

Part-Based High Accuracy Recognition of Serial Numbers in Bank Notes

Bo-Yuan Feng¹, Mingwu Ren¹, Xu-Yao Zhang², and Ching Y. Suen³

¹ School of Computer Science, Nanjing University of Science and Technology,
Nanjing 210094, P.R. China
fengboyuannj@gmail.com

² National Laboratory of Pattern Recognition, Institute of Automation of Chinese
Academy of Sciences, Beijing 100190, P.R. China
xyz@nlpr.ia.ac.cn

³ Centre for Pattern Recognition and Machine Intelligence, Concordia University,
Montreal H3G 1M8, Canada
suen@cenparmi.concordia.ca

Abstract. This paper proposes a novel part-based character recognition method for a new topic of RMB (renminbi bank note, the paper currency used in China) serial number recognition, which is important for reducing financial crime and improving financial market stability and social security. Given an input sample, we first generate a set of local image parts using the Difference-of-Gaussians (DoG) keypoint detector. Then, all of the local parts are classified by an SVM classifier to provide a confidence vector for each part. Finally, three methods are introduced to combine the recognition results of all parts. Since the serial number samples suffer from complex background, occlusion, and degradation, our part-based method takes advantage of both global and local character structure features, and offers an overall increase in robustness and reliability to the entire recognition system. Experiments conducted on a RMB serial number character database show that the test accuracy boosted from 98.90% to 99.33% by utilizing the proposed method with multiple voting based combination strategy. The part-based recognition method can also be extended to other types of banknotes, such as Euro, U.S. and Canadian dollars, or in character recognition applications with complex backgrounds.

Keywords: RMB serial number, part-based, character recognition, classifier outputs combination.

1 Introduction

In the community of handwriting recognition, much attention has been paid to online and offline handwriting recognition and printed character recognition. As RMB circulation management becomes a serious problem to China's financial industry in recent years, a high reliability RMB serial number recognition system is needed. However, little research has been done on bank notes serial number



Fig. 1. RMB images scanned by contact image sensor

recognition (e.g. [1] [2]). In this paper, we investigate the recognition of RMB serial numbers which consist of 10 printed characters (including alphabetic and numeru characters). Fig. 1 shows two scanned RMB images with serial numbers (marked by red rectagles), which have been designed uniquely and used as the identification of RMB.

In recent years, a few RMB character classification techniques have been proposed to improve the recognition performance by artificial neural network (ANN) and support vector machine (SVM). However, no previous study has achieved high accuracy so far. In [1], the back propagation (BP) artificial neural network based on genetic algorithm training method achieved the accuracy of 95%. The serial number identification system based on SVM [2] yields high recognition result on brand new printed RMB, however, the recognition of used RMB serial number with complex background is much more complicated and not studied in that paper.

To obtain RMB serial number characters, we proposed a RMB serial number extraction method in [3]. First, the skew correction and orientation identification are used to detect the region containing RMB serial number. Then we binarize the text region and extract RMB characters by a local contrast average scheme. Overlap recall rate of 79.68% and precision rate of 98.10% are achieved. Recently, a RMB serial number database has been released [4]. In [4], we comprehensively compared different types of feature extraction methods, classifiers, multiple classifier combination strategies, distortion methods, and rejection schemes on the new database, and provided a large amount of experiment results, which could greatly profit further research of RMB banknote recognition.

The character samples in the RMB serial number database contain complex security texture and suffer from inaccurate extraction, various illumination and contamination. Specifically, there are some small circles located around the character strokes, which will strongly affect the classification process and increase the challenge on the recognition task. These color-marked circles shown in Fig. 2 called “EURion constellation” [5] are designed for the prevention of counterfeiting using color photocopiers. As we observed, there are always some parts in the character sample without circles or complex background textures which can provide much more discriminative features for classification than the other parts. It is possible to recognize the input sample by taking advantage of these local parts.

