - W_{HH} \\ c &= W_{LL} + W_{HL} - W_{LH} - W_{HH} \\ d &= W_{LL} - W_{HL} - W_{LH} + W_{HH} \end{aligned} \quad \text{(Inverse transform)}$$

Fig. 4. Haar wavelet transform.

used to express the layer of frequency components. A DC value is given to the whole plane of the LL component image, and this gives an average brightness to the entire watermarked area. The HC value for the HH component image is provided to every $(N/2) \times (N/2)$ component block, and this yields the 1-b binary information as watermarking data. The phase of HC is used to express binary data, i.e., "0" or "1." All component values for the LH component image and the HL component image are provided to "0." With $(M/2) \times (M/2)$ pixels as the size of each component image, a watermarked image of $M \times M$ pixels is produced using a inverse DWT. We used Haar DWT as the algorithm for the DWT. Fig. 4 has details on the Haar DWT. The equations for forward and inverse Haar DWT are in this figure, which are simple linear equations. Watermarked image data generated by Haar DWT become equivalent watermarked image data generated by WHT with the same DC and HC values.

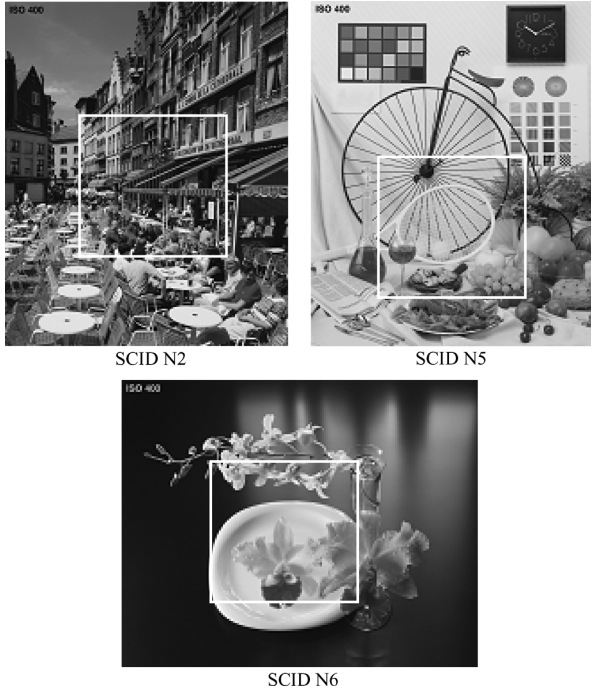


Fig. 5. Images used in experiments: Squares with white frames were used for areas irradiated with watermarks.

III. EXPERIMENTS

We produced watermarked image data in the experiments that had blocks of 4×4 , 8×8 , and 16×16 pixels and evaluated the accuracy with which embedded watermarking information could be detected. As the entire watermarked area was 128×128 pixels, the number of blocks in the watermarked area was 32×32 when a block with 4×4 pixels was used, 16×16 when a block with 8×8 pixels was used, and 8×8 when a block with 16×16 pixels was used. The watermarked images were generated as binary information that was embedded as blocks of “0” and “1,” which were alternately positioned in a checkerboard pattern. They were projected onto pictures as objects with a projector, and the image data including an optical watermark were captured with a digital camera.

A digital light processing (DLP) projector was used as a light source that had a resolution of 800×600 pixels. The objects were printed A4 images of the standard image data. The images in Fig. 5 were used for the pictures in the experiment, where the squares enclosed by the white frames indicate portions that were irradiated with a watermark in each image. The value for DC was fixed at 150, and the values of HC were varied as these were the experimental parameters. The size of the projected watermarked area was about $105 \text{ mm} \times 105 \text{ mm}$ on the object image, which was about 650×650 pixels taken with a digital camera that had a resolution of 4288×2848 pixels. Fig. 6(a)–(f) shows images of parts of objects embedded with watermarking under the condition of $\text{HC} = 15$, which is a rather strong embedding condition; therefore, the patterns in the watermarked images can easily be seen. However, it was difficult to recognize them under



