CHARACTER RECOGNITION FOR THE MACHINE READER ZONE OF ELECTRONIC IDENTITY CARDS

Hyeogjin Lee and Nojun Kwak

Graduate School of Convergence Science and Technology,
Seoul National University, Korea
{hjinlee, nojunk}@snu.ac.kr

ABSTRACT

This paper proposes an overall procedure of recognizing the machine reader zone of a real world picture of a passport. To begin with, the proposed method finds the area of passport from the input image and its rotation angle is determined. With the rectified passport image by counter-rotating the area of passport, the machine reader zone is found and an inverse projective transform is performed to remove projective distortion. Then, each code is extracted and enhanced by using adaptive posterization. Template matching with improved similarity measure is applied to classify the codes. To classify the number 0 and the character O, a support vector machine is used. The experimental results show the correct character recognition rate of 99.77% and the correct recognition rate of 83.84%.

Index Terms— E-passport, Optical Character Recognition, Machine Reader Zone.

1. INTRODUCTION

The e-passport has the machine reader zone (MRZ) in the bottom part [1]. The MRZ consists of 88 characters or codes which contains information on name, nationality, date of birth, passport number, personal number and check digit for the fabrication. E-passport reader must recognize the MRZ for the security of the e-passport. A security guard certifies identification of the owner and the information on the passport by comparing the data obtained from the MRZ and the data obtained from the IC chip.

In this paper, we propose an overall procedure of recognizing the MRZ of a real world picture of a passport, *e.g.*, the image from the mobile passport reader. Rotation, projective warping, reflected light and brightness difference appear in our passport images as shown in Fig. 1.

There have been many works on e-passport recognition, which can be divided into two main categories depending on whether the passport is snapped or scanned. The quality of snapped e-passport is worse than that of the scanned. The former has perspective distortion and is influenced by illumination. Also it may have reflected light effect.

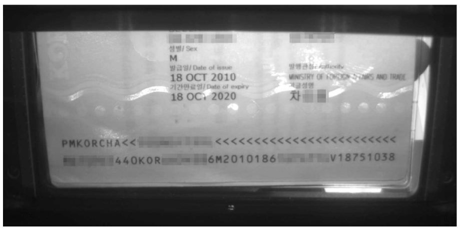


Fig. 1. An example of passport image used in this work. Personal information is blurred on purpose.

Kim *et al.* proposed the methods for optical character recognition (OCR) of the MRZ in scanned e-passport [2, 3, 4]. A smearing method was used to find the MRZ and the angle was adjusted using the longest code line in the MRZ. Each code was extracted using Sobel operator, horizontal smearing, and 8-neighborhood contour tracking algorithm. They recognized each code by variants of clustering algorithms.

Wang *et al.* proposed a method for finding a skew angle in snapped e-passport [5]. They minimize within-class divergence of projected images to find the skew angle. Bessmeltsev *et al.* proposed an effective algorithm for portable passport readers [6]. They used vertical profile projection, and detected skew angle using individual code position to recognize codes. To find the code position, they made a binary image using the Otsu's method.

In this work, we recognize the MRZ of a snapped passport such as Fig. 1. We firstly find the area of passport in an image containing passport. In our case, the passport area is surrounded by a black area, so we find a white passport area from the image. Next, we find a rotational angle by using a method that does not need binarization. More specifically, fast Fourier transform (FFT) and principal component analysis (PCA) are used. Then, the MRZ is extracted. To cope with projective warping, we apply inverse projective transformation to the MRZ. Then, each code, whose quality is enhanced by the proposed adaptive posterization technique, is extracted. This posterization method is more robust to reflected light effect, and simpler than histogram equalization. For character recognition, we use template matching and support vector machine

(SVM). We propose new measure for template matching that reduces edge effect. The SVM is used to classify the number 0 and the character O from the results of template matching.

The paper is organized as follows. Section 2 deals with the preprocessing of the passport image which includes important steps such as extraction and alignment of MRZ, segmentation and quality enhancement of each code. Recognition of each code is dealt with in Section 3 and experimental results follow in Section 4. Finally, conclusions are made in Section 5.

