A Fast and Robust Digital Watermark Detection Scheme for Cellular Phones

Takao Nakamura†, Atsushi Katayama, Ryo Kitahara, and Kenji Nakazawa

Abstract

We describe a fast and robust digital watermark detection scheme for camera-equipped cellular phones that enables users to follow links embedded invisibly in printed photographs to related information on the Internet. The scheme consists of two processes: correcting geometric distortions of the captured image and detecting watermark information in the corrected image. We implemented this scheme as a Java application on a cellular phone. Experimental results showed that it could be performed in less than one second with sufficient robustness against geometric distortions and noise, which satisfies the requirements for practical use.

1. Introduction

Digital watermarking technology is used to embed additional information into content imperceptibly and then extract that information later. Although watermarking technology is primarily used for copy protection and security, if we bind a content identification (ID) to the content, it can be used to enable end users to get information related to the content via the Internet [1]. Watermark detection generally requires a lot of processing power, so personal computers (PCs) are usually used as the terminals. However, mobile terminals equipped with a camera, such as cellular phones, would be more convenient for users because they could easily retrieve related information anywhere and anytime (**Fig. 1**). Aiming to provide this convenience, we initially designed a client/server system that required the user to transmit an image captured by a cellular phone to the server so that the image processing could be performed there [2], [3]. However, it is expensive for users to transmit the large amount of data, so a more-cost effective scheme was required.

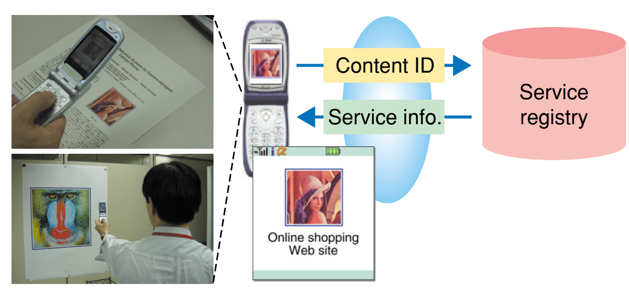

Fig. 1. System for providing related information through digital watermarking technology.

In this paper, we describe a watermark detection scheme for camera-equipped cellular phones that is both fast and robust against geometrical distortions and noise. The scheme consists of two processes: correcting geometric distortions of the captured image and detecting watermark information in the corrected image. Experimental results confirmed the scheme's efficiency.

2. Requirements

2.1 High speed

There are two approaches to watermark detection

† NTT Cyber Space Laboratories
Yokosuka-shi, 239-0847 Japan
E-mail: nakamura.takao@lab.ntt.co.jp

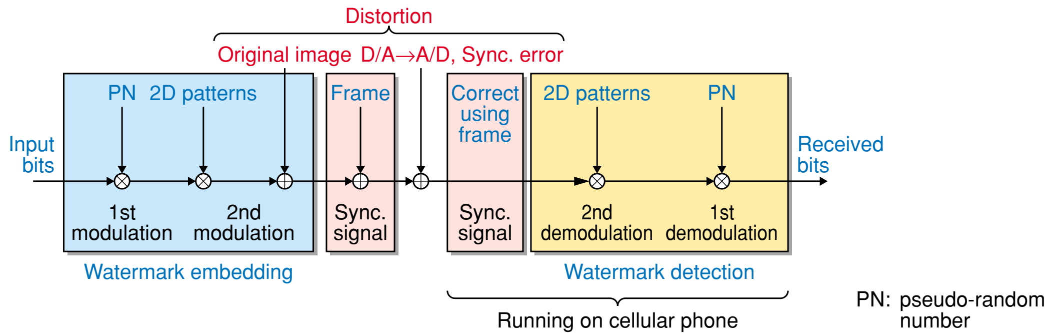

Fig. 2. Model of the watermarking scheme.

on cellular phones. One is to include the detection process as a built-in feature of the design. This provides a higher processing speed, but involves more preparatory work and will thus delay the commercial introduction of the product. The other approach is to implement the detection process as loadable software on existing terminals, such as an "i-appli" [4]. In this case, detection takes more time, but from a business strategy perspective, uncertainties over a product's marketability have led to businesses tending to prefer i-appli implementation, at least initially, because it is a faster way to provide services. Therefore, we selected i-appli implementation as the basis for our scheme. Because previous studies have shown that the time delay must be less than 2 seconds to be acceptable for human interfaces [5], we chose this as the target time from taking a picture to obtaining the detection result.