Fig. 6. Part of magnified image with watermarking: (a) 4×4 pixels, DCT, DC = 150, HC = 15; (b) 8×8 pixels, DCT, DC = 150, HC = 15; (c) 16×16 pixels, DCT, DC = 150, HC = 15; (d) 4×4 pixels, WHT and Haar DWT, DC = 150, HC = 15; (e) 8×8 pixels, WHT and Haar DWT, DC = 150, HC = 15; and (f) 16×16 pixels, WHT and Haar DWT, DC = 150, HC = 15.

the condition of $\text{HC} = 5$ or less which was revealed in our previous study [16].

A rectangle was clipped out from the captured image data as a watermarked area that was brighter than its neighbors. The clipped area was then reduced to 128×128 pixels with image processing and was divided into each block of $N \times N$ pixels, and a forward orthogonal transform was carried out on all blocks. When DCT was used, (3) was used as the forward transform

$$F_{i,j}(u,v) = \frac{C(u)C(v)}{N \times N} \sum_x \sum_y^{N-1} f_{i,j}(x,y) \cdot \cos \left\{ \frac{(2x+1)u\pi}{2N} \right\} \cos \left\{ \frac{(2y+1)v\pi}{2N} \right\}. \quad (3)$$

Here, $C(u)$ and $C(v)$ are the same as those in (1).

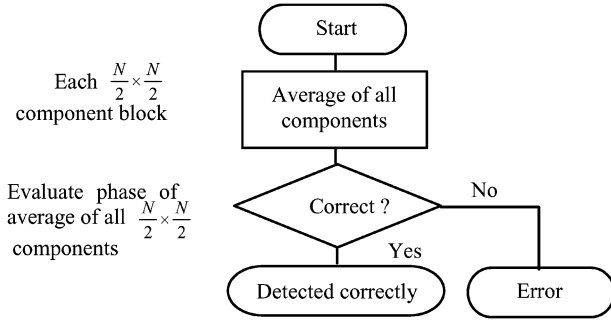


Fig. 7. Procedure for detecting optical watermarking using Haar DWT.

When WHT was used, (4) was utilized for the forward transform, where the values in Fig. 2 were used as the components of matrix $wh(i, j)$

$$F_{i,j}(u, v) = \frac{1}{N} \sum_x \sum_y^{N-1} f_{i,j}(x, y) wh(u, x) wh(y, v) \quad (4)$$

The accuracy with which the embedded data were read out was evaluated by checking the sign of the $F_{i,j}(N-1, N-1)$ components for all blocks where either DCT or WHT was used.

When Haar DWT was used, the watermarked area that was clipped out was also transformed to 128×128 pixels. As forward Haar DWT was carried out on the entire watermarked area of 128×128 pixels, a multi-resolution image was obtained, and the HH component image was separated from this multi-resolution image. If the embedded watermarking information was correctly read out, a +HC value or -HC value appeared on every $(N/2) \times (N/2)$ component block of the HH component image. However, if the spatial-frequency component of the object image contained an HH frequency element, the coefficients in a $(N/2) \times (N/2)$ component block may be disrupted by noise derived from this element. We therefore used the procedure in Fig. 7 to read out the embedded watermarking data. First, the mean value of all coefficients of every $(N/2) \times (N/2)$ component block in the HH component image was calculated, where 1-bit binary information was embedded, and the mean value was evaluated. If the phase of the mean value we obtained was negative the watermarking data that was read out was “0”, and if the phase was positive, it was “1.”

The accuracy of detection of embedded data read out from the watermarked image we obtained was evaluated with the rate of correctly read out data to whole embedded data in the watermarked image where blocks of “0” and “1” were alternately positioned like those in a checkerboard pattern.

