Proceedings of the IIEEJ Image Electronics
and Visual Computing Workshop 2012
Kuching, Malaysia, November 21-24, 2012

OPTICAL WATERMARKING TECHNIQUE APPLIED TO MOVING OBJECTS

Yasunori Ishikawa, Kazutake Uehira, and Kazuhisa Yanaka

Kanagawa Institute of Technology

ABSTRACT

We propose the application of optical watermarking to real moving objects like the faces of people who are walking. Optical watermarking technology uses spatially modulated light containing invisible watermarked data that irradiates real objects to make photographed data contain the embedded watermarking. We performed experiments where illumination containing time-sequential data embedded with watermarking information in each frame was irradiated onto the model of a human face with motion, and the watermarking information was extracted from moving image data captured with a video camera. A technique was used to correct geometrical distortion in the image of the watermarked area that occurred on the curved surface of a 3D object. As a result, we found that the embedded time-sequential watermarking information was extracted from the moving image data with a sufficient degree of accuracy.

1. INTRODUCTION

Optical watermarking technology has several important features that cannot be obtained with conventional digital watermarking methods [1]-[7]. One of these features is that it can be applied to 3D objects like human faces, and another is its ability to update watermarked data in real time. If optical watermarking is irradiated onto a moving 3D object with these features and the object is captured as moving image data with a video camera, time-sequential watermarking information is embedded into the moving image data, and the embedded watermarking information can be extracted from the captured moving image data with appropriate image processing. We have already proposed a method of optical watermarking irradiated onto still 3D objects where the embedded watermarking information was extracted by correcting geometrical distortions that occurred on the surface of a 3D object [8]. We have also proposed a method where illumination containing time-sequential watermarking information was irradiated onto objects on plain surfaces and the embedded watermarking information was extracted as time-sequential data from moving image data captured with a video camera [9].

We performed experiments assuming a practical situation where the optical watermarking was projected onto a moving person and the embedded watermarking information was extracted from the captured moving image data using the knowledge we obtained from these earlier proposals. We also propose a new method of correcting geometrical distortion for more practical environments.

2. BASIC TECHNIQUE OF OPTICAL WATERMARKING

Fig. 1 outlines the basic concept underlying the technology of optical watermarking, where an object is illuminated by light that contains invisible information. As the illumination includes watermarking data, any photographed image of the object illumined with this lighting will also include watermarking. By digitizing this photographed image, watermarking information can be extracted in the same way as with the conventional watermarking technique. The light source provides a distribution of 2D-illumination like that with a projector, and the watermarking data are expressed in the form of this 2D-illumination distribution. However, the spatial modulation in illumination has to be imperceptible to the human-visual system. The brightness distribution given by this light source then looks uniform to the observer over the object, the same as that with conventional illumination. The brightness of the object's surface is proportional to the product of the reflectance of the object's surface and illumination by incident light.