2.2 Robustness

Capturing images on analog media using a cellular-phone camera introduces distortion caused by 1) digital-to-analog (D/A) conversion, such as degradation due to printing, 2) analog-to-digital (A/D) conversion, such as degradation due to camera performance, and 3) geometric transformation, such as perspective distortion due to the shooting angle. Consequently, the watermarking scheme must be robust against all these causes of distortion.

3. Watermarking scheme

3.1 Design policy

A model of our watermarking scheme is shown in **Fig. 2**. Considering the watermarking technology as a communication model, we must transmit modulated watermark information through a channel with severe distortion, i.e., with (a) additive and quantization noise due to D/A and A/D conversion, (b) severe synchronization errors such as perspective transformation caused by the shooting angle, and (c) small synchronization errors such as phase distortion due to lens aberration and a slight curvature of the printed material, and/or errors resulting from the perspective distortion correction process. We cope with the severe synchronization errors by detecting a quadrilateral region in the captured image and recovering synchronization. By applying 2D pattern modulation with stable characteristics to each small block, we also cope with the small synchronization errors. Furthermore, we designed our scheme to be robust against additive and quantization noise by using the spread spectrum as the first-order modulation [6], [7].

3.2 Watermark embedding

The watermark embedding algorithm is shown in **Fig. 3**. The input parameters for the embedding process are the original image I, k-bit watermark information ID, and embedding strength a that controls the balance between the image quality and robustness.

Step E1 Using error correction/detection coding (ECC/EDC) for the watermark information ID, we generate an n-bit codeword.

Step E2 Multiplying the codeword by a pseudo-random number (PN) sequence, we modulate the n-bit codeword into the sequence $\{t_i\}$ with predetermined length N^2 (direct sequence spread spectrum (DS-SS) modulation).

Step E3 Depending on the value of each element of the sequence $\{t_i\}$, a block pattern is selected out of the two 2D sine curves with 90° rotational symmetry; hence, a block sequence with length N^2 is generated. Then, the block sequence is rearranged to form an $N \times N$ block matrix P that is the same size as the original image I. Finally, we obtain a watermark pattern as P, whose watermark signal can be distributed over the whole original image I.

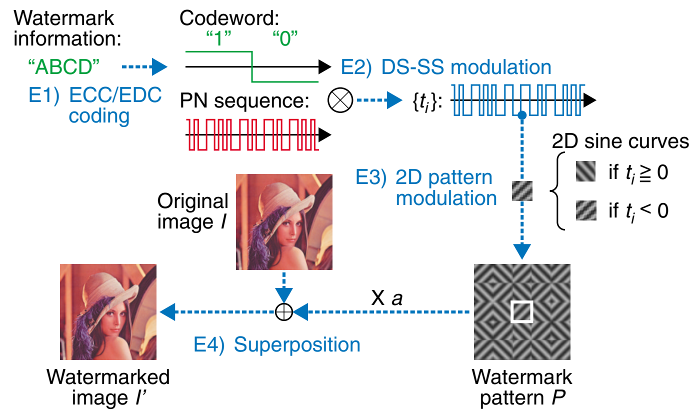

Fig. 3. Watermark embedding process.

