

# Extraction of Embedded Class Information from Universal Character Pattern

Seiichi Uchida  
Kyushu Univ., Japan

Megumi Sakai  
Kyushu Univ., Japan

Masakazu Iwamura  
Osaka Pref. Univ., Japan

Shinichiro Omachi  
Tohoku Univ., Japan

Koichi Kise  
Osaka Pref. Univ., Japan

## Abstract

This paper is concerned with a universal pattern, which is defined as a character pattern designed to have high machine-readability. This universal pattern is a character pattern printed with stripes. The cross ratio calculated from the widths of the stripes represents the character class. Thus, if the boundaries of the stripes can be detected for measuring the widths, the class can be determined without ordinary recognition process. Furthermore, since the cross ratio is invariant to projective distortions, the correct class will be still determined under those distortions. This paper describes a practical scheme to recognize this universal pattern. The proposed scheme includes a novel algorithm to detect the stripe boundaries stably even from the universal pattern image contaminated by non-uniform lighting and noise. The algorithm is realized by a combination of a dynamic programming-based optimal boundary detection and a finite state automaton which represents the property of the universal pattern. Experimental results showed the proposed scheme could recognize 99.6% of the universal pattern images which underwent heavy projective distortions and non-uniform lighting.

## 1. Introduction

In upcoming ubiquitous computing era, character patterns in 3D-scene will be often exposed to cameras and sometimes recognized by machines. Thus, character patterns can be promising media for ubiquitous computing, if they possess not only perfect human-readability but also high machine-readability. In this paper, the character pattern which possesses both readabilities is called *universal pattern*.

OCR/MICR fonts [1] are classical universal patterns. Their machine-readabilities, however, are not sufficient in the ubiquitous computing era; they were designed for scanners or magnetic readers and thus not for cameras. As listed in [2], character patterns acquired by cameras undergo various distortions which do not appear in character patterns acquired by scanners. For example, projective distortion and

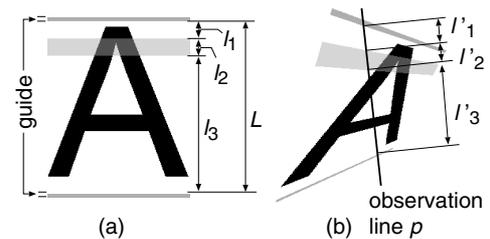


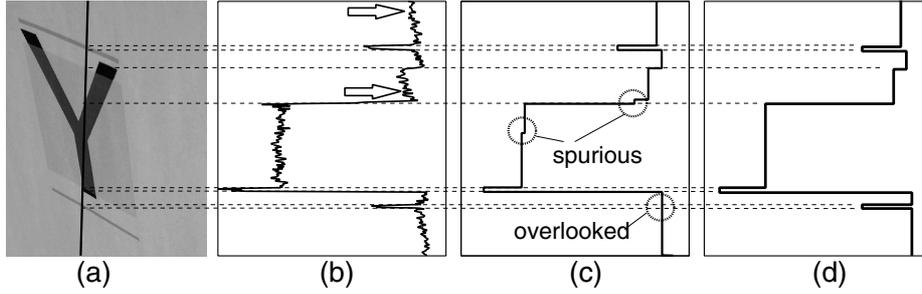
Figure 1. (a) The universal pattern “A” proposed in [3, 4]. (b) Observation line to extract class information.

non-uniform lighting will seriously disturb the recognition of not only ordinary character patterns but also OCR/MICR fonts.

