Electronic Imaging

SPIEDigitalLibrary.org/jei

Optical watermarking technology for protecting portrait rights of threedimensional shaped real objects using one-dimensional high-frequency patterns

Mizuho Komori Kazutake Uehira



Optical watermarking technology for protecting portrait rights of three-dimensional shaped real objects using one-dimensional high-frequency patterns

Mizuho Komori Kazutake Uehira Kanagawa Institute of Technology Shimo-Ogino 1030 Atsugi 243-0292, Japan E-mail: uehira@nw.kanagawa-it.ac.jp

Abstract. We propose a new technology of optical watermarking that uses a one-dimensional high-frequency pattern to protect the portrait rights of three-dimensional (3-D) shaped real objects by preventing the use of images captured illegally with cameras. We conduct experiments using human faces as real 3-D objects assuming that this technology would be applied to human faces to protect their portrait rights. We utilize the phase difference between two color component patterns, i.e., binary information was expressed if the phase of the high-frequency pattern was the same or opposite. The experimental results demonstrate that this technique was robust against the pattern being deformed due to the curved surface of the 3-D shaped object. and the highest degree of accuracy of 100% in reading out the embedded data was possible by optimizing the conditions under which data were embedded. As a result, we can confirm that the technique we propose is feasible. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1 .JEI.22.3.033004]

1 Introduction

The protection of copyrights of digital-image content has become more important because increasingly more digital-image content is being distributed throughout the Internet and it can be copied exactly the same as that of the original because it is digital. Digital watermarking is an effective way of protecting copyrights from being illegally copied. Various techniques of digital watermarking for digital images have been developed.^{1–6} Digital watermarking has also been recently used in printed images, where digital watermarking is embedded in the digital data before it is printed.^{7,8} This is to prevent images copied by digital cameras or scanners from being illegally used.

However, whether digital watermarking is in the displayed image on an electronic display or on a printed image, conventional digital watermarking rests on the premise that people who want to protect the copyrights of their content have the original digital data because it has been embedded by digital processing. However, there are some cases where this premise does not apply. One such case can arise for images that have been illegally produced by people taking photographs of real objects that are invaluable as portraits, e.g., art works at museums that have been painted by a famous artists or faces of celebrities on a stage. The images produced by malicious people capturing these real objects with digital cameras or other image-input devices have been vulnerable to illegal use since they have not contained digital watermarking.

We have previously proposed a technology that could prevent the images of objects from being used in such cases.^{9,10} It uses illumination that contains invisible information on watermarking. As the illumination contains the watermarking, the images of photographs of objects that are illuminated by such illumination also contain watermarking. We refer to this technique as "optical watermarking" in this paper.

We also treated flat objects in our previous study^{9,10} assuming famous paintings were being illegally photographed and used two-dimensional (2-D) high-frequency patterns as watermarking patterns. We demonstrated that this technique effectively embedded watermarking in the captured image.

However, if we use optical watermarking for three-dimensional (3-D) shaped objects that have curved surfaces, the embedded information for watermarking cannot be read out correctly because the projected patterns on the surface of objects are deformed. We therefore tried to correct deformation in the projected watermarking pattern using 2-D grid patterns.¹¹ We found that this technique made it possible to correct deformation and accurately read out watermarking information. However, it was difficult to make the grid pattern invisible. Projecting the watermarking pattern and the grid pattern separately may have been a solution to the problem; however, it could not be applied to moving objects.

The problems with the illegal use of portraits of celebrities and character goods that have been captured with cameras by general fans have recently become serious and there is a strong demand to prevent the illegal use of these images. One of the main reasons for these to have become serious problems is because the performance of recent digital cameras such as their resolution has improved drastically and the image quality of these illegally captured images has dramatically increased.

Paper 12179 received May 12, 2012; revised manuscript received Apr. 10, 2013; accepted for publication Jun. 20, 2013; published online Jul. 15, 2013.

This paper proposes new optical watermarking that can be used even for 3-D shaped objects that have curved surfaces. It does not need additional projecting patterns for correction. It uses one-dimensional (1-D) high-frequency patterns as watermarking patterns. An 1-D pattern is expected to be robust to the projected pattern being deformed because of its simplicity compared to 2-D patterns. We used a human face as a real object in this study assuming its portrait rights were protected. This paper also presents results obtained from experiments where we evaluated the accuracy with which the watermarking could be read out and the invisibility of the projected 1-D pattern on the face.