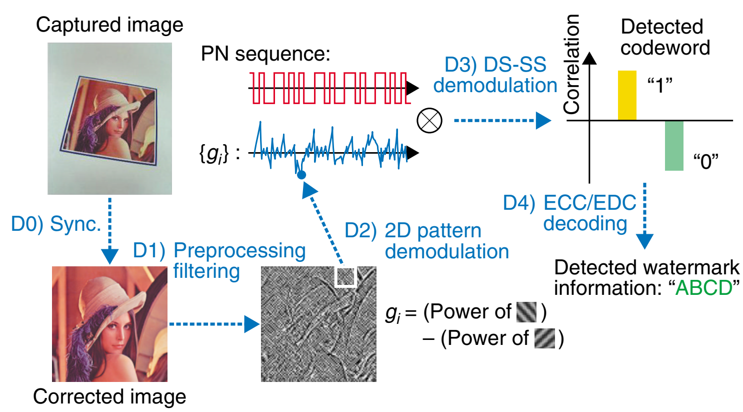

Fig. 4. Watermark detection process.

Step E4 The watermark pattern *P* is multiplied by embedding strength *a*, and the watermarked image *I'* is obtained by superposing *P* on the original image *I*. We can also apply adaptive pattern superposition to improve the balance between the image quality and the robustness of the watermark [7].

3.3 Watermark detection

After the watermarked image has been generated, a frame is placed around it to facilitate the correction of the geometric distortion that occurs during photography [8]. Then, the framed image is output to an analog medium, for example, by printing it. The frame improves usability because it acts as both a calibration pattern and as an indicator that this image contains a watermark and that related information can therefore be retrieved. The operation flow for watermark detection is shown in **Fig. 4**.

Step D0 Geometric distortion correction, that is synchronization, must be performed before the watermark can be detected. First, we must find four corners to estimate the parameters of the perspective transformation as follows (**Fig. 5**).

(1) Search for the edge points from the midpoints on each side of the captured image, moving towards the center.

(2) When the edge point has been found, trace a line perpendicular to the search direction.

(3) Set candidate line segments between the end points of each traced edge, while ignoring traced edges that are shorter than the threshold.

(4) Obtain all intersections of horizontal and vertical candidate line segments, assuming them to be endless lines. Based on the corner connectivity, select line segments that have both end points near the intersection and that can be formed into quadrilaterals. Finally, select the largest quadrilateral out of the ones formed.

By giving a four-point template that corresponds to the four corners of a detected quadrilateral, we can get a normalized target image by using inverse perspective transformation.

Our method finds the side lines of a quadrilateral

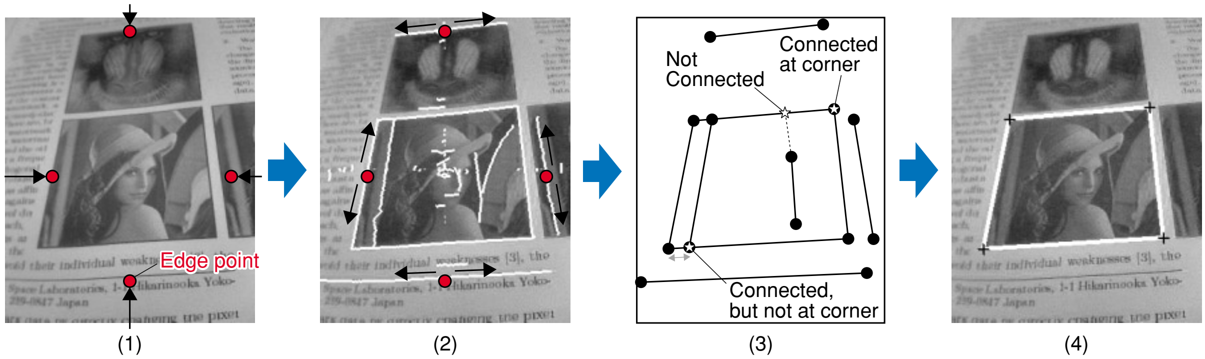

Fig. 5. Quadrilateral detection process.

using the contrast between the background and the edges of the target image. Therefore, the additional frame mentioned above is not necessary if the contrast is sufficiently large. However, the frame ensures that quadrilateral detection is still robust even when the contrast is insufficient, e.g., for a bright image on white paper. Our method is a heuristic approach that depends on the fact that users tend to center and enlarge the target framed image with the camera zoom. This heuristic approach makes it possible to detect quadrilaterals by referring to only a few pixels (typically less than one tenth of the whole captured image), which contributes to faster processing because access to the image frame buffer is a time-consuming process for an i-appli [8].