2. PREPROCESSING

2.1. Finding the passport area and rotational angle

The passport image is surrounded by black areas due to our machine characteristic. To find the passport area only, a binary image is obtained using the Otsu's method [7]. As a result, the passport background is bit 1 and the codes and black areas are bit 0.

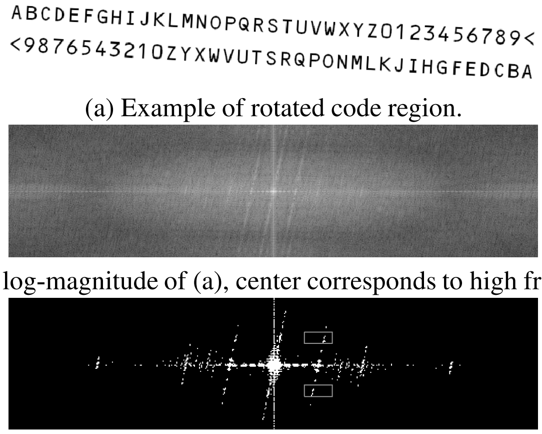
Then, the binary image is projected to horizontal and vertical directions respectively. It is expected that the passport exists in the large value area of each direction [6]. Let T be the threshold obtained by the Otsu's method, and $I(r,c)$ be the pixel value at pixel (r,c). Then one-dimensional projections P_{col} (column-wise), P_{row} (row-wise) are obtained as

$$\begin{array}{ll} P_{col}(k) = \sum_{r=1}^{R} u(I(r,k) - T), & k = 1, \cdots, C \\ P_{row}(k) = \sum_{c=1}^{C} u(I(k,c) - T), & k = 1, \cdots, R. \end{array} \quad (1)$$

Here, R and C are the number of rows and columns, respectively, $u(\cdot)$ outputs 1 if the input is positive, otherwise, outputs 0. We threshold the projection P_{col} and P_{row} with a half of the average of each direction; then we find horizontal and vertical passport area using one-dimensional connected-component labeling. The biggest area is regarded as passport area in both directions.

The snapped or scanned passport image has rotation. There are several methods that correct rotation such as [2, 5, 6]. However, all those methods need a binary image. But in our case, a binary image is difficult to obtain because of image intensity distortion by brightness. Adaptive threshold [8] can make a nice binary image in the code region. But this method makes many errors in a background region of the passport. So we propose a new method to find the rotational angle of the passport.

The proposed method uses the fast Fourier transform (FFT). In the passport image, the MRZ exists in the bottom part; roughly between 5/9 and 1 ratio of the height. The FFT is applied to this rough MRZ and the log-magnitude image is truncated using 50 times the average intensity. Then, it is shifted such that high frequency components corresponds to the center. Finally, the Otsu's method is used to obtain the binary image.


ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789<
<9876543210ZYXWVUTSRQPONMLKJIHGFEDCBA

(a) Example of rotated code region.

(b) FFT log-magnitude of (a), center corresponds to high frequency.

(c) Truncated image of (a) with gray ROI boxes.

Fig. 2. Finding rotation angle by FFT.

Figure 2(b) shows the log-magnitude FFT image of the rotated MRZ in Fig. 2(a). It shows that the rotation of the MRZ affects sub-lines in the magnitude image of FFT. The binary image by the truncation is shown in Fig. 2(c) where the gray boxes are the predefined region of interest (ROI). Rotational angle can be found by the principal component analysis (PCA) [9, 10] of bit 1 locations in ROI, that is eigenvalue decomposition of the covariance matrix:

$$\mathbf{C} = \begin{bmatrix} C_{xx} & C_{xy} \\ C_{xy} & C_{yy} \end{bmatrix}. \quad (2)$$

From [9, 10], the angle between x-axis and the principal component is obtained by

$$\theta = \text{arctan2}(\, 2C_{xy}, C_{xx} - C_{yy}). \quad (3)$$

However, the rotation angle needs correction because the MRZ is not square. Also, in our case, the desired rotation angle is not the angle between principal component and x-axis, but y-axis. Thus, final rotation angle can be obtained as

$$\theta' = \frac{C_M}{R_M} \times (\frac{\pi}{2} - \theta), \quad (4)$$

where C_M and R_M are the numbers of columns and rows of the rough MRZ image. By affine transformation using the resultant rotational angle, we can rectify the passport image.