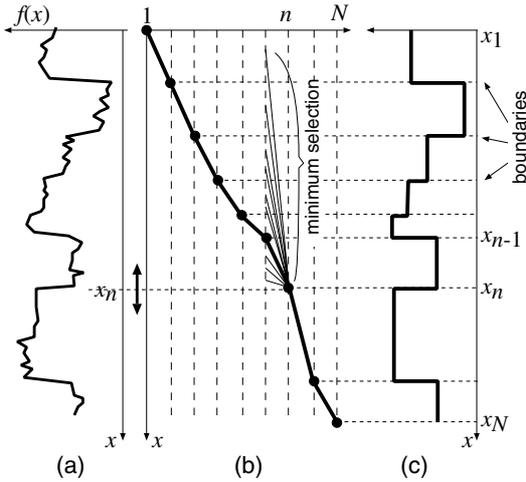
In [3, 4], a universal pattern has been proposed for camera-based recognition. Hereafter, the term “universal pattern” refers to the character pattern proposed in [3, 4], although we can consider another universal pattern such as [5]. Figure 1 (a) shows the universal pattern of [3, 4]. This universal pattern is a character pattern printed with a set of horizontal stripes, called a *cross ratio pattern*, to realize high machine-readability. The cross ratio calculated from the widths of the stripes represents the character class. This fact means the character class can be determined without ordinary recognition process if the cross ratio is extracted correctly. In addition, the correct class will be determined under projective distortions because the cross ratio is invariant to projective distortions.

This paper describes a practical scheme to recognize the universal pattern in real scene. For determining the widths of the stripes, the boundaries of the stripes should be detected. One of the main contributions of this paper is the proposal of a novel algorithm to detect the boundaries stably even from universal pattern images contaminated by non-uniform lighting and noise<sup>1</sup>. The algorithm is realized by a combination of dynamic programming (DP) and a finite

<sup>1</sup>Note that the trial in [3, 4] was a simulation experiment without any grayscale change and therefore the boundaries can be detected just by checking the grayscale value at each pixel.



**Figure 2. Contamination of grayscale value by non-uniform lighting and noise. (a) Character image and an observation line. (b) Grayscale value  $f(x)$  along the observation line. (c) A boundary detection result without FSA. (d) A result with FSA.**



**Figure 3. (a) Grayscale value  $f(x)$  along an observation line. (b) Search for optimal segmentation by DP. (c) Piecewise-constant approximation of  $f(x)$  and boundary sequence  $x_1, \dots, x_N$ .**

state automaton (FSA) which represents the property of the universal pattern. Note that the proposed boundary detection algorithm is rather general one and it can be applied to other various tasks. Another contribution is to report the results of a recognition experiment of universal pattern images really acquired by a camera. The results showed that even universal patterns with heavy projective distortions and non-uniform lighting could be recognized with a very high accuracy.

## 2. Universal Pattern for Camera-Based Character Recognition

### 2.1. Embedment of class information

The universal pattern proposed in [3, 4] is a character pattern printed with the cross ratio pattern of Fig 1 (a) to realize high machine-readability. The cross ratio pattern is comprised of five horizontal stripes. The top and the bot-

tom stripes are *guides*, which have a fixed width and define the beginning and the end of the cross ratio pattern. The remaining three stripes have variable widths,  $l_1$ ,  $l_2$ , and  $l_3$  and provide a *cross ratio*  $r$  by

$$r = \frac{(l_1 + l_2)(l_2 + l_3)}{l_2(l_1 + l_2 + l_3)}. \quad (1)$$

In this paper, we assume that all the universal patterns of class  $k$  are printed with the same cross ratio pattern with  $r = r_k$ . In addition, we also assume  $r_k \neq r_{k'}$  if  $k \neq k'$ . Under these assumptions, the class of a universal pattern can be determined, theoretically, just by extracting the cross ratio  $r_k$ .