Step D1 The preprocessing filter emphasizes the watermark pattern in the corrected image.

Step D2 The filtered image is divided into $N \times N$ blocks. For each block, we calculate the energy difference of two frequencies that correspond to the two 2D sine curves. Then, each element of a sequence $\{g_i\}$ with length of N^2 is determined based on this difference. This process is robust against small geometric distortions of the image because the energy of sine curves is stable against small geometric shifts.

Step D3 By applying DS-SS demodulation to the sequence $\{g_i\}$, we can obtain the detected codeword.

Step D4 ECC/EDC decoding is applied to the detected codeword. Then, correctable bit errors are properly corrected, and the watermark information ID of the k-bits is decoded.

3.5 Reliability of watermark detection

In general, one of the most important requirements for watermarking technology is to have a low probability of false positive detection errors [7], [9], which consist of the following two cases:

A) Detection is reported when the image does not contain watermark information and

B) Detection is successful, but the decoded information is incorrect.

To meet this requirement, we developed a design theory based on statistical analysis of the probability distribution of the correlation value at spectrum spreading demodulation [9]. In the analysis, we applied a statistical test for type A errors and examined the bit error rate distribution and the characteristics of ECC/EDC for type B errors. Based on the analysis, we introduced a detection reliability index that enables us to quantitatively guarantee the detection reliability for both type A and B errors using a threshold test of the index value. Because evaluation for type A errors corresponds to the existence of the watermark, the application software should prompt the user to retake the picture if the threshold test is successful for type A errors but fails for type B errors.

4. Experimental results

To validate the effectiveness of the scheme, we implemented it as an i-appli and conducted various experiments. For the error correction/detection code, we used (31 + 1, 16) extended BCH code (2-tuple bit error correction). Therefore, the length of the watermark information k was 16 bits, the code length n was 32 bits, and the resolution of the captured image on the cellular phone was 288×352 pixels.

4.1 Processing speed

The total processing time was about 0.7 seconds on average, which satisfies the speed requirements mentioned in Section 2.1.

4.2 Robustness

4.2.1 Quadrilateral detection

The robustness of quadrilateral detection against

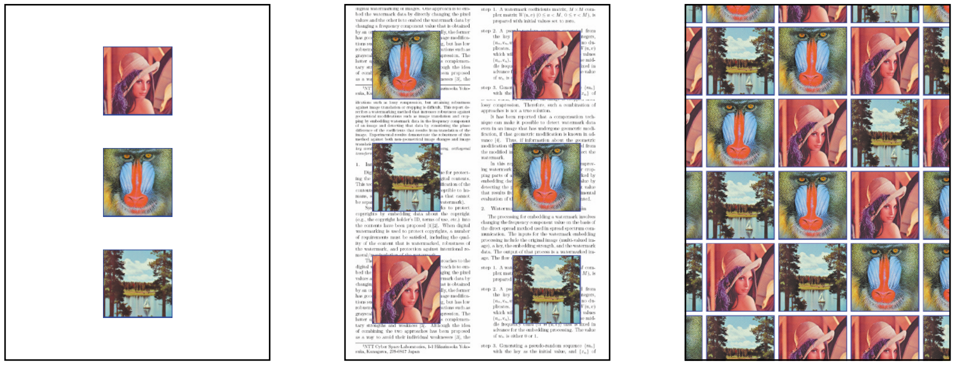

Fig. 6. Printed materials used to evaluate quadrangle detection (21 cm × 29 cm).