2 Optical Watermarking and Proposed 1-D Optical Watermarking

Figure 1 outlines the basic concept underlying our technology of watermarking that uses light to embed information. An object is illuminated by light that contains invisible information on watermarking. As the illumination itself contains the watermarking information, the image of a photograph of an object that is illuminated by such illumination also contains watermarking. By digitizing this photographic image of the real object, the watermarking information in binary data can be extracted in the same way as that with the conventional watermarking technique. To be more precise, information to be embedded is first transformed into binary data, "1" or "0," and it is then transformed into a pattern that differs depending on whether it is "1" or "0." This pattern is transformed into an optical pattern and projected onto a real object. It is this difference in the pattern that is read out from the captured image.

Some applications that use invisible patterns utilize infrared light¹²; however, infrared light cannot be used for our purposes because cameras usually have a filter that cuts off infrared light and the invisible pattern is not contained in the captured image of the object although it is contained in the optically projected image on the object. Therefore, the technique we propose uses visible light, and the pattern is made invisible by using fine patterns or low contrast patterns both of which are under the resolving power of the human visual system. Using this method, the pattern can be made invisible in both an optically projected image on the object and the image of the object captured with the camera.

> Information to be embedded ID: KAIT301 Binary data Invisible pattern Light source Light ID: KAIT301 Light Light Light Read out by image processing Digital camera Captured image Real object

Fig. 1 Basic concept underlying technology of watermarking that uses light to embed data.

The light source used in this technology projects the watermarking pattern similar to a projector. Since the projected pattern 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, which is the same as that with the conventional illumination. The brightness of the object's surface is proportional to the product of the reflectance of the surface of the object and illumination by an incident light. Therefore, when a photograph of this object is taken, the image on the photograph contains watermarking information, even though this cannot be seen.

The main feature of the technology we propose is that the watermarking can be added by light. Therefore, this technology can be applied to objects that cannot be electronically embedded with watermarking, such as pictures painted by the artists.

We used a method in our previous study that used a 2-D inverse discrete cosine transform to produce the watermarking pattern.⁹ The illumination area was divided into numerous numbers of blocks. Each block had 8×8 pixels. We expressed 1-bit binary data as "1" or "0" by using the sign of the high-frequency component. This method, however, is not robust to the projected pattern being deformed when projected onto a curved surface.

In contrast, we propose the 1-D high-frequency pattern shown in Fig. 2, in this study, where each vertical (or horizontal) pixel line has 1-bit binary information of "1" or "0." We use two color components to express "0" or "1," i.e., if the phases of the high-frequency pattern of two color components are the same, the binary information is "1," and if not, it is "0." Even if we cut out a small area of the pattern shown in Fig. 2 using this pattern, we can determine whether the binary information is "1" or "0" because we just need to check if the phase of the patterns of the two color components in the line is the same or not.

We assumed that for a small area, displacement of the projected pattern caused by the curvature of the object surface would not be very large. Since we captured the object image with double resolution (as will be explained later), the phase calculated from some pixel lines was not affected by the deformation seen at the right of Fig. 2. That is, even if some areas of the captured image were slanted due to the curved surface of the objects shown at the right of the figure,



Fig. 2 One-dimensional optical watermarking pattern.

as long as there were few pixels in the area of the checked phase in the line, a large part of the area that was checked whose width was 1 pixel still remained on the original line of the projected pattern. Therefore, the phase of the 1-D pattern could be read out correctly. This is the reason that this technique was expected to be robust to deformation of the projected pattern. Moreover, we could use an arbitrarily small area vertically (or horizontally) while for a 2-D pattern, the area where 1-bit binary data was embedded was fixed and the effect of deformation was significant.

Moreover, this technique is robust to trapezoidal distortion since it has the features mentioned above. This should be beneficial, especially for the applications we assumed in this study because people capture photographs from all directions.