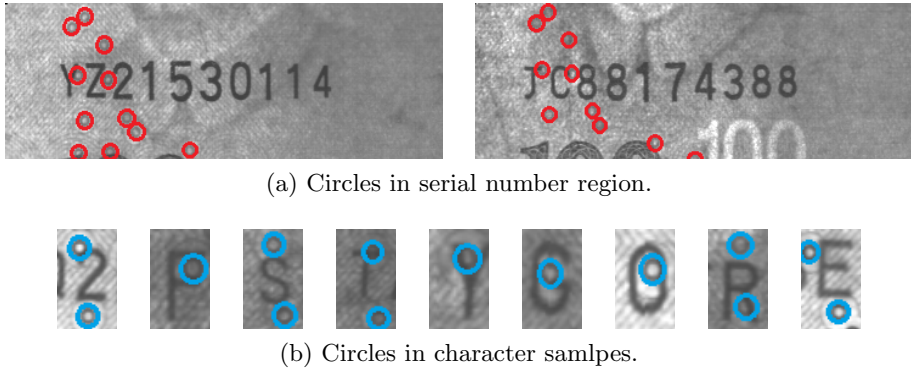


Fig. 2. EURion constellation circles

In this paper, we present a novel part-based method for the recognition of RMB serial number characters. For each training sample, a set of local image parts are synthesized by Difference-of-Gaussians (DoG) keypoint detector [6]. We train an SVM classifier [7] with both the original and local-part images. During the test, each part of the test sample is recognized by SVM, and the final category is determined via combining the classification results of all parts. Suen *et al.* [8] [9] proposed an alphanumeric handprints recognition method by parts, in which they produce a more effective character recognizer based on the probability of occurrence of the patterns. However, they manually choose the local image parts instead of using an automatic interest point detector.

Compared with the previous part-based handwritten character recognition methods [10] [11], our method has some advantages. First, given an interest point, instead of describing the local part by complicated speeded-up robust features (SURF) [12], we extract eight-direction gradient features directly from the fixed-size local image, which is more effective. Second, to speed up the training process, we only generate eight local image parts for each training sample, while the method in [10] uses 60 local parts. Third, the structure of our recognition system is more efficient. In the training step, we do not need to extract the SURF feature vector and build a dictionary. Here, the generated local parts are simply trained together with the original data. In the test step, the SVM model provides a confidence vector for each image part, and the final category is produced by considering all of these local recognition results.

The rest of the paper is organized as follows. In Section 2, we briefly introduce the collected RMB character database. Section 3 details the strategy of part-based character recognition. Section 4 summarizes the experimental results. Finally, we conclude our paper in Section 5.

2 Database

To evaluate the performance of various algorithms for RMB serial number recognition, we collect a database from daily-used RMB images. Scanned RMB images



Fig. 3. NUST-RMB2013 dataset

(Fig. 1) are captured by the contact image sensor (CIS) which installed in the money counting machine with an output resolution of 200×180 dpi. RMB serial numbers are located on the bottom left corner of the scanned image and the characters are extracted straightforward by variance contrast and normalized by BMN [6] to the size of 36×60 . As the extraction process is not so accurate, we manually choose the complete and human readable extraction results to compose the RMB database.

We name our RMB serial number database as NUST-RMB2013 [4] which represent all the different categories of RMB characters by separating into a training set of about 500 samples for each of the 35 classes (numeral 0-9 and alphabet A-Z except V) and a testing set of 200 samples for each class. Fig. 3 shows some RMB character samples containing circles in a complex background.

3 Part-Based Character Recognition

The flow of part-based character recognition approach is depicted in Fig. 4. The recognition procedure has three main steps. We first utilize the DoG keypoint detector to locate the interest points, and extract gradient features of the corresponding local image parts. Then, the trained SVM provides a confidence vector for each part. Finally, the recognition results of all parts are aggregated via three types of combination strategies. In the following, the procedures of image partition, feature extraction, and the integration of local parts are detailed.