IV. RESULTS AND DISCUSSION

Fig. 8(a)–(c) presents the results we obtained from the experiments when the SCID N2 image was used as an object image, where the unit of the accuracy is the rate of the number of correctly detected blocks to the number of whole blocks in these figures. When the block of 16×16 pixels was used, a detection accuracy of 1.0 was obtained using DCT and WHT for all the

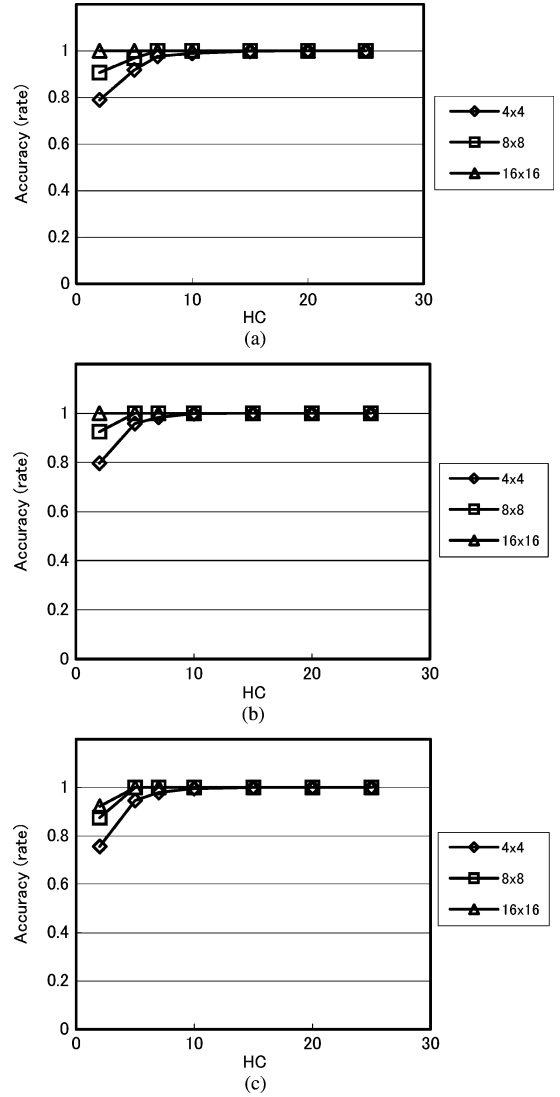


Fig. 8. Accuracy with which data were read out: (a) (SCID N2): DCT; (b) (SCID N2): WHT; and (c) a (SCID N2): Haar DWT.

HC values. In regions of $HC = 10$ and under, the detection rate when a block of 8×8 pixels was used was superior to that when a block of 4×4 pixels was used, when DCT and WHT were employed. Here, WHT demonstrated a slightly more accurate detection rate than DCT. The results when Haar DWT was used indicated that blocks with more pixels had better accuracy in detection. These were the same results as when DCT and WHT were used. However, when $HC = 2$, the accuracy with which data were read out did not reach 1.0.

Figs. 9(a)–(c) and 10(a)–(c) also have the results we obtained from experiments when SCID N5 and N6 images were used as object images. The tendency for the accuracy of detection when SCID N5 was used as an object image was almost the same as that when SCID N2 was used, although the overall accuracy of detection using SCID N5 was better than that for SCID N2. It should be noted that when SCID N5 and N6 were used as object images the accuracy of detection when Haar DWT was used in the block of 16×16 was 1.0 for $HC = 2$, which was different from the case of SCID N2. When SCID N6 was used, higher