2.2. Finding and rectangularizing MRZ

In the rectified passport, a rough MRZ still exists between 5/9 and 1 ratio of the height. Adaptive threshold is used to obtain a binary image. Then it is morphologically dilated, and is applied the Canny edge detector [11]. This method is much precise than applying the Canny edge detector directly to the original image.

Many edges are in the code region because codes are black and others are white. If the size of the MRZ is known,

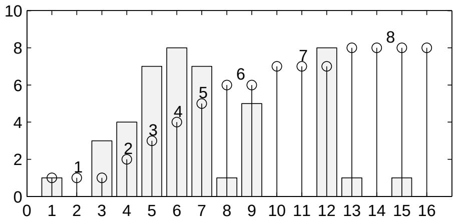


Fig. 3. Toy example of adaptive posterization using 16 histogram bins. The bar shows input histogram and the circle shows the result of adaptive posterization. The same height of circles corresponds to the same value after adaptive posterization.

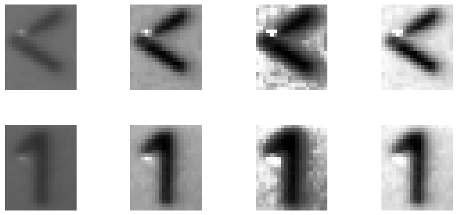


Fig. 4. Comparison of different point processing methods. From left to right: original image, contrast stretching, histogram equalization and adaptive posterization.

the exact location of the MRZ can be obtained by finding the location that has the maximum sum of the edge image in this size. In our case, the size of the MRZ is approximately 1200 by 100 and an integral image [12] of the rough binary MRZ is used to pinpoint the location of MRZ fast.

The MRZ has the projective warping distortion. Thus we perform inverse projective transform, which needs 4 points from a source image and a destination image respectively. The destination image points are four corners of a rectangle, *e.g.*, [0, 0], [0, 99], [1199, 99], [1199, 0].

To find source image points, adaptive threshold is used to obtain the binary image. Morphological dilation is applied to avoid separating one code into two or more. Then, we can find each code's location using two-dimensional connected-component labeling. Inverse projective transform with source and destination image points makes the rectangular MRZ image free from projective warping.

2.3. Separating codes and adaptive posterization

Adaptive threshold is used to obtain the binary image of the rectangularized MRZ. The size and the location of each code are found by the two-dimensional connected-component labeling.

All the code has to have the same aspect ratio, and we choose a reference aspect ratio of 5/6 (width/height). Labeling errors can be handled by using the size and the location of each code. Labels with too big or too small size are deleted. Also, labels with too much distance to the average vertical location are also deleted. After this process, each code is separated with the size of 24 by 20. Then, for character recognition, the quality of each code image is enhanced by a point processing described below.

The famous image enhancement technique is the histogram equalization. But, the histogram equalization merges many bins which may be meaningful into one bin. So, we propose a new method to enhance image modifying the posterization method [13].

The proposed adaptive posterization method also uses histogram bin values. If the bin value is bigger than, or equal to a threshold, this bin is meaningful and it is retained. Otherwise, this bin is considered as non-meaningful and it is merged with other nearby meaningful bin. The threshold can be chosen to be the average value of histogram, *i.e.*, the number of pixels divide by the number of bins.

Fig. 3 shows a toy example of the proposed method with threshold 2.9 (average histogram value). Bins that have smaller value than the threshold (*e.g.*, bin 1 and 2) merge to the nearby meaningful one that has higher value than the threshold (*e.g.*, bin 3).

Fig. 4 compares the proposed adaptive posterization with the conventional point processing methods such as contrast stretching and histogram equalization. Histogram equalization of the special character $<$ has been affected by reflected light, while the proposed method has not. For the character $<$, contrast stretching outputs darker background than the proposed method which makes the character recognition more difficult. Also since the number 1 is thin, histogram equalization makes the code thicker. However contrast stretching and the proposed method keeps it thin. The proposed method performs fine in the presence of reflected light errors.