This technique is expected to have the same level of robustness against realistic processing, like that in compression and scaling, as that of the conventional watermarking technique where data are embedded in the frequency domain.

3 Experiments

The key to optical watermarking is for the embedded watermarking to simultaneously remain both readable and invisible when it is contained in light as well as in the captured image. We carried out experiments to evaluate these characteristics in the performance of the proposed technique.

3.1 Test Pattern, Equipment, and Layout

We made test patterns where each vertical pixel line was assigned the 1-bit information in Fig. 2. We used a color image, the G-color component was used for the signal pattern and the B-color component was used for the reference pattern. We set the average brightness of the pattern to 180 and the highest-frequency component (HC) was changed from 20 to 70 as an experimental parameter. These values were on an 8-bit gray scale with a maximum of 255. The magnitude of HC indicated the degree of pattern contrast that is the brightness difference between the bright and the dark pixel. Readability was considered to decrease as HC decreased.

Figure 3 has the layout for the experiment. We used a digital light processing projector that had 1024×768 pixels for each color to project the watermarking image. The distance between the projector and the object, which was a human face, was 110 cm. The distance between the light source and the object will be much longer for many cases in practice than that in these experiments and may be several meters or more. However, the most important index is not the distance



Fig. 3 Layout for experiment on invisibility.

but the pixel density of the projected image on the surface of the object. This depends on the specifications of the lens that projects the patterns, especially, its f-number. Even if the distance between the light source and the object is several times or more, we can make the pixel density on the object the same as that in these experiments by choosing an appropriate f-number on the lens.

We used the faces of five people as real 3-D objects. We also used a manikin's face as a reference. Figure 4 is a photograph where a watermarking image has been projected onto the manikin's face. The whole projected area was 80×60 cm; therefore, the faces were part of the whole projected area.

Although the projector had 1024×768 pixels and its projected area was 80×60 cm, the signal patterns that indicated the binary data were projected using 50×50 pixels. Therefore, the signal pattern area had 50 lines and the binary data were assigned to 40 lines except for five lines to the left and right of them. We embedded data of "1" and "0" in turn to enable the accuracy of readability to be evaluated more easily using the pattern in Fig. 2. This signal pattern area was projected onto the three areas of the image on the humans' and manikin's faces in Fig. 4. One was an image on the forehead that was relatively flat and the two others were images on the right and left cheeks that were largely curved.

The projected patterns were captured with a digital camera that had 3906×2602 pixels. The captured image had over twice the pixel density of the projected image. This was because over twice as many pixels were needed according to the sampling theorem to restore the original high-frequency patterns in the projected image.

3.1.1 Evaluation of readability

After it was captured, we changed this ratio to just twice by digital processing. The signal pattern area in Fig. 4, which had 100×100 pixels in the captured image, was cut out and read out from the binary data. We manually cut out the signal pattern area from the whole image area to simplify the experiment although it should be done automatically when this technique is used in practice in the future, which should not be too difficult by finding the edge of high-frequency areas.

The frequency components, HC1(x) and HC2(x), whose phases differ by 90 deg were calculated with Eqs. (1) and (2) for each line in the captured image to obtain the phase of the high-frequency pattern using *N* pixels:

$$HC1(x) = H1I(x) \tag{1}$$





Fig. 4 Manikin's face used as 3-D real object in experiment. Parts of areas in each square were cut out to read out embedded binary data.

and

where I(x) indicates the 1-D matrix of image data consisting of N pixels whose center is x, and H1 and H2 are 1-D matrices to detect the frequency components of the embedded 1-D high-frequency patterns. Since the cycle of the highfrequency patterns in the captured image was 4 pixels long, we used a matrix whose *i*'th elements, H1_i and H2_i, are given as

 $H1_i = 1$ (if the remainder of *i* by four is 0 or 1),

-1 (if the remainder of *i* by four is 2 or 3),

 $H2_i = 1$ (if the remainder of *i* by four is 1 or 2),

-1 (if the remainder of *i* by four is 0 or 3).