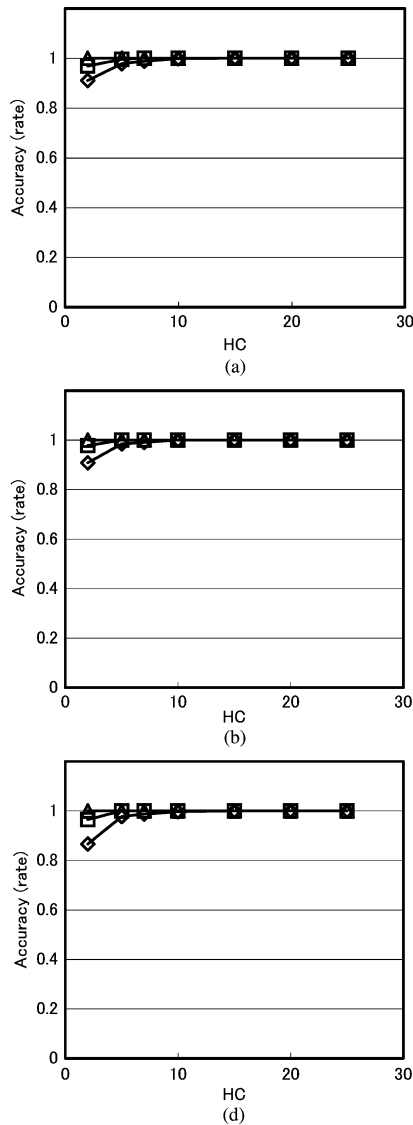


Fig. 9. Accuracy with which data were read out: (a) (SCID N5): DCT; (b) (SCID N5): WHT; and (c) (SCID N5): Haar DWT.

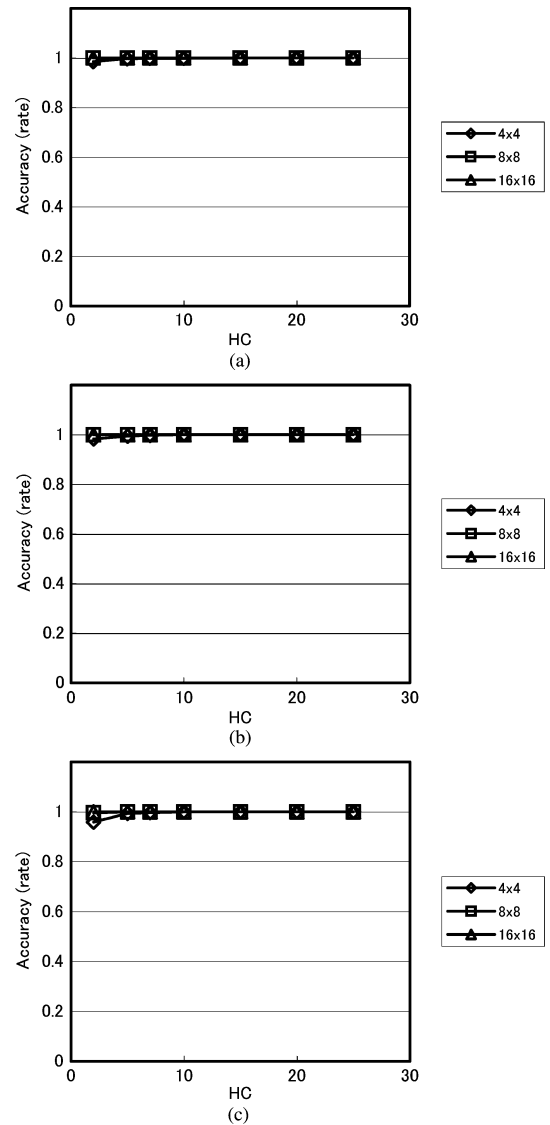


Fig. 10. Accuracy with which data were read out: (a) (SCID N6): DCT; (b) (SCID N6): WHT; and (c) (SCID N6): Haar DWT.

detection accuracies were obtained than when SCID N2 and N5 were used. That is, when DCT, WHT and Haar DWT were used, detection accuracies under all the HC values were over 0.99 in the blocks of 16×16 and 8×8 pixels. Also, they were over 0.95 for all the HC values even in the block of 4×4 pixels. The reason for the excellent accuracy of detection for SCID N6 was considered to be because it had a simpler structured picture than the N2 and N5 in Fig. 5. In other words, N6 contained rather smaller numbers of high-frequency components than N2 and N5.

