Information Embedment with Cross Ratio of Areas for Accurate Camera-Based Character Recognition

Shinichiro Omachi†, Masakazu Iwamura††, Seiichi Uchida†††, and Koichi Kise††

†Graduate School of Engineering, Tohoku University
Aoba 6-6-05, Aramaki, Aoba-ku, Sendai-shi, 980-8579 Japan
E-mail: machi@ecei.tohoku.ac.jp
††Graduate School of Engineering, Osaka Prefecture University
1-1 Gakuen-cho, Naka-ku, Sakai-shi, Osaka, 599-8531 Japan
E-mail: {masa, kise}@cs.osakafu-u.ac.jp
†††Faculty of Information Science and Electrical Engineering, Kyushu University
744 Motooka, Nishi-ku, Fukuoka-shi, 819-0395 Japan
E-mail: uchida@is.kyushu-u.ac.jp

Extended Abstract

With the widespread use of digital cameras and camera phones, character recognition in scene images has become an important topic in the field of character recognition. Low resolution, uneven lighting condition and geometrical distortion make it difficult to recognize characters in real scene images accurately[1]. In order to achieve a highly accurate recognition of characters in scene images, there are some attempts on offering supplementary information for recognition[2] with a character image. The information should be human-friendly and robust against geometrical distortions. Uchida et al. proposed a method for embedding category information into character images using stripes[3]. They also proposed an algorithm for extracting the information stably with a finite state automaton[4]. Omachi et al. proposed a method for embedding information as a ratio of areas[5]. Since the ratio of areas is affine-invariant, even if a character image is geometrically distorted, theoretically the embedded information will be extracted correctly as long as the distortion can be regarded as an affine transformation. However, in general, geometrical distortions of character images taken by cameras are not affine transformations but projective transformations.

In this paper, we propose a method for embedding supplementary information into a character pattern as a cross ratio of areas by designing a character pattern in two colors. It is known that a cross ratio defined by four points on a line is projective-invariant[6], and a cross ratio of areas is also projective-invariant. A character image with a shadow is shown in Figure 1(a). The dark-grey part represents the character and the light-grey part is the shadow. In order to define a cross ratio of areas, we need at least three regions. Therefore, we use the character part, the shadow part and the convex full of the character part. Let the areas of the character part and the shadow part be *s* and *t*, respectively. The convex hull of the character part is shown as the trapezoid in Figure 1(b). Let the area of the convex hull excepting the character part be *u* as shown in Figure 1(c). The cross ratio of areas can be defined as $r = t(s+t+u)/(s+t)(t+u)$. We embed supplementary information by designing a character pattern so that *r* is the value of the information. Figure 2 shows an example of calculating cross ratios. The ratio of the pattern of Figure 2(a) is 0.907. Figure 2(b) shows a projective-transformed pattern of Figure 2(a). The ratio of this pattern is 0.906. Except for the quantization error, the cross ratio is projective-invariant.

Figure 3 shows an example of the character patterns where the information is embedded into each character pattern as the cross ratio of areas. The value of the ratio starts at 0.20 and increases by 0.01. Therefore the ratio of “A” is 0.20 and that of “Z” is 0.45. Figure 4 shows the quantization error for embedding the information using four fonts. Since the maximum error was 0.005 and the interval of the values was 0.01, the information can be extracted correctly. Using the generated pattern, a simulation experiment of recognizing the character patterns only by extracting the embedded information was carried out. The character patterns were transformed projectively, and the embedded information was extracted from each pattern. According to the extracted value, the category of the pattern was determined. Figure 5

shows the result. The value of δ is the degree of deformation defined in [3] and the horizontal axis shows the interval of the embedded values. The figure clarifies that if the degree of distortion is small and the interval of the embedded value is large, the embedded information is correctly extracted and the character pattern is correctly recognized.

Key words Camera-based character recognition, Supplementary information, Information embedment, Cross ratio.

References

[1] J. Liang, D. Doermann, and H. Li, "Camera-based analysis of text and documents: a survey," *International Journal of Document Analysis and Recognition*, vol. 7, nos.2-3, pp. 84-104, 2005.

[2] M. Iwamura, S. Uchida, S. Omachi, and K. Kise, "Recognition with Supplementary Information —How Many Bits Are Lacking for 100% Recognition?—," In *Proc. of the First International Workshop on Camera-Based Document Analysis and Recognition (CBDAR2005)*, pp. 68-75, Seoul, Korea, 2005.

[3] S. Uchida, M. Iwamura, S. Omachi, and K. Kise, "OCR Fonts Revisited for Camera-Based Character Recognition," In *Proc. of 18th International Conference on Pattern Recognition (ICPR 2006)*, vol. 2, pp. 1134-1137, Hong Kong, China, 2006.

[4] S. Uchida, M. Sakai, M. Iwamura, S. Omachi, and K. Kise, "Extraction of Embedded Class Information from Universal Character Pattern," In *Proc. of the 9th International Conference on Document Analysis and Recognition (ICDAR 2007)*, vol. 1, pp. 437-441, Curitiba, Brazil, 2007.

[5] S. Omachi, M. Iwamura, S. Uchida, and K. Kise, "Affine Invariant Information Embedment for Accurate Camera-Based Character Recognition," In *Proc. of 18th International Conference on Pattern Recognition (ICPR 2006)*, vol. 2, pp. 1098-1101, Hong Kong, China, 2006.

[6] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2nd edition, 2004.

(a) (b) (c)

Figure 1. Information embedment into a character pattern.

(a) r =0.907 (b) r =0.906

Figure 2. Change in the cross ratio of areas.

Figure 3. Example of character patterns with supplementary information.

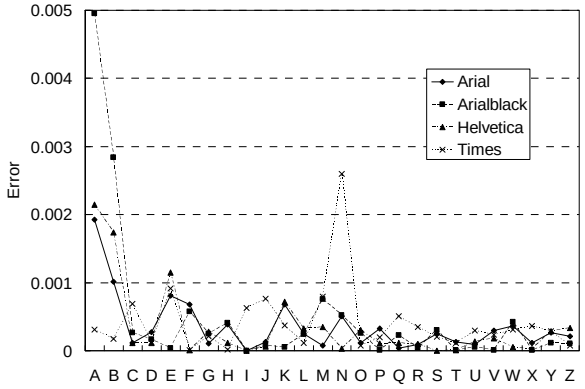

Figure 4. Error of the embedded value.

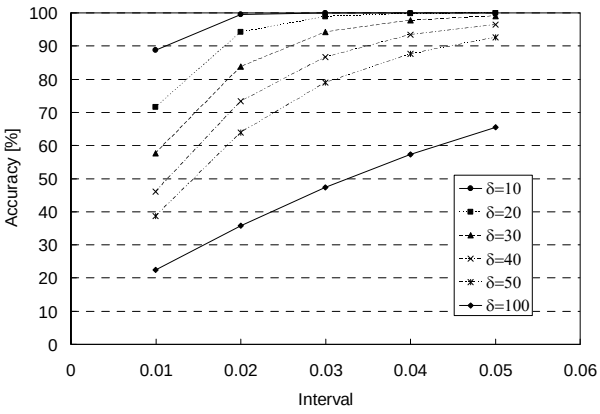

Figure 5. Accuracy of recognition by extracting the embedded information.