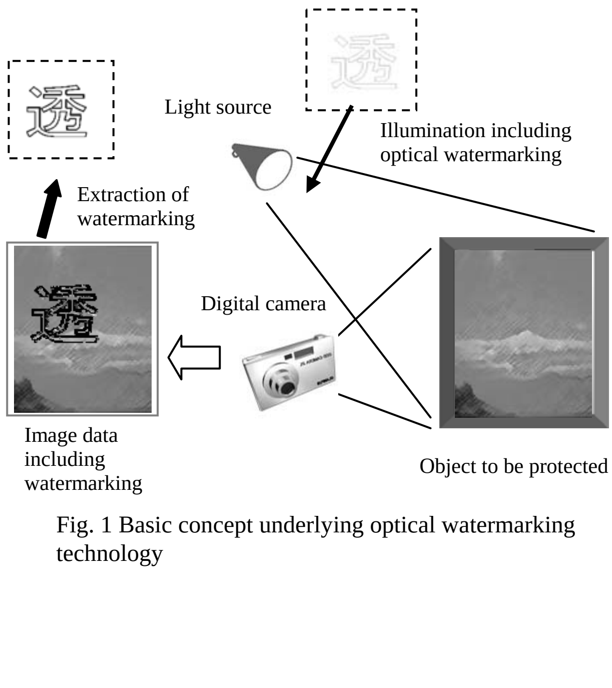

Fig. 1 Basic concept underlying optical watermarking technology

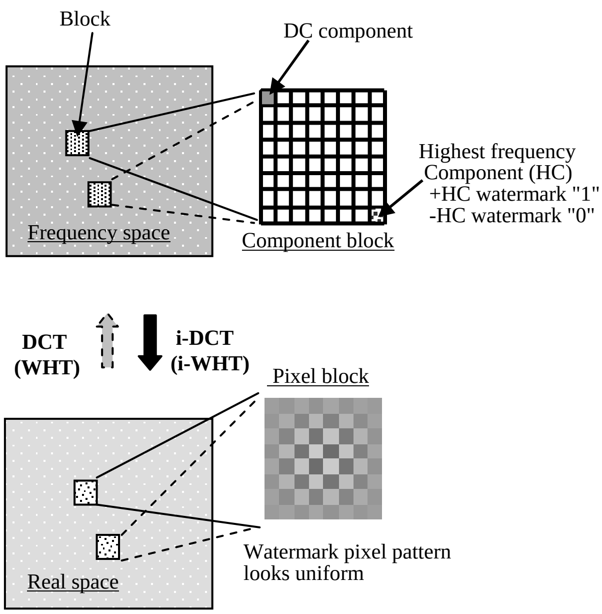

Fig. 2 Producing watermarks using DCT and WHT

Table 1 Walsh-Hadamard matrix

(a) 8×8 matrix

1	1	1	1	1	1	1	1
1	1	1	1	−1	−1	−1	−1
1	1	−1	−1	−1	−1	1	1
1	1	−1	−1	1	1	−1	−1
1	−1	−1	1	1	−1	−1	1
1	−1	−1	1	−1	1	1	−1
1	−1	1	−1	−1	1	−1	1
1	−1	1	−1	1	−1	1	−1

(b) 16×16 matrix

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	−1	−1	−1	−1	−1	−1	−1	−1
1	1	1	1	−1	−1	−1	−1	−1	−1	−1	−1	1	1	1	1
1	1	1	1	−1	−1	−1	−1	1	1	1	1	−1	−1	−1	−1
1	1	−1	−1	−1	−1	1	1	1	1	−1	−1	−1	−1	1	1
1	1	−1	−1	−1	−1	1	1	−1	−1	1	1	1	1	−1	−1
1	1	−1	−1	1	1	−1	−1	−1	−1	1	1	−1	−1	1	1
1	1	−1	−1	1	1	−1	−1	1	1	−1	−1	1	1	−1	−1
1	−1	−1	1	1	−1	−1	1	1	−1	−1	1	1	−1	−1	1
1	−1	−1	1	1	−1	−1	1	−1	1	1	−1	−1	1	1	−1
1	−1	−1	1	−1	1	1	−1	−1	1	1	−1	1	−1	−1	1
1	−1	−1	1	−1	1	1	−1	1	−1	−1	1	−1	1	1	−1
1	−1	1	−1	−1	1	−1	1	1	−1	1	−1	−1	1	−1	1
1	−1	1	−1	−1	1	−1	1	−1	1	−1	1	1	−1	1	−1
1	−1	1	−1	1	−1	1	−1	−1	1	−1	1	−1	1	−1	1
1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1

The main attribute of the technology is that the watermarking can be added by light. Therefore, it can be applied to objects that cannot be electronically embedded with conventional watermarking, such as paintings created by renowned artists. Moreover, it offers the possibility of being applied to 3D objects, such as sculptures, merchandise, and even the human body, as well as being applied to 2D objects.

Fig. 2 illustrates the procedure for watermarking using DCT or WHT. The watermarking 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 watermarking 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-bit binary information for watermarking. We used the phase of HC to express binary data i.e., "0" or "1". When a 2D inverse DCT (i-DCT) is used to produce watermarking images, this is mathematically expressed by Eq.(1).

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

where $f_{i,j}(x,y)$ are the watermarking 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

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

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

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

where $wh(i,j)$ denotes a component of the Walsh-Hadamard matrix in Table 1.

3. EXPERIMENTS

We carried out experiments on detecting optical watermarking from the image of a moving object taken with digital video cameras, where the optical watermarking was embedded with time-domain sequential data and was irradiated onto the moving object. The experiments was focused on whether watermarking could be efficiently detected in the sequential digital images of the moving object.