### 2.2. Extraction of class information — Overview

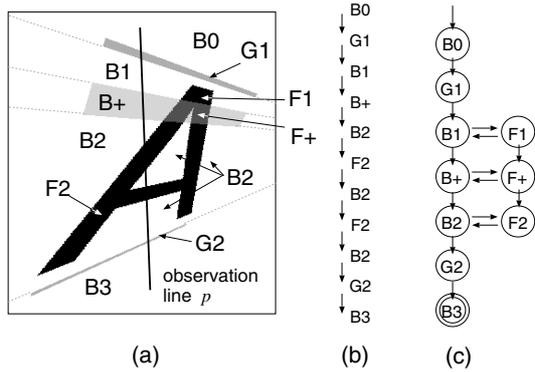
The cross ratio  $r_k$  can be obtained by drawing an *observation line*  $p$  on the universal pattern. As shown in Fig. 1(b), the observation line  $p$  passes across the cross ratio pattern. Since the stripes are parallel originally and the cross ratio is a projective invariant, we can obtain the original cross ratio  $r_k$  by using measured widths  $l'_1$ ,  $l'_2$ , and  $l'_3$  instead of  $l_1$ ,  $l_2$ , and  $l_3$  in (1) regardless of camera angle and the position and the slope of the observation line  $p$ .

The extracted cross ratio may be erroneous due to insufficient camera resolution, blurring, non-uniform lighting, etc. Thus, we use the following voting strategy for a robust estimation of  $r_k$ : (i) we draw the observation line  $p$  on the character pattern  $P$  times changing its position and slope randomly, (ii) obtain  $P$  cross ratio values, (iii) quantize each of those values into one of  $\{r_k\}$ , and (iv) choose the most frequent  $r_k$  as the cross ratio embedded.

## 3. FSA-Guided Boundary Detection for Extraction of Class Information

### 3.1. Optimization-based boundary detection

For measuring the widths  $l'_1$ ,  $l'_2$ , and  $l'_3$ , it is necessary to detect the boundaries on  $f(x)$  which represents the



**Figure 4. (a) Decomposition of the universal pattern into ten region types. (b) Transition of region types along the observation line in (a). (c) FSA representing all possible transitions.**

grayscale value at the position  $x$  ( $x_s \leq x \leq x_e$ ) on an observation line. Figure 2 (a) and (b) show a character pattern with an observation line and  $f(x)$  along the observation line. Simple threshold-based techniques are insufficient to detect the boundaries because the grayscale value on the observation line will be contaminated by non-uniform lighting and other noises. As shown in Fig. 2 (b), the originally different grayscale values pointed by two arrows become similar due to non-uniform lighting.

An alternative is an optimization-based boundary detection technique whose basic idea is optimal piecewise-constant approximation of  $f(x)$ . Figures 3 (a) and (c) show an example of  $f(x)$  and its piecewise-constant approximation. The ends of constant segments will indicate a boundary of the stripes.

The objective function of the approximation is formulated as

$$J_1 = \sum_{n=2}^N \int_{x_{n-1}}^{x_n} \|f(x) - \bar{f}_{x_{n-1}, x_n}\| dx, \quad (2)$$

and to be minimized with respect to the boundary sequence,  $x_1, \dots, x_n, \dots, x_N$ , subject to the conditions,  $x_{n-1} < x_n$ ,  $x_1 = x_s$ , and  $x_N = x_e$ . The value  $\bar{f}_{x_{n-1}, x_n}$  is the average grayscale value within the segment  $[x_{n-1}, x_n]$ . The minimization can be represented as an optimal path problem from  $x_1$  to  $x_N$  as shown in Fig. 3 (b) and solved by a DP algorithm efficiently. While we omit the details of the DP algorithm here, readers can find similar DP algorithms as recognition-based segmentation algorithms for handwritten text [6] and speech [7].

### 3.2. Incorporation of FSA

The above boundary detection technique still has two drawbacks for our task. First, it cannot distinguish the boundaries of the stripes from the boundaries between character stroke and background. (The latter boundaries are not

necessary to extract the cross ratio.) Second, the above technique still produces spurious boundaries due to noise and overlooks genuine boundaries.