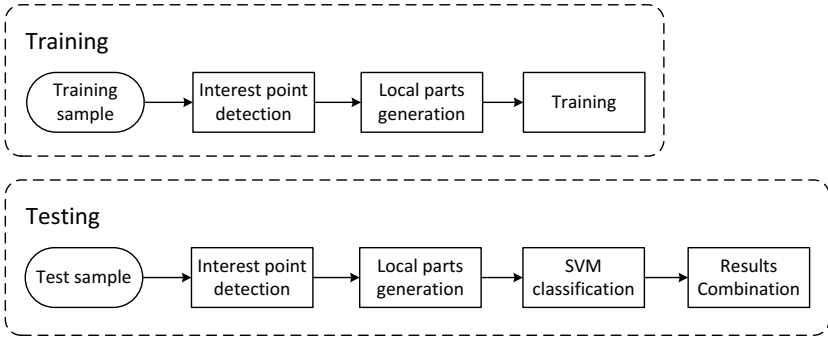


Fig. 4. Working flow of the proposed method

3.1 Image Partition and Feature Extraction

As we mentioned, the character samples may contain a complex background, uneven illumination, smear, and contamination. Specially, there are small circles located around the distinctive part of character strokes, which makes these samples hard to be recognized. However, it is difficult to remove these circles from character samples. In this paper, we propose an alternative way to solve this problem by producing some “clean” local parts from the input samples.

In order to obtain some potential character parts which may not contain circles, we employ DoG keypoint detector [6] to locate the interest points, which are the scales-space extremas in the convolutional result of Difference-of-Gaussian function and the image.

$$DoG(x, y, \sigma) = \left(\frac{1}{2\pi k\sigma^2} e^{-(x^2+y^2)/2k\sigma^2} - \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \right) \otimes I(x, y), \quad (1)$$

where $I(x, y)$ is an input image, σ represents the standard deviation of the Gaussian function, and k denotes a constant multiplicative factor. We estimate σ and k according to paper [6]. The keypoints are the local maxima and minima of $DoG(x, y, \sigma)$, specifically, the sample point will be selected as keypoint only if it is larger or smaller than its eight neighbors in the current image and nine neighbors in the images of scale above and below. Fig. 5 shows some keypoints detected in character samples.

Each training sample can generate W (depends on the internal structure of character) keypoints, which could enlarge the size of the training dataset W times. However, there are some keypoints located close to each other, and the relevant synthetic parts are very similar. To remove those redundant interest points and reduce the training time of SVM model, for each sample, we cluster the number of keypoints to K using the k-means [13] clustering algorithm (Fig. 5). Considering the trade-off between training cost and recognition rate, we empirically set K as eight. In the experiment, we also tested K with larger values such as 10 and 12, however, both of them barely provide any improvement.

The local image parts are extracted centered by the interest points. As we aim to remove the circles while keeping the integrity of character strokes, the

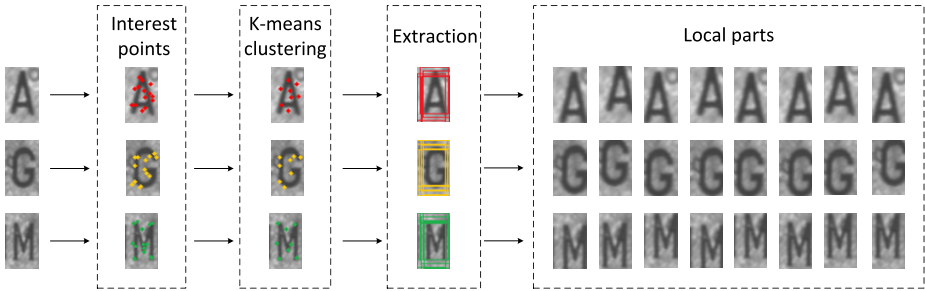


Fig. 5. Interest points detection and local parts generation

size of local parts can not be too small. We intentionally fixed it to $5/6$ of the original image (30×50) by considering the size of characters and the circles in the RMB database. For the interest points which are located at the corners of the sample, we shift the extraction window according to the image boundaries. As shown in Fig. 5, some local parts without circles can be generated, which will help to promote the recognition performance in the subsequent procedure. To make sure the features extracted from each local part has the same dimension, and also facilitate the classification process, we stretch all local parts to the same size of 36×60 .