(a) (b) (c) (d)

Fig. 7. Examples of detection results. (a) and (b): both quadrangle detection and watermark detection were successful. (c) and (d): quadrangle detection was successful, but watermark detection failed.

surrounding noise, such as letters or other images near the target, was evaluated using the three printed material layouts shown in **Fig. 6**. The watermarks were embedded into all images with the maximum strength γ. The only instruction given to test subjects was to capture the entire target image within the photographable area. The success rate for detecting both the quadrilateral and the watermark was 97%, without any significant difference among the layouts. The images used in this experiment are shown in **Fig. 7**. Figures 7(a) and (b) are examples for which watermark detection was successful, while (c) and (d) are examples for which it failed. The failure was caused by image blurring due to hand movements (Fig. 7(c)) or severe perspective distortion (Fig. 7(d)). We checked these images and found that the quadrilaterals could be detected, even though watermark detection failed. In other words, our quadrilateral detection method has sufficient robustness.

4.2.2 Watermark detection

We used three images (Lena, Baboon, and Sailboat) with three levels of embedding strengths a ($a = \alpha, \beta$, or γ, where $\alpha < \beta < \gamma$). The peak signal-to-noise ratios (PSNRs) between the original and watermarked images are shown in **Table 1**. The framed watermarked images with $a = \beta$ are shown in **Fig. 8**. Even though the same embedding strength was used in all the images, there were differences among the PSNRs due to the adaptive embedding process described in Section 3.2.

We conducted a robustness experiment by capturing the printed images with a camera-equipped cellular phone. We used a color laser printer and printed watermarked images 5 cm × 5 cm. To determine the conditions for successful detection, we used the detection reliability index in Ref. 9. When the index value was greater than the threshold and no error was detected by the ECC/EDC decoding process, the detection was deemed successful. We chose a threshold value that could guarantee that the probability of type A errors would be less than 3×10^{-10} and type B errors less than 0.001.

In the first stage, we gave no instructions to the test subjects, except to capture the entire frames because

Table 1. PSNRs for each embedding strength.

	α	β	γ
Lena	39.1	37.8	36.7
Baboon	33.2	32.0	30.9
Sailboat	37.1	35.8	34.7

Fig. 8. Watermarked images (from left to right: Lena, Baboon, Sailboat).

this is necessary for quadrilateral detection. The results of watermark detection are given in the middle column of **Table 2**. We found that detection easily failed when there was motion blurring due to camera movement, and/or when the degree of perspective distortion was large, as in Figs. 7(c) and (d).

Next, the test subjects were instructed to capture the image from a position directly in front of them to decrease perspective distortion. As shown by the results in the right column of Table 2, detection accuracy improved in this case.

From these results, for an image with $a = \beta$, captured by the test subjects without any special care, the probability for detection failure was 11% for Lena, 6% for Baboon, and 16% for Sailboat. However, when the test subjects were instructed to re-capture from directly in front of the image, they successfully detected the watermark information in all cases.

No cases of type B errors occurred throughout the entire set of experiments of 1680 trials.

Table 2. Detection success rates.

	a	Without instruction	With instruction
Lena	α	81%	100%
	β	89%	100%
	γ	98%	100%
Baboon	α	90%	100%
	β	94%	100%
	γ	99%	100%
Sailboat	α	48%	90%
	β	84%	100%
	γ	93%	100%

5. Conclusion

Our fast and robust digital watermark detection scheme for camera-equipped cellular phones utilizes geometric distortion correction employing frames and a watermarking algorithm that is robust against small geometric distortions and noises. We implemented this scheme as a Java application on cellular phones. In experiments, we found that the scheme enabled users to detect embedded information in less than one second with sufficient robustness. The results confirmed that this scheme satisfies the requirements for practical use. A commercial service using this scheme has just started [10]. In the future, we will analyze the image dependence of the scheme's robustness for various analog output methods, such as commercial printing, CRTs (cathode ray tubes), and LCDs (liquid crystal displays). By analyzing the results, we hope to improve both visual quality and robustness.