Incorporation of an FSA is very useful to remedy those drawbacks since it can exploit the property of the universal pattern. For designing the FSA, we first consider ten regions, B0, B1, B2, B3, B+, F1, F2, F+, G1, and G2 of Fig. 4 (a). The first five regions correspond to background regions, the next three regions correspond to character stroke regions, and the remaining two regions correspond to the two guides. The region F+ has a different grayscale value from the regions F1/F2 due to the overlay of the middle stripe of the cross ratio pattern. Similarly, the region B+ has a different grayscale value from the regions B0/B1/B2/B3. Figure 4 (b) shows the transition of the regions along the observation line of Fig. 4 (a). It is easy to confirm that every possible transition like Fig. 4 (b) will be described as a state transition of the FSA of Fig. 4 (c).

We consider an incorporation of the FSA into the technique of 3.1. By the incorporation, we can assign a state (i.e., a region) to each segment and thus distinguish the necessary boundaries from the unnecessary boundaries. In addition, we can expect far better results as discussed later. The objective function after the incorporation of the FSA is defined as

$$J_2 = \sum_{n=2}^N \left[ \int_{x_{n-1}}^{x_n} \|f(x) - \bar{f}_{x_{n-1}, x_n}\| dx + q(x_n, x_{n-1} | s_n, s_{n-1}) \right], \quad (3)$$

where  $s_n$  represents the state to which the segment  $[x_{n-1}, x_n]$  corresponds. The objective function  $J_2$  should be minimized with respect to not only the boundary sequence  $x_1, \dots, x_N$  but also state sequence  $s_1, \dots, s_N$ , where the state transition  $s_{n-1} \rightarrow s_n$  should be allowed by the FSA.

The function  $q$  ‘‘penalizes’’ erroneous combinations of the two consecutive segments  $[x_{n-2}, x_{n-1}]$  and  $[x_{n-1}, x_n]$  and the two states  $s_{n-1}$  and  $s_n$ . Assuming  $s_{n-1} = B+$  and  $s_n = B2$ , the two segments should satisfy the condition  $\bar{f}_{x_{n-2}, x_{n-1}} < \bar{f}_{x_{n-1}, x_n}$ . This is because the region of type B+ should have a lower (darker) grayscale value than the region of type B2. (See Fig. 4(a).) If the condition does not hold, the function  $q$  returns  $\infty$  to prohibit this segmentation; otherwise  $q$  returns 0 to allow the segmentation. Thus, by the incorporation of the function  $q$ , we can exclude the segmentation which does not agree with the property of the universal pattern.

It should be noted that the function  $q$  checks the relative brightness of two consecutive segments and does not check any absolute brightness. The relative brightness is far more stable against non-uniform lighting and noise than the absolute brightness. The idea of this relative evaluation

is novel one and cannot be found in the conventional DP-based segmentation algorithms, such as recognition-based segmentation [6, 7].

The foregoing DP algorithm for the minimization of  $J_1$  with respect to  $\{x_n\}$  can be extended for the minimization of  $J_2$  with respect to  $\{(x_n, s_n)\}$ . Simply speaking, this extension can be done by re-organizing the  $x$ - $n$  search space of Fig. 3(b) to be an  $(x, s)$ - $n$  search space. After the optimal sequence of  $\{(x_n, s_n)\}$  is obtained,  $l'_1, l'_2$ , and  $l'_3$  are obtained using  $x_n$  related to stripe boundaries.

Details of the extended DP algorithm is omitted here except for the important point that  $\bar{f}_{x_{n-2}, x_{n-1}}$  in  $q$  is approximated as  $\bar{f}_{x_{n-1}-\epsilon, x_{n-1}}$ , where  $\epsilon$  is a positive constant. Without this approximation,  $q$  depends on  $x_{n-2}$  in addition to  $x_{n-1}$  and  $x_n$ . Thus, the minimization of  $J_2$  becomes a problem on a second-order Markovian process and requires prohibitive computations.

### 3.3. Estimation of the number of boundaries

The correct number of boundaries,  $N$ , is not known in advance. Thus, all possible  $N$ s should be examined; that is, the above algorithm is performed repeatedly for each  $N$ . Fortunately, the Markovian nature of the DP-algorithm allows us to obtain the optimal solution at  $N$  from that at  $N - 1$  with very few computations. If a certain  $N$  gives a boundary detection result which is acceptable by the FSA and provides the minimum  $J_2$ , this  $N$  is considered as the correct one.