Watermarking images of all frames in the time-domain sequential data were produced that consisted of 80 × 60 blocks in this experiment, where each block had 8 × 8 pixels, i.e., the watermarking images had 640 × 480 pixels. Binary watermarking information was embedded as blocks of "0" and "1" that were alternately placed horizontally and vertically in each frame. We created "Even pattern frame" and "Odd pattern frame" that were changed of the phase of embedding information of "1" and "0", and sequential watermarking information data were then produced using "Even pattern frame" and "Odd pattern frame" that were alternately displayed with a frame rate of 10 frames per second (fps). When this moving picture data were input into the projector, "Even pattern frame" and "Odd pattern frame" were alternately displayed on the object every 1/10 sec.

We used a liquid crystal projector with a resolution of 1024×768 pixels, and the moving image mode of a digital camera with a resolution of 1280×720 pixels and a motion-JPEG format of 30 fps to take a video picture of moving watermarked objects. We fixed the value for DC to 100 and changed HC, which were the experimental parameters. The HC value controlled the strength of embedded watermarking. The motion of the model of a human face was created in the experiments by manually moving it several centimeters horizontally and vertically. Figs. 3 (a) and (b) are magnified images of the model of a human face that were irradiated with optical watermarked illumination.

Fig. 3 (a) Magnified image with watermarking (DCT,HC=15)

Fig. 3 (b) Magnified image with watermarking (WHT,HC=15)

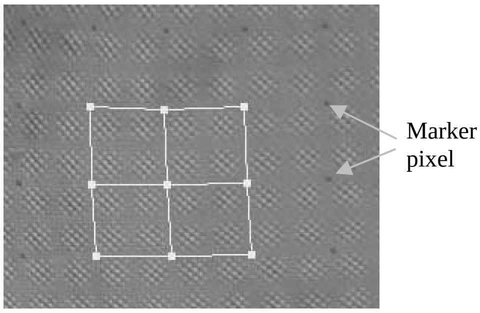

Fig. 4 (a) Discriminating the area of 2×2 blocks (DCT)

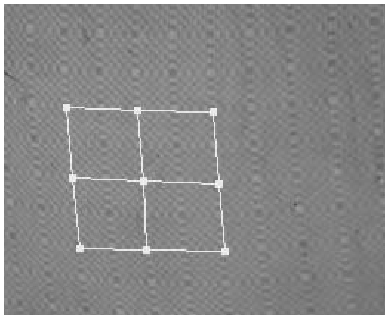

Fig. 4 (b) Discriminating the area of 2×2 blocks (WHT)

The video data we obtained were converted to sequential still image data of 30 fps, and embedded watermarking information was extracted from each separated frame. As the timing between projected image data (10 fps) and the image data we obtained (30 fps) were not synchronized in the time domain, one out of three consecutive frames of the obtained image data did not correctly capture the projected frame. We used the frame in the three consecutive frames that had the best detection rate to evaluate the accuracy of detection.

We used the following method to discriminate each 8×8 pixel block in the moving image data we obtained. That is, when DCT was used to produce watermarked data, marker pixels were dotted on the boundary of 2×2 blocks, and using these marker pixels the boundaries could be determined by manually measuring the captured watermarked data of each frame (Fig. 4(a)). When WHT was used, it could easily be discriminated using the pixel pattern of WHT itself; it could also easily be discriminated by measuring it manually (Fig. 4(b)).

We used 4×4 blocks of the watermarked area that were cut out from each frame using the boundary information mentioned above. Then, 2×2 blocks were segmented in the cut out area, and the coordinates of the corner points of all 2×2 blocks segmented areas were measured respectively. These segmented areas were distorted geometrically on the curved surface of the model of a human face. We used following method to correct the geometrical distortion by using the coordinates of the corner points of the segmented area, as each segmented region might become a precise square, assuming that the segmented region was a plane surface.

The transformation from an undistorted coordinate system (x, y) to a geometrically distorted system (x', y') is generally expressed by following equations.

$$x' = h_1(x, y)\,,\ \ y' = h_2(x, y) \qquad (3)$$

If the distortion is perspective, the transformation is expressed by the following linear equations.

$$x' = ax + by + d\,,\ \ y' = dx + ey + f \qquad (4)$$

The shape of the image containing the generated watermark is a precise rectangle, and all segmented areas are also precise rectangles. However, if these were distorted in perspective, they become general quadrangles. When the coordinates of all corner points of a quadrangle are given, the coefficients of the above-mentioned linear equations can be determined from three of these and the corresponding coordinates of the original undistorted rectangle. Using these equations the value of pixels in the distorted quadrangle can be transformed to the value of pixels in the undistorted rectangle. However, because the coordinates of transformed pixels do not generally become integers, an interpolation technique is utilized to determine the density value of the nearest pixel. Linear transformation using the four nearest neighboring pixels was used in the experiments, which is so called "bi-linear interpolation".