3. CHARACTER RECOGNITION

3.1. Template matching

After preprocessing steps, segmented code images are obtained. Template matching is used to recognize each code where the template is set to the average of training images. Consider there are C characters to be matched, then the segmented code image is compared with the C templates and it is classified as the code that results in a smallest difference from the corresponding template.

As a measure of difference, the sum of squared difference (SSD) can be used. This measure is similar to the variance of pixel values in difference image.

In this paper, we propose a slightly different variance measure which will be denoted as NM (new measure). Extracting codes from the MRZ and resizing the code image can make small pixel deviation. If there is a small pixel deviation, errors often occur at the edge of the template image. Thus, the squared difference is divided by 2 if that pixel is on the

Table 1. Top 5 error on character recognition.

	TM			TM+NM			TM+SVM			TM+NM+SVM		
	Original	Classified	#	Original	Classified	#	Original	Classified	#	Original	Classified	#
1	0	O	319	0	O	319	Y	T	25	1	T	23
2	O	0	79	O	0	93	1	T	24	Y	T	20
3	Y	T	25	1	T	23	<	K	17	<	K	18
4	1	T	24	Y	T	20	<	1	17	O	D	14
5	<	K	17	<	K	18	O	D	16	<	5	13

edge of the template image. This method of NM has an effect of reducing high frequency of images.

Let $\mathbf{p} = \{p_1, \ldots, p_n\}$ with $p_k(k = 1 \ldots n)$ be each pixel location and n be the number of pixels. We define $\mathbf{D}^{(c)}(p) = |\mathbf{I}(p) - \mathbf{T}^{(c)}(p)|$, where $\mathbf{I}$ is an input image, and $\mathbf{T}^{(c)}$ is a template image of code c. Let $\mathbf{E}^{(c)}$ be the Canny edge of each template, then $\mathbf{E}^{(c)}(p_k)$ is 1 if p_k is on the edge, and 0 otherwise.

Finally, the template matching error $\mathbf{TM}^{(c)}$ with the template code c is obtained as

$$\mathbf{TM}^{(c)} = \sum_{p \in \mathbf{p}} \frac{1}{\mathbf{E}^{(c)}(p) + 1} \left[\mathbf{D}^{(c)}(p) - \overline{\mathbf{D}^{(c)}} \right]^2, \qquad (5)$$

where $\overline{\mathbf{D}^{(c)}} = \frac{1}{n} \sum_{p \in \mathbf{p}} \mathbf{D}^{(c)}(p)$. As a result of template matching, the most similar character $c^* = \arg\min_c \mathbf{TM}^{(c)}$ is found. If $\mathbf{E}^{(c)}$ is zero in all pixel, then $\mathbf{TM}^{(c)}$ is original variance measure.

3.2. Support vector machine (SVM)

In spite of the font made for the OCR, the the number 0 and the character O are hard to recognize. The number 0 is bigger and wider than the character O, but template matching cannot classify those correctly. Thus, we use the support vector machine (SVM) [14] to classify the number 0 and the character O. If the two best results of the template matching are the number 0 or the character O, then the SVM classifies that code image.

4. EXPERIMENTAL RESULT

The e-passport images taken from the e-passport reader are used for training and test. The training phase consists of making template and training SVM. The template is made by average image of each code images. SVM is trained with randomly chosen 500 images out of 4,000 number 0 images and 2000 character O images from preprocessing. Each code image is obtained after preprocessing steps and resized to 24 by 20. Running time for 708 images is 160.7 second and the average time execution for an image is 0.227 second in C++ using OpenCV.

Table 1 and 2 are obtained based on 708 test images which includes over 62,000 characters. Table 1 shows the top 5 error types of character recognition, which shows the major cause of error without SVM is 0/O confusion. On the other hand, after the inclusion of SVM to cope with 0/O confusion, the major error is incurred by 1/T, Y/T, </K and so on. Note that the number of misclassification decreased significantly by the inclusion of SVM.