## 4. Experimental Results

### 4.1. Experimental setup

The universal patterns of 26 English capital letters (Fig. 5) were printed on a paper. The original intensity values were: 255 for B0/B1/B2/B3, 215 for B+, 0 for F1/F2, 80 for F+, and 175 for G1/G2. Then they were acquired as grayscale images by a camera from 9 different angles as shown in Fig. 6. The lighting condition was not controlled with the angles and thus the images underwent the distortion by non-uniform lighting as well as the projective distortion. The average size of the universal pattern images was  $381 \times 291$  (pixels). Then,  $P = 81$  observation lines were drawn on each universal pattern image and the proposed boundary detection algorithm was applied to each observation line. Thus, there were  $26 \times 9 = 234$  universal pattern images and  $234 \times 81 \sim 19,000$  observation lines.

### 4.2. Qualitative evaluation of boundary detection result

Figures 2 (c) and (d) show the results of boundary detection on an observation line. The result (c) was obtained by



Figure 5. Universal patterns used in our experiment.

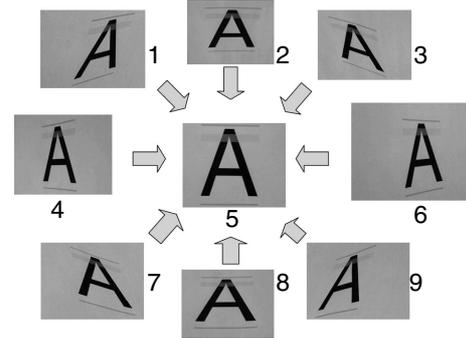


Figure 6. Camera angles.

solving the minimization problem of  $J_1$ . It is observed that two spurious boundaries were detected and the two boundaries of the lower guide were overlooked. In contrast, the proposed technique provided the result (d) and could avoid those spurious and overlooked boundaries. This improvement was achieved by the incorporation of the FSA; it could exclude inadmissible segmentation in (c) where the average grayscale values decreases gradually on five consecutive segments.

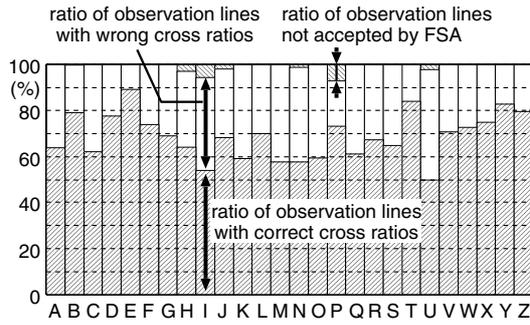
### 4.3. Quantitative evaluation of segmentation result

The ratio of the observation lines whose cross ratio  $r_k$  were extracted correctly was 69.2%. The extracted cross ratio becomes erroneous even if only a single boundary is detected wrongly and the remaining  $N - 1$  boundaries are detected successfully. Thus, the ratio of boundaries detected correctly will be far higher than 69.2%.

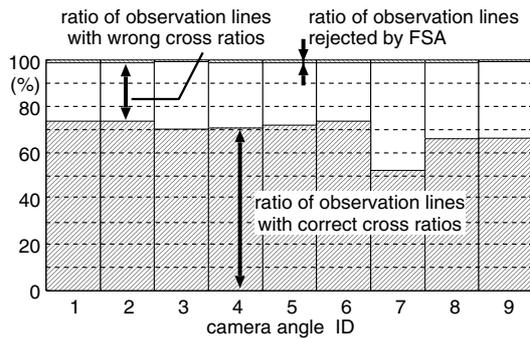
Figure 7 shows the correct detection ratio of each class and Fig. 8 shows the ratio of each camera angle. Both figures reveal that the correct extraction ratio is stable against character shapes and camera angles. Especially, Fig. 8 proves the invariance of the cross ratio against projective distortions.