The cut out watermarked area was then transformed to just 64×64 pixels and divided into 4×4 blocks, each of which had 16×16 pixels. We also carried out DCT on all blocks using Eq. (5),

$$F_{i,j}(u,v) = \frac{C(u)C(v)}{M \times M} \sum_{x}^{M-1} \sum_{y}^{M-1} f_{i,j}(x,y) \bullet \cos\{\frac{(2x+1)u\pi}{2M}\} \cos\{\frac{(2y+1)v\pi}{2M}\} \qquad (5)$$

We also utilized Eq. (6) for WHT, using the values in Table 1(b) as the components of matrix $wh(i, j)$.

$$F_{i,j}(u,v) = \frac{1}{M} \sum_{x}^{M-1} \sum_{y}^{M-1} f_{i,j}(x,y) wh(u,x) wh(y,v) \qquad (6)$$

where M is the number of pixels in the u and v directions in frequency space, which was 16 in the experiments.

The information with which the embedded data were restored was evaluated by checking the sign of the $F_{i,j}(7,7)$ components for all blocks. The restored information was compared with the embedded information, and the accuracy of detection was calculated with the number of blocks that were correctly restored to 4×4 = 16 blocks.

4. RESULTS AND DISCUSSION

Figs. 5 (a) -- (f) plot the time-series for the accuracy of detection of embedded watermarking under the

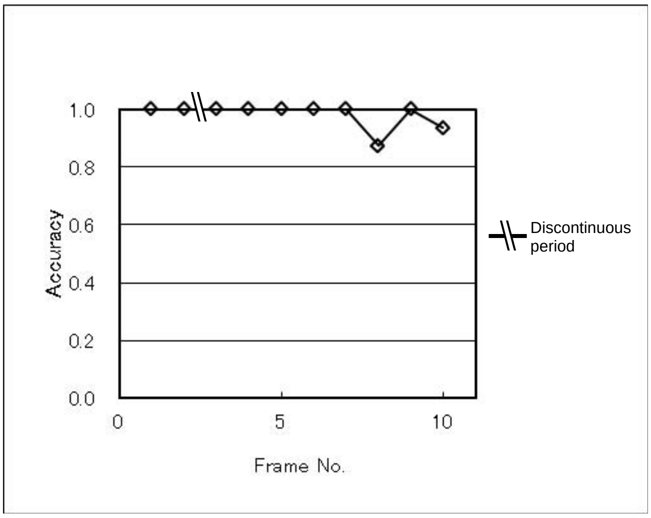

Fig. 5 (a) Sequential accuracy of detection (DCT, HC=05)

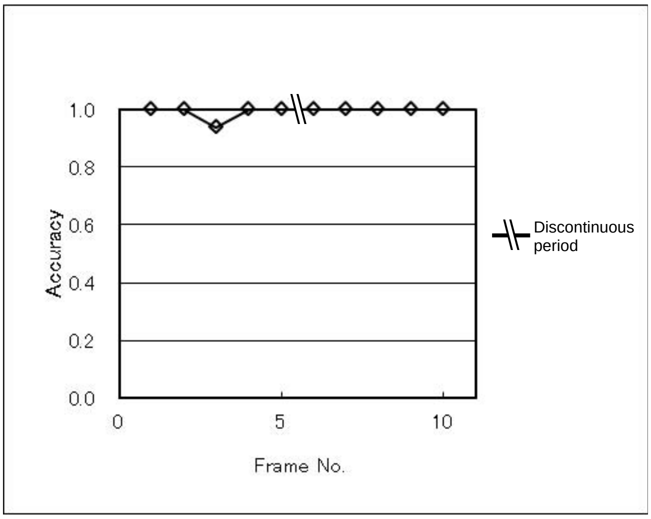

Fig. 5 (b) Sequential accuracy of detection (DCT, HC=10)