For example, when N = 8, these matrices are those given in Eqs. (3) and (4):

$$H1 = [11001100] \tag{3}$$

and

$$H2 = [01100110]. \tag{4}$$

The inner product of the matrix in Eqs. (1) and (2) was done over N pixels in the y-direction. We chose 8, 12, 16, 32, and 60 as N. The start pixels for the inner product of the matrix in Eqs. (1) and (2) were determined arbitrarily because we simply had to check if there was a phase difference between the patterns in the G- and B-components.

We calculated the phase of the patterns in the captured image with Eq. (5):

$$\theta_c(x) = \arctan[\text{HC2}(x)/\text{HC1}(x)], \quad (5)$$

where $c = \{G, B\}$.

We obtained $\theta_G(x)$ for the signal pattern in the G-component image at x and $\theta_B(x)$ for the reference pattern in the B-component image at x. We determined the binary data at x to be "1" if the absolute value of the difference between the phases of the two patterns, $|\theta_G(x) - \theta_B(x)|$, was <90 deg and we determined it to be "0" if it was >90 deg. This was because we set the phase of the original signal and reference pattern to be the same when we assigned the "1" of binary data, and we set the difference between two patterns to be 180 deg when we assigned the "0" of binary data, as can be seen in Fig. 2.

Moreover, to improve the accuracy in reading out embedded data, we tested a redundancy method where the data were read out from three different parts on the same line and binary data were determined by the majority of



Fig. 5 Example of images of patterns projected onto the cheek of one of five people. Data were read out from three different parts on same line and binary data were determined from the majority of three data.

the three data. Figure 5 explains this method. This method is a kind of error-correction technique that utilizes one of the features of the proposed technique where an arbitrarily small vertical area can be cut out because we just have to check if the phases of the lines of two color components are the same or not.

3.1.2 Evaluation of invisibility

We evaluated the invisibility of the projected patterns with a subjective test. The watermarking patterns were projected onto a human face. The pixel density of the projected image was the same as that used in evaluating readability as described previously. Subjects viewed the projected pattern at a distance of 1 m. Under these conditions, the angle of view between the adjacent pixels was 2 min. The subjects were asked if they could see the pattern on the face. HC was chosen as an experimental parameter and patterns with different HC values were presented at random to the subjects. Five subjects participated in the evaluation and they all had corrected eyesight of over 1.0 at least.

4 Results and Discussion

Figure 6 has an example of color component images from which we evaluated readabilities. They are patterns projected onto the right cheek of one of five people. They are magnified images, and the magnitude of HC of the projected pattern was 60. We can see that they are deformed due to the curve of the surface of the face. Moreover, both patterns became misshaped in detail due to the unevenness of the surface of the face or the absorption of light by the skin. Figure 6(a) shows the signal pattern in the G-component. Binary data of "1" and "0" were arranged in turn. The line where the phase of the pattern is the same as that of the reference pattern in the B-component shown in Fig. 6(b) indicates data of "1" and the line where this is opposite indicates "0."

4.1 Readability

Figures 7 and 8 plot the experimental results for the accuracy with which the embedded data were read out. The accuracy is indicated by the percentage of data that was read out correctly from the entire amount of data. The data indicated by the open circles and squares are averaged from those obtained from the patterns projected onto the five people's faces and those indicated by the closed circles and squares are those obtained from the patterns projected onto the manikin's face. We used 40 lines of data for each person's and manikin's face to calculate accuracy; therefore, there were 200 measurements for each mark for people's faces and 40 for the manikin's face in Figs. 7 and 8. Figure 7 plots the dependence of accuracy on the number of pixels, N,



(a) Signal pattern (G-component)

(b) Reference pattern (B-component)

Fig. 6 Example of images of patterns projected onto the cheek of one of five people.



Fig. 7 Accuracy with which embedded data were read out. Dependence of accuracy on number of pixels, *N*.

and Fig. 8 plots the dependence of accuracy on the high-frequency component, HC.

We identified four main findings from Figs. 7 and 8: (1) Accuracy was very high for small numbers of N, (2) it was poor for small HC, (3) it was poorer for data from the forehead than those from the cheeks of human faces, and (4) it was poorer for the human faces than that for the manikin.