The proposed technique failed to extract correct cross ratios due to the following three reasons.

- Failure to estimate the number of boundaries on the observation line. In fact, this failure was found in 19.1%



**Figure 7. Extraction accuracy of cross ratios for each class.**



**Figure 8. Extraction accuracy of cross ratios for each camera angle. Each angle ID (1~9) corresponds to the number in Fig. 6.**

of the 19,000 observation lines.

- Erroneous detection of boundaries. The boundary detected at a slightly different position was sometimes accepted by the FSA.
- Unexpected  $f(x)$ . For example, if an observation line passes a point where a character boundary crosses a stripe boundary, such an unexpected region change (e.g., B1  $\rightarrow$  F+) will happen; the FSA will reject this change. Fortunately, this problem is not serious; as shown in Figs. 7 and 8, most ( $\sim 98\%$ ) observation lines were accepted by the FSA.

#### 4.4. Recognition accuracy

As noted in 2.2, the final recognition result of a character pattern was determined by the voting of  $P = 81$  extraction results. The overall character recognition rate was  $99.6\% = 233/234$ , that is, only one universal pattern image was misrecognized. This near-perfect recognition rate results from the fact that the ratio of the observation lines with correct  $r_k$  was  $69.2\% (> 50\%)$ .

The misrecognized image was “O” whose recognition result was “N”. Among its 81 observation lines, 27 lines

were “N” and 26 lines were “O”. As shown in Fig. 5, the cross ratio patterns of “N” and “O” are similar and their small difference caused the misrecognition.

## 5. Conclusion

A practical scheme of recognizing a universal pattern has been proposed. The universal pattern is a character pattern printed with a set of horizontal stripes, called cross ratio pattern, for realizing high machine-readability. The cross ratio calculated from the widths of the stripes represents class information and is invariant to projective distortions. Thus, if the boundaries of the stripes are detected correctly, we can determine the character class regardless of projective distortions.

A novel boundary detection technique has been proposed to recognize the universal pattern. This is an optimization-based technique with a guide of a finite state automaton (FSA) for stable extraction of class information under heavy change of grayscale values due to non-uniform lighting and noise. Again, note that this technique is rather general and thus applicable to other tasks.

An experiment of recognizing 234 universal pattern images acquired by a camera in scene was conducted. The obtained recognition rate was  $99.6\% = 233/234$ , that is, a near-perfect rate. The boundary detection performance was also very reasonable even though the character patterns were heavily degraded by non-uniform lighting, noise, and projective distortions.

## References

- [1] The British Computer Society. Character Recognition 1967, Unwin Brothers Limited, 1966.
- [2] 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, pp. 84–104, 2005.
- [3] S. Uchida, M. Iwamura, S. Omachi and K. Kise, “Data embedding for camera-based character recognition,” Proc. First Int. Workshop on Camera-Based Doc. Anal. Recog., pp.60–67, 2005.
- [4] S. Uchida, M. Iwamura, S. Omachi and K. Kise, “OCR fonts revisited for camera-based character recognition,” Proc. 18th Int. Conf. Pat. Recog., vol. 2 of 4, pp.1134-1137, 2006.
- [5] S. Omachi, M. Iwamura, S. Uchida, and K. Kise, “Affine invariant information embedment for accurate camera-based character recognition,” Proc. 18th Int. Conf. Pat. Recog., vol. 2 of 4, pp.1098-1101, 2006.
- [6] R. G. Casey and E. Lecolinet, “A survey of methods and strategies in character segmentation,” IEEE Trans. Pat. Anal. Mach. Intell., vol. 18, no. 7, pp. 690–706, 1996.
- [7] H. Ney and S. Ortmanms, “Progress in dynamic programming search for LVCSR,” Proc. IEEE, vol. 88, no. 8, pp. 1224–1240, Aug. 2000.