We found from these results that blocks with more pixels had a better detection rate. When DCT or WHT was used in blocks with 16×16 pixels, a detection accuracy of 1.0 was obtained with a value of $HC = 2$ that enabled extremely weak embedding. We found from our previous study that watermarked images were almost invisible at $HC = 2$ [16]. When Haar DWT was used, the accuracy of detection also reached 1.0 in a block with 16×16 pixels with a value of $HC = 2$ when SCID N5 and N6 images were used as object images.

Because watermarking information is embedded under high frequency elements, if the blocks contain more pixels, embedded information is distributed within a wider area of the object image by optical watermarking, and there is the possibility that interference by a specific part of the object with low reflectivity will be low. As a result, the accuracy with which watermarking information is detected is improved, and can offer a method of optical watermarking with tolerance against interference as high as that of the entire pixel block. However, it is clearly advantageous to have fewer pixels in blocks, and this involves a trade-off in the detection rate and tolerance of watermarking against interference that enables embedding to be optimized.

From the results obtained from evaluating the accuracy of detection with Haar DWT, we also found it was inferior to that with DCT and WHT when a complicated structured picture like SCID N2 was used as the object image. This was due to the following causes. The conversion base for Haar DWT was 2×2 pixels per block, and this was fewer than the number of pixels

in the blocks for DCT and WHT in these experiments. There may be a possibility that interference in a particular block with 2×2 pixels can not be avoided, although an algorithm was used that compensated for the entire HH component image. However, DWT has an advantage in that it offers a high degree of freedom in the number of pixels per block, for instance, 6×6 and 10×10 pixels can be in one block.

V. CONCLUSION

We proposed an optimal condition for the size of pixel blocks of an orthogonal transform that was used for a technique of robust optical watermarking. The experimental results proved that it was practical and that the accuracy of detection of data embedded with optical watermarking could be improved with more pixels in each block. They revealed that under conditions of very weak embedded watermarking, the accuracy of detection using a block with 16×16 pixels reached 100%, except when Haar DWT was used to produce watermarked images and a complicated structured image was used as an object image. We also clarified that robustness against various disturbances became a trade-off in optimizing embedded watermarking data, as the volume of information using blocks with 16×16 pixels that could be embedded into data for the watermarked image was lower than that using blocks with 4×4 or 8×8 pixels. As a result, we concluded that the maximum volume of embedded bits per unit block size under conditions of 100% accuracy of detection could be determined in optical watermarking.

When Haar DWT was used, the accuracy of detection was rather inferior to that with DCT and WHT. However, as the general features of DWT indicated that the pixel resolution in real space and the spatial-frequency resolution in frequency space were independent, the accuracy of detection could be improved when more pixels were used in a block of the conversion base for DWT. We next intend to evaluate the optimal pixel size in the conversion base to obtain sufficiently accurate detection with DWT.