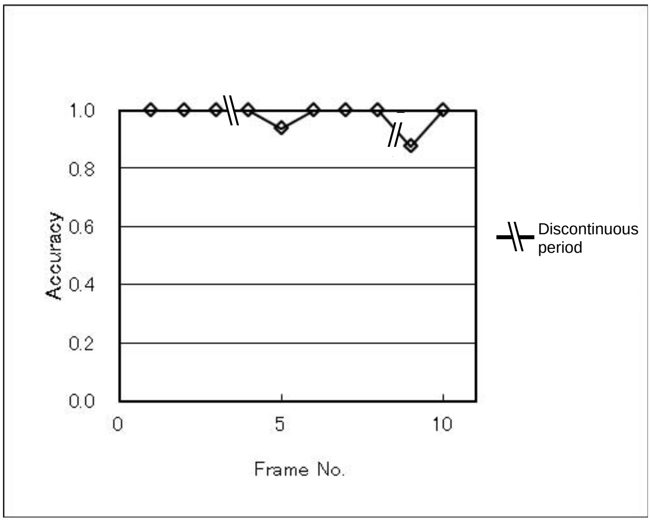

Fig. 5 (c) Sequential accuracy of detection (DCT, HC=15)

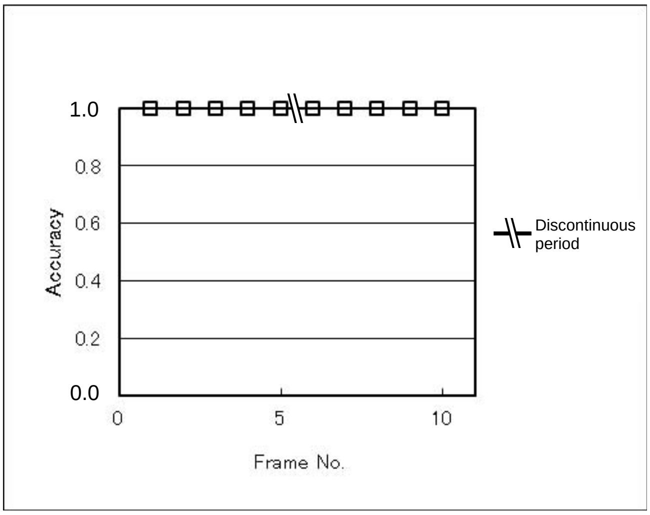

Fig. 5 (d) Sequential accuracy of detection (WHT, HC=05)

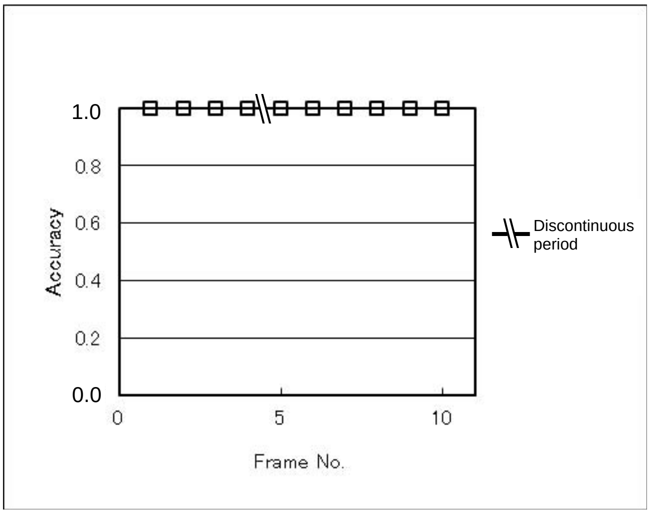

Fig. 5 (e) Sequential accuracy of detection (WHT, HC=10)

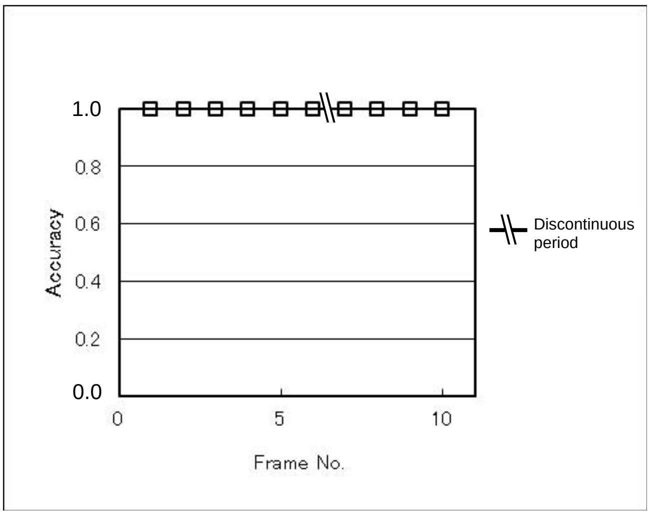

Fig. 5 (f) Sequential accuracy of detection (WHT, HC=15)