For each part of the image, the eight-direction gradient features are extracted [14]. First, eight gradient direction feature maps of the input sample are obtained by Sobel operator. Then, each gradient direction image is divided into 6×10 blocks. Gradient magnitude of each block is extracted by Gaussian blurring, hence the feature dimensionality is 480 (60×8). The feature vector is reduced to 34-dimensional (class number minus one) learned by Linear Discriminant Analysis (LDA) [15].

3.2 Local Parts Combination

System Training. We trained the SVM classifier with eight direction gradient features extracted from both the original and local images. The SVM models with Linear and RBF kernel are tested for fair comparison and analysis. We implemented the SVM via LibSVM software [16] which can provide the confidence weight of each category for the test sample. As our database has 35 classes, the output of SVM classifier is a 35-dimensional confidence vector.

Classifier Outputs Combination. The classifier combination applies on the outputs of SVM without knowledge of the internal structure of classifiers and their feature vectors. Given a test sample, we extract gradient features from the local parts. By feeding these features to SVM, a 35-dimensional confidence vector can be obtained from each part. We combine the classifier outputs of both the original and local images at abstract and measurement levels [17].

Three methods are used to aggregate the output measurements, namely, max rule, major voting, and multiple voting, respectively. For a test image \mathbf{x} , assume

M local image parts are obtained, together with the original image, we have $M + 1$ input images in total. The SVM classifier provides an N -dimensional confidence vector $V_i = \{v_{i1}, v_{i2}, \dots, v_{iN}\}$, $N = 35$ for each image part.

The max rule performs on the measurement level. It finds the maximum confidence value among all local parts for each class, then classifies the text sample \mathbf{x} to the corresponding category with the highest measurement.

$$\mathbf{x} \in \arg \max_{j=1}^N \{ \max_{i=1}^{M+1} v_{ij} \} \quad (2)$$

The major voting strategy combines the outputs of classifier on abstract level. Assuming each image part only belongs to one class, it counts the recognition results of all parts, and sorts the sample to the class which has the maximum votes. Let $R = \{r_1, r_2, \dots, r_{M+1}\}$ denotes the classification results for all local parts, the value of r_i is set between 1 to N . The major voting method can be described as:

$$\mathbf{x} \in \arg \max_{j=1}^N \sum_{i=1}^{M+1} I(r_i), \quad (3)$$

where

$$I(\cdot) = \begin{cases} 1 & \text{if } r_i = j \\ 0 & \text{else} \end{cases} \quad (4)$$

Both of these two methods introduced above provide outstanding recognition results in spite of their simplicity. As the local parts which are classified to the correct category always belong to the side of the majority, the major voting can restrain the influence of the circles by taking advantage of statistical distribution. Meanwhile, the lower complexity combination approach will compensate for the representational information lost during combinations.

The multiple voting method [11] can also be called as the sum rule strategy, which intergrates the recognition results on measurement level. Unlike the major voting method which assumes each part only belongs to one class, given a local image, multiple voting method considers its distribution to all categories. In the test, we obtain 35-dimensional confidence vectors $V_i = \{v_{i1}, v_{i2}, \dots, v_{iN}\}$, $i = 1, 2, \dots, M + 1$ for the local parts of sample \mathbf{x} . Then, the confidence weights of different image parts are summed according to their categories. \mathbf{x} belongs to the class with the maximum sum.

$$\mathbf{x} \in \arg \max_{j=1}^N \{ \sum_{i=1}^{M+1} v_{ij} \} \quad (5)$$

Individual recognition results of all image parts contribute to the final decision by summing the confidence weights for each class. The benefit of using such strategy is that it effectively suppresses the influences of circles and other factors to the classification. For instance, the SVM classifier may have difficulties in distinguishing the original sample contaminated by circles while some synthetic

Table 1. Test accuracies (%) of local parts combination methods on NUST-RMB2013

Combination	SVM kernel	
	Linear	RBF
Conventional	98.90	99.31
Max rule	99.27	99.43
Major voting	99.07	99.30
Multiple voting	99.33	99.51

local parts may not. They can provide correct recognition results with high confidence weights helping to find the proper category after sum operation.