References

[1] H. Sakamoto, H. Fujii, S. Irie, and H. Yamashita, "Service Gateway to Enable the Introduction of Content Related Services," Proc. IEEE International Conference on Multimedia and Expo 2001 (ICME2001), pp. 637-640, 2001.

[2] T. Nakamura, A. Katayama, H. Miyachi, H. Yamashita, and M. Yamamuro, "A Watermark Detection Scheme for the Service Offering System using Camera-equipped Mobile Phone," Proc. Forum on Information Technology 2003 (FIT2003), N-020, pp. 409-410, 2003.

[3] "CyberSquash: Internet Access Platform using Digital Watermarks-Enables Users to Access Related Information Easily by Reading Printed Images," NTT Technical Review, Vol. 1, No. 6, pp. 96-97, 2003.

[4] NTT Docomo, Inc., "Specifications of Java for i-mode," http://www.nttdocomo.co.jp/english/p_s/imode/sites/i_appli/index.html

[5] R. B. Miller, "Response time in man-computer conversational transactions," Proc. AFIPS Fall Joint Computer Conference, Vol. 33, pp. 267-277, 1968.

[6] W. Bender, D. Gruhl, and N. Morimoto, "Techniques for Data Hiding," Proc. SPIE Conference on Storage and Retrieval for Image and Video Databases III, Vol. 2420, pp. 164-173, Feb. 1995.

[7] T. Nakamura, H. Ogawa, A. Tomioka, and Y. Takashima, "Improved Digital Watermark Robustness against Translation and/or Cropping

of an Image Area," IEICE Trans. Fundamentals, Vol. E83-A, No. 1, Jan. 2000.

[8] A. Katayama, T. Nakamura, M. Yamamuro, and N. Sonehara, "New High-speed Frame Detection Method: Side Trace Algorithm (STA) for i-appli on Cellular Phones to Detect Watermarks," Proc. Mobile And Ubiquitous Multimedia 2004 (MUM2004), pp. 109-116, 2004.

[9] T. Nakamura, A. Katayama, M. Yamamuro, and N. Sonehara, "Fast Watermark Detection Scheme for Camera-equipped Cellular Phone," Proc. Mobile And Ubiquitous Multimedia 2004 (MUM2004), pp. 101-108, 2004.

[10] NTT DATA Inc. and NTT Communications, Inc., "Pattobi," http://pattobi.jp, 2005.

Takao Nakamura

Research Engineer, Visual Media Communications Project, NTT Cyber Space Laboratories.

He received the B.S. degree in mathematics from Waseda University, Tokyo in 1994 and joined NTT Human Interface Laboratories, Kanagawa in 1994. He received the FIT2003 Young Researcher Award. He is a member of the Information Processing Society of Japan.

Atsushi Katayama

Senior Research Engineer, Visual Media Communications Project, NTT Cyber Space Laboratories.

He received the B.E. and M.E. degrees in mechanical engineering from Kyushu University, Fukuoka in 1984 and 1986, respectively. He joined NTT Integrated Communication Laboratories, Kanagawa in 1986. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan, the Japan Society of Mechanical Engineers, and the Robotics Society of Japan.

Ryo Kitahara

Research Engineer, Visual Media Communications Project, NTT Cyber Space Laboratories.

He received the B.E. and M.E. degrees in information engineering from Tohoku University, Miyagi in 2001 and 2003, respectively. He joined NTT Cyber Space Laboratories, Kanagawa in 2003.

Kenji Nakazawa

Senior Research Engineer, Visual Media Communications Project, NTT Cyber Space Laboratories.

He received the B.E., M.E., and Ph.D. degrees in materials sciences and engineering from Kanazawa University, Ishikawa in 1981, 1983, and 1991, respectively. He joined the Musashino Electrical Communications Laboratories of Nippon Telegraph and Telephone Public Corporation (now NTT) in 1983. He is a member of IEICE, the Institute of Image Information and Television Engineers of Japan, and the Society for Information Display.