The reason for result (1) that accuracy was excellent for small numbers of N and poor for large numbers was because the results from calculations were largely affected by the object surface being deformed when the area used for the calculations was elongated. As N decreased, on the other hand, the frequency components in Eqs. (1) and (2) decreased, and this caused a decrease in accuracy. However, the results revealed that accuracy for small N was very high. The human face had uniform characteristics with regard to its image signal and it did not have numerous high-frequency components; therefore, the values were almost all obtained



Fig. 8 Accuracy with which embedded data was read out. Dependence of accuracy on high-frequency component, HC.

from the projected pattern. These were the reasons that accuracy was very high for small N.

The reason for result (2) that accuracy was poor for small HC was what we had expected because the small frequency component meant the watermarking pattern had low contrast and it was less detectable.

Result (3) was unexpected because the forehead is relatively flat. The reason for result (3) was that strong light was reflected in the direction of the reflex angle on the surface of the human face similarly to that with mirror reflections due to sebum on the human face. Therefore, if the camera is in this direction, the area where light is reflected is so bright that the brighness level is saturated and the pattern there disappears. The faces in this experiment were turned toward the projector, i.e., the surface of the forehead was perpendicular to an incident light from the projector, and the camera was in the same direction as that of the projector. Therefore, the reflected light went in the direction of the camera. Therefore, this was



Fig. 9 Results for redundancy method.

not due to the forehead but due to the location the light from the projector was reflected toward the camera.

The reason for result (4) that the manikin's face had higher accuracy than that of the human faces was that its reflectance had more uniform characterisitics than that of human faces because it was made of resin although both had similar 3-D shapes. Moreover, part of the light did not reflect from the surface of the skin but it was reflected in the skin, and this was considered to be part of the reason for result (4) because this caused a decrease in contrast of the projected pattern.

Figure 9 plots the results for the redundancy method. We chose 8 and 20 as the number of pixels, N, in this experiment because if we used large numbers of N, three different parts on a vertical pixel line would overlap because the vertical area that was cut out was 100 pixels, as shown in Fig. 2. We can see from Fig. 7 that accuracy can be improved and the highest degree of accuracy of 100% in reading out the

embedded data are possible for N = 8 and HC of >60 from the cheeks.

The results mentioned above are those obtained under the condition where there was no other light source. In practice, the method will probably be used in rooms where there is other lighting than that for this technique, e.g., lighting from the ceiling. We need to consider how such lighting will affect readability. Conventional lighting just raises the averaged brightness on the surface of the object. This means a decrease in the contrast of the pattern. Therefore, we can estimate readability using the dependence of readability on HC that indicates contrast by establishing the increase in brightness on the surface of the object and calculating the decrease in contrast, as shown in Figs. 7 or 8. Brightness due to illumination for special objects is usually much brighter than that due to illumination for rooms especially on the surface of special objects. This is typically over 10 times. Therefore, great accuracy of close to 100% can be retained if we choose adequate conditions for N and HC.

4.2 Invisibility

Table 1 summarizes the results for invisibility of the projected watermarking patterns evaluated with subjective tests. It can be seen that the projected patterns were invisible for HCs of <70.

The projected pattern pitch was 2 min in the angle of view and this was over the resolving power of the human-visual system. The reasons that the projected watermarking patterns were invisible, even though their pattern pitches were over the resolving power of the human-visual system, were considered to be as follows. First, the contrast in the original patterns was low, i.e., it was 30% even when HC was 60. Second, as mentioned before, part of the light did not reflect from the surface of the skin but it was reflected in the skin, and this decreases the contrast of the projected pattern on the human face. As a result, it became less visible.

We could confirm the feasibility of the proposed technique from the results given in Secs. 4.1 and 4.2.

Based on the results and discussion of readability and invisibility mentioned above, some requirements for the light source were clarified. First, we need more resolution. Ideally, it would be best if the pixel density of the projected image on the surface of the object was half that of the captured image from the point of view of satisfying both the

 Table 1
 Results for invisibility of projected watermarking patterns obtained from subjective tests. A–E indicate subjects.

	Α	в	с	D	E	
HC = 20	×	×	×	×	×	
30	×	×	×	×	×	
40	×	×	×	×	×	
50	×	×	×	×	×	
60	×	×	×	×	×	
70	0	0	×	×	0	

conditions of readability and invisibility. As recent digital cameras have ~ 4000 pixels in the x-direction, the light source needs to have ~2000 pixels. This requirement can be satisfied as recent high-end model projectors have more than 2000 pixels.