4 Experiments

4.1 Baseline Recognition Results

Both the Linear and RBF kernel based SVM classifiers have been used to construct the baseline classifier for the recognition of RMB serial number. Before classification, the features were reduced to 34 dimensions by LDA. The hyperparameters of SVM were selected via cross-validation on the training data. The NUST-RMB2013 database contains the serial number characters of 35 categories with 17,262 training samples and 7000 testing samples in total. The recognition rate on the original database using SVM is 98.90% with Linear kernel and 99.31% with RBF kernel, respectively [4].

4.2 Part-Based Recognition Results

In the training process, the part-based method enlarged the training data about eight times by generating eight local parts for each training sample. Along with the original data, the total size of training dataset became 155,358.

Since a larger number of local parts implies a higher probability that there exists some local images without circles, given a test sample, we created as many local images as we can to ensure the robustness of our method. Eight-direction gradient features of these local images were fed to SVM classifiers (Linear and RBF kernel) to get the 35-dimensional confidence vectors. After that, three combination methods aggregated the individual recognition results to produce higher accuracies. Table 1 compares the performances of different combination methods on NUST-RMB2013 database.

We find out that all of the part-based methods lead to approving results. Multiple voting is the most competitive strategy, which achieves the highest classification accuracy on SVM trained with both Linear and RBF kernels. The accuracy on test data boosts by 0.43 % and 0.2% using Linear and RBF kernel, respectively. The max rule also helps to improve the recognition rates to reach

the accuracies of 99.27% and 99.43% by different SVM kernels. However, the advantages of major voting is relatively inferior to the other methods. It merely provides 0.17% improvement on the model trained by Linear SVM, and barely works on the model with RBF kernel because the accuracy is identical to the conventional recognition method.

4.3 Discussion

Since the max rule and multiple voting methods combine the outputs of the SVM classifier on the measurement level by utilizing confidence weights of each local part, it is reasonable that they outperform the major voting scheme which assumes each local part only belongs to one category and intergrates the recognition results on the abstract level. The experimental results prove our part-based method works well on the database which contains uneven illumination, contrast variation, smear, and complex background texture. Especially, the misclassification problem caused by the anti-counterfeiting circles appeared in the RMB character samples are properly solved. The best recognition rate of 99.51% is given by multiple voting combination method with the SVM model trained with RBF kernel.

According to Table 1, the part-based strategies is less sensitive to noise than the conventional recognition method. The reasons for the superior performance are twofold. First, we train the SVM classifier with both the full-size and local image part samples, which makes the SVM model considering not only the global but also local character features. Second, the part-based method makes capital of the statistical distribution of the recognition outputs obtained from individual local parts. The various local parts generated from the test sample can eliminate the influences caused by the circles and complex background, and help the classifier to make the right determination.

The distortion method has been proved very helpful in [4] and [18]. Compared with the distortion method which also expands the training dataset by generating additional synthetic training samples, our part-based method has its own merit. Since the elastic distortion randomly chooses its scaling and rotation parameters, and the translation distortion shifts the input pattern one or two pixels towards eight directions, neither of them considers the internal structure of training sample. On the contrary, we make use of the samples' characteristic structure information by utilizing DoG keypoint detector to locate the interest points and extracting the surrounding area as local parts. The distinctive principle of the DoG detector helps to extract more distinguishing local parts than distortion methods, which promises a higher recognition rate.

Fig. 6 demonstrates some samples misclassified by the part-based recognition method using multiple voting combination strategy. There are some circles touching the distinctive part of character strokes, which makes these samples extremely hard to be recognized. We are not able to remove these circles even by extracting the local parts. To deal with this problem, we plan to find a method to detect the circles in character samples using image processing technique in

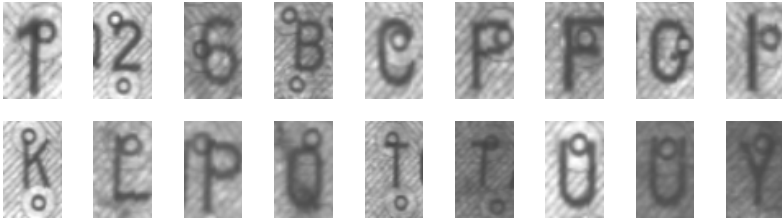


Fig. 6. Misrecognition samples

future studies. After removing the circles from the data samples, a new recognition system with higher accuracy can be conducted on these “clean” character samples.