Table 2. Character and passport recognition rate.

	Correct Ratio	All Match Ratio
TM	99.11%	65.51%
TM+NM	99.10%	66.23%
TM+SVM	99.76%	82.54%
TM+NM+SVM	99.77%	83.84%

Table 2 shows that template matching (TM) with the proposed measure (NM) in combination with SVM performs best. In the table, correct ratio is per character and all match ratio is computed based on the 88 characters in a passport. Template matching with new measure shows a lower recognition rate than the template matching only case. However, Table 1 shows that in TM+NM, O/0 confusion is increased by 14 than TM only case. It is because proposed measure restricts information on the edge and the differentiation between 0 and O by the new measure is more difficult than that by the conventional variance measure. However, other errors are reduced due to new measure. Due to the introduction of the SVM to classify 0 and O, TM+NM+SVM performs better than TM+SVM.

5. CONCLUSION

An overall procedure of MRZ recognition in a passport image is proposed. The proposed procedure consists of two parts, preprocessing and character recognition. The contribution of this paper is 1) FFT is used to find the rotational angle, 2) adaptive posterization is proposed to cope with the reflected light effect, 3) new measure of template matching which permits slight translation is proposed for character recognition for MRZ in a passport. In the character recognition step, SVM is used to classify the character O and the number 0 after template matching. As a future work, we will use SVM to classify all the confusing characters that cause errors.

6. REFERENCES

[1] Jean Monnerat, Serge Vaudenay, and Martin Vuagnoux, "About machine-readable travel documents," in *RFID Security*, 2007, number LASEC-CONF-2007-051.

[2] Kwang-Baek Kim, Jae-Hyun Cho, and Cheol-Ki Kim, "Recognition of passports using FCM-based RBF network," in *Advances in Artificial Intelligence*, pp. 1241–1245. 2005.

[3] Kwang-Baek Kim and Sungshin Kim, "A passport recognition and face verification using enhanced fuzzy ART based RBF network and PCA algorithm," *Neurocomputing*, vol. 71, no. 16, pp. 3202–3210, 2008.

[4] Kwang-Baek Kim, Am-suk Oh, and Young Woon Woo, "PCA-based face verification and passport code recognition using improved FKCN algorithm," in *Int'l Conf. Intelligent Systems Design and Applications*, 2008, vol. 2, pp. 51–57.

[5] Zhiwen Wang, Guoqing Xie, Shaozi Li, and Zhi Liu, "Tilt detection algorithm for OCR credentials image based on directional projection and divergence," in *Int'l Conf. Computer Science and Information Technology*, 2010, vol. 2, pp. 212–215.

[6] Victor Bessmeltsev, Evgeny Bulushev, and Nikolay Goloshevsky, "High-speed OCR algorithm for portable passport readers," *Int'l Conf. Computer Graphics and Vision*, vol. 1, pp. 29–32, 2011.

[7] Nobuyuki Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285-296, pp. 23–27, 1975.

[8] Francis HY Chan, Francis K Lam, and Hui Zhu, "Adaptive thresholding by variational method," *IEEE Trans. Image Processing*, vol. 7, no. 3, pp. 468–473, 1998.

[9] Gary R Bradski, "Real time face and object tracking as a component of a perceptual user interface," in *Proc. IEEE Workshop on Applications of Computer Vision*, 1998, pp. 214–219.

[10] William T Freeman, Ken-ichi Tanaka, Jun Ohta, and Kazuo Kyuma, "Computer vision for computer games," in *Proc. Int'l Conf. Automatic Face and Gesture Recognition*, 1996, pp. 100–105.

[11] John Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, , no. 6, pp. 679–698, 1986.

[12] Paul Viola and Michael Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, 2001, vol. 1, pp. I–511.

[13] Oh-Yeol Kwon and Sung-Il Chien, "Improved posterized color images based on color quantization and contrast enhancement," in *Proc. Int'l Conf. Machine Vision, Image Processing, and Pattern Analysis*, 2011, vol. 5, pp. 1203 – 1206.

[14] Corinna Cortes and Vladimir Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.