condition of HC=5, 10, and 15, using DCT and WHT, that was derived from 10 frames of watermarked image. We used the frame in the three consecutive frames of obtained video data of 30fps that had the best detection rate to evaluate the accuracy of detection as mentioned before. A detection rate of 1.0 was achieved for all frames when using WHT, although the detection rate did not reach 1.0 for all frames when using DCT. However, it seemed that the accuracies of detection had almost same tendency under the condition of HC=05, 10 and 15 when using DCT. As shown in Figs. 5 (a) -- (f), these 10 frames did not necessarily have complete continuity in time-sequence. In the discontinuous period in the figures the accuracies of detection were seriously degraded. We observed that the pattern of optical watermarking obtained with video camera became periodically indistinct because defocusing of image occurred periodically, when we took a video of the moving model of a human face. As a result the accuracies of detection were degraded periodically.

The digital camera used in the experiments seemed to not have enough stability in its auto-focusing capabilities for the moving image mode from our observation of the moving image data we obtained. On the other side the reason of instability of the accuracy of detection by using DCT other than WHT was supposed to be inexactness of manually measuring the marker points that identified the area of 4×4 blocks.

We embedded the same watermarking information into every 4×4 blocks in the whole watermarked area in the experiments. Therefore, embedded 16 bits watermarking information was correctly restored if at least one of 4×4 blocks could be determined.

5. CONCLUSION

We evaluated the detection rate of watermarked image of moving real objects that were irradiate with optical watermarking embedding with time-domain sequential watermarking information. We introduced a new method to discriminate each 4×4 blocks in whole watermarked area that were geometrically distorted. That is, when DCT was used to produce watermarked image, marker pixels were dotted on the corner points of 2×2 blocks and they were used to find the boundary of 2×2 blocks. When WHT was used, the pixel pattern of WHT itself was used to discriminate the boundary of 2×2 blocks. Using this method, when all 4×4 blocks are embedded with the same 16 bits watermarking information in whole optical watermarking area, embedded watermarking information can be correctly restored from at least one of 4×4 block that is detected on a moving object.

As a result we concluded that enough accuracy was obtained on detecting time-domain sequential watermarking information from the video data of the moving model of a human face taken with a digital video camera, if sufficient ability of auto-focusing was equipped to take a moving picture. However, the moving picture mode of the digital camera we used in the experiments did not occasionally show enough stability of focusing of image, we could not achieve 100% on the accuracy of detection. We next intend to carry out experiments using a video camera with more stable auto-focusing capabilities.

6. REFERENCES

[1] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia", IEEE Trans. Image Processing, vol. 6, No. 12, pp.1673-1687 (1997).

[2] J. Haitsma and T. Kalker, "A Watermarking Scheme for Digital Cinema", Proc. ICIP 2001, No. 2, pp.487-489 (2001).

[3] "Digital cinema system specification V1.2", Digital Cinema Initiatives, Mar. (2008).

[4] 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", KES2005, LNAI3682, pp.1099-1107 (2005)

[5] S. G. Narasimhan, S. J. Koppal, and S. Yamazaki, "Temporal Dithering of Illumination for Fast Active Vision", Proc. ECCV, pp.830-844 (2008).

[6] K. Uehira and M. Suzuki, "Digital watermarking technique using brightness-modulated light", Proc. ICME2008, pp.257-260 (2008).

[7] Y. Ishikawa, K. Uehira and K. Yanaka, "Practical evaluation of Illumination watermarking technique using orthogonal transforms", J. Display Technology, vol.6, No.9, pp.351-358 (2010).

[8] Y. Ishikawa, K. Uehira and K. Yanaka, "Protection of 3D objects against illegal photography using optical watermarking technique with spatially modulated illumination", Proc. IMAGAPP2011, pp.49-52 (2011).

[9] Y. Ishikawa, K. Uehira and K. Yanaka, "A Technique of Time Domain Sequential Data Embedding into Real Object Image Using Spatially Modulated Illumination", Proc. ISSPIT2011, pp.62-66 (2011).