The bandwidth of the spectra of each light component is important because those of the camera usually overlap. This means that the projected G-component pattern is captured in the R- and B-component image in addition to the G-component. Moreover, since objects are 3-D shaped, an extended depth of field is required. It needs to be >1 m, although this depends on applications.

5 Conclusion

We proposed 1-D optical watermarking to protect the portrait rights of 3-D shaped real objects. Focusing on protecting the portrait rights of human faces, we conducted an experiment using human faces and a manikin's face as real 3-D objects in practice. We used a method of phase difference where two out of R-, G-, and B-color components were used and binary information was expressed as to whether the phase of the high-frequency pattern was the same or its opposite. The experimental results demonstrated that this technique was robust to the pattern being deformed due to the curved surface of 3-D shaped objects, and the highest degree of accuracy of 100% in reading out the embedded data was possible by using the redundancy method. Moreover, we demonstrated that the patterns were invisible. As a result, we could confirm that the proposed technique is feasible to protect the portrait rights of 3-D shaped real objects.

Acknowledgments

This work was supported by JSPS Kakenhii (No. 23650055).

References

- 1. I. J. Cox et al., "Secure spread spectrum watermarking for multimedia,"
- D. Score and Special spectral spectral water matching for mathine dia, *IEEE Trans. Image Process.* 6(12), 1673–1687 (1997).
 M. D. Swanson et al., "Multimedia data-embedding and watermarking technologies," *Proc. IEEE* 86(6), 1064–1087 (1998).
 M. Barni et al., "A DCT-domain system for robust image watermark-ing," *Signal Process.* 66(3), 357–372 (1998).
 J. Bitse, "A method for watermark acting on disited image," *IEEE*
- I. Pitas, "A method for watermark casting on digital image," *IEEE Trans. Circ. Syst. Video Technol.* 8(6), 775–780 (1998).

- M. Hartung and M. Kutter, "Multimedia watermarking techniques," *Proc. IEEE* 87(7), 1079–1107 (1999).
 F. Y. Shih and S. Y. Wu, "Combination image watermarking in the spa-
- tial and frequency domain," Pattern Recognit. 36(4), 969-975 (2003).
- T. Mizumoto and K. Matsui, "Robustness investigation of DCT digital watermark for printing and scanning," Trans. IEICE A J85-A(4), 451-459 (2002), in Japanese
- 8. M. Ejima and A. Miyazaki, "Digital watermark technique for hard copy image," Trans. IEICE A J82-A(7), 1156-1159 (1999), in Japanese.
- K. Uehira and M. Suzuki, "Digital watermarking technique using brightness-modulated light," in *Proc. IEEE 2008 Int. Conf. on Multimedia and Expo*, pp. 257–260, IEEE, Hanover, Germany (2008)
- 10. Y. Ishikawa, K. Uehira, and K. Yanaka, "Practical evaluation of illumination watermarking technique using orthogonal transforms," *IEEE/* OSA J. Disp. Technol. 6(9), 351–358 (2010).
- 11. Y. Ishikawa, K. Uehira, and K. Yanaka, "Optical watermarking technique robust to geometrical distortion in image," in Proc. IEEE Int. Symposium on Signal Processing and Information Technology, IÉEÉ, Bilbao, Spain (2011).
- O. Matoba et al., "Optical techniques for information security," *Proc. IEEE* 97(6), 1128–1148 (2009).



Mizuho Komori received her BS and MS in engineering from Kanagawa Institute of Technology, Japan. She studied digital watermarking in her master's course. She joined NTT Advanced Technology in 2012.



Kazutake Uehira received his BS, MS, and PhD in electronic engineering from Osaka Prefecture University, Japan, in 1979, 1981, and 1990. Dr. Uehira joined the Electrical Communication Laboratories of NTT in 1981, where he started work on RD in video communication systems and IO systems including display systems. Dr. Uehira was appointed a full professor of Kanagawa Institute of Technology in Atsugi, Japan, in 2001. He is a fellow of the Institute of Image Information and Television Engineers.