REFERENCES

- [1] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1673–1687, Dec. 1997.
- [2] M. D. Swanson, M. Kobayashi, and A. H. Tewfik, "Multimedia data-embedding and watermarking technologies," *Proc. IEEE*, vol. 86, no. 6, pp. 1064–1087, Jun. 1998.
- [3] M. Hartung and M. Kutter, "Multimedia watermarking techniques," *Proc. IEEE*, vol. 87, no. 7, pp. 1079–1107, Jul. 1999.
- [4] G. C. Langelaar, I. Setyawan, and R. L. Lagendij, "Watermarking digital image and video data," *IEEE Signal Process. Mag.*, vol. 17, no. 5, pp. 20–46, 2000.
- [5] J. Haitisma and T. Kalker, "A watermarking scheme for digital cinema," in *proc. ICIP2001*, 2001, vol. 2, pp. 487–489.
- [6] "Digital cinema system specification V1.2," Digital Cinema Initiatives Mar. 2008.
- [7] S. Goshi, H. Nakamura, H. Ito, R. Fujii, M. Suzuki, S. Takai, and Y. Tani, "A new watermark surviving after re-shooting the images displayed on a screen," in *Proc. KES2005, LNAI3682*, 2005, pp. 1099–1107.
- [8] K. Okihara, Y. Inazumi, and H. Kinoshita, "A watermark method that improve the relationship between of the number of embedded bits and image degradation," *J. IEIJ*, vol. 58, no. 10, pp. 1465–1467, 2004.
- [9] T. Mizumoto and K. Matsui, "Robustness investigation of DCT digital watermark for printing and scanning," *Trans. IEICE (A)*, vol. J85-A, no. 4, pp. 451–459, 2002.
- [10] M. Ejima and A. Miyazaki, "Digital watermark technique for hard copy image," *Trans. IEICE (A)*, vol. J82-A, no. 7, pp. 1156–1159, 1999.
- [11] Y. Horiuchi and M. Muneyasu, "Information embedding to the printing images based on DCT," in *Proc. ITC-CSCC2004*, 2004, p. 7F3P50-1-4.
- [12] Z. Liu, "New trends and challenges in digital watermarking technology: Application for printed materials," in *Multimedia Watermarking Techniques and Applications*, B. Furht and D. Kirovski, Eds. London, U.K.: Auerbach, 2006, pp. 289–305.
- [13] K. Uehira and M. Suzuki, "Digital watermarking technique using brightness-modulated light," in *Proc. ICME2008*, 2008, pp. 257–260.
- [14] K. Uehira, Y. Ishikawa, and K. Yanaka, "Optical watermarking robust to object with low-reflectance," in *Proc. ICME2009*, 2009, pp. 1756–1759.
- [15] Y. Ishikawa, K. Uehira, and K. Yanaka, "Illumination watermarking technique using orthogonal transforms," in *Proc. IAS2009*, 2009, pp. 257–260.
- [16] Y. Ishikawa, K. Uehira, and K. Yanaka, "Practical evaluation of illumination watermarking technique using orthogonal transforms," *J. Display Technol.*, vol. 6, no. 9, pp. 351–358, Sep. 2010.
- [17] T. Yamada, S. Gohshi, and I. Echizen, "Re-shooting prevention based on difference between sensory perceptions of humans and devices," in *Proc. ICIP 2010*, Hong Kong, Sep. 2010, pp. 993–996.
- [18] Y. Ishikawa, K. Uehira, and K. Yanaka, "Optical watermarking technique robust to geometrical distortion in image," in *Proc. ISSPIT2010*, 2010, pp. 67–72.
- [19] Y. Ishikawa, K. Uehira, and K. Yanaka, "Embedding watermarking into real object image data using QR-code and optical watermarking technique," in *Proc. ICIP2011*, 2011, pp. 34–38.



Yasunori Ishikawa received the B.S. degree in measurement and instrumentation engineering from the University of Tokyo, Japan, in 1977, and the Ph.D. degree in information engineering from Kanagawa Institute of Technology, Japan, in 2011.

He worked for Fuji Electric Corporation and Ricoh Corporation from 1977 to 2001, where his work involved the research and development of embedded systems for consumer products, image processing and communication products, and research on image compression. Since 2001, he has been the president

of YIT Consulting Corporation, offering technical consulting and technology transfer in his fields of expertise.



Kazutake Uehira received the B.S. and M.S. degrees in electronics in 1979 and 1981, and his Ph.D. in 1994, all from the University of Osaka Prefecture, Osaka, Japan.

He joined NTT Electrical Communication Laboratories in Tokyo, Japan, in 1981. Since then, he has been engaged in the research and development of image acquisition technologies, display systems, and high-reality video communication systems. In 2001, he joined Kanagawa Institute of Technology as a professor and is currently engaged in research

on 3-D displays and optical watermarking.



Kazuhisa Yanaka received the B.S., M.S., and D.E. degrees from the University of Tokyo in 1977, 1979, and 1982, respectively.

In 1982, he joined NTT Electrical Communication Laboratories in Tokyo, Japan. In 1997, he joined Kanagawa Institute of Technology, Kanagawa, Japan, where he is currently a professor. For more than 30 years, he has been researching various aspects of images such as image processing, image communication, image input/output systems, and optical watermarking.