4.4 Rejection

High reliability is required in the problem of RMB serial number recognition which tolerates no classification error. Even a 0.1% error rate will cause significant economic loss in the banking system. The reliability [19] is defined by:

$$\text{Reliability} = \frac{\text{Number of correct recognitions in accepted samples}}{\text{Number of accepted samples}} \times 100\%, \quad (6)$$

We use the first two rank measurement (FTRM) rejection method to improve the reliability by rejecting confusing patterns. The FTRM is based on the gap between the top-2 outputs of classifiers. Fig. 7 shows the rejection tradeoff curves of the recognition methods using conventional SVM and the part-based scheme with multiple voting combination strategy. It indicates that the part-based method produces a higher reliability than the traditional recognition method, which achieves 99.92% reliability with 2.91% rejection rate.

5 Conclusions

In this paper, a novel part-based character recognition approach for RMB serial number recognition has been proposed. According to the characteristics of the samples in RMB dataset, we automatically generate a set of local parts for each input sample using the Difference-of-Gaussians (DoG) keypoint detector. The feature vectors are extracted by eight-direction gradients and reduced to 34 dimensions by LDA. Both the original and local image parts are used to train an SVM classifier. In the test step, we first obtain the confidence vector for each part. Then, three different types of methods are investigated to combine the recognition results of all image parts.

Experiments conducted on a large serial number character database named NUST-RMB2013 show the superior performance of part-based character recognition method. It exploits both the global and local character structure features,

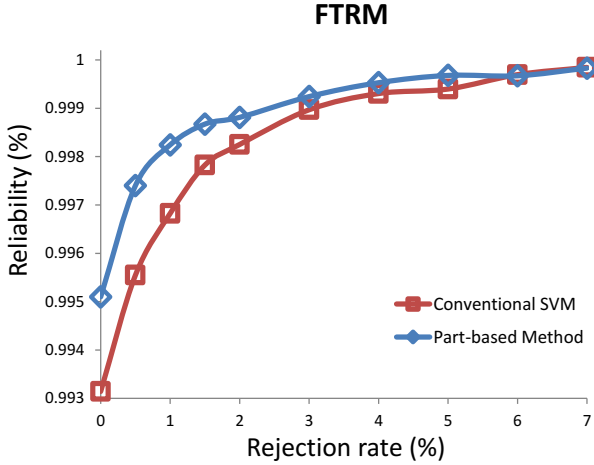


Fig. 7. Rejection-reliability tradeoff for FTRM

and offers an overall increase in robustness, performance and reliability to the entire recognition system. Apart from its obvious advantages of simplicity and completeness, it has been proved to be able to recognize the imprecisely extracted characters which have lost their global structure by occlusion, decoration, contamination, and other degradations (Fig. 3).

Three local parts combination methods have been compared and analysed. Since the measurement level combination approaches such as max rule and multiple voting consider the confidence vectors of test samples, they show better performances than the method based on abstract level (major voting). The highest recognition rate 99.51% is achieved by the RBF kernel based SVM cooperating with the multiple voting combination scheme. Compared to the conventional recognition method, the error reduction rate is 39.09% and 28.99% using Linear and RBF kernel based SVM, respectively. We find that the major voting is not a promising solution to our problem since it only leads to a slight improvement of the recognition accuracy.

What is more, the concepts of the proposed part-based character recognition method can also be used in other types of banknotes, such as Euro, U.S. and Canadian dollar, or in the recognition of document character samples with complex backgrounds.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant 61231014 and 60875010. We also wish to thank Prof. Cheng-Lin Liu, the colleagues of Centre for Pattern Recognition and Machine Intelligence (CENPARMI) and National Laboratory of Pattern Recognition (NLPR) for their great support and help in our work.

References

1. Zhao, T.T., Zhao, J.Y., Zheng, R.R., Zhang, L.L.: Study on RMB number recognition based on genetic algorithm artificial neural network. In: Proc. International Congress on Image and Signal Processing, pp. 1951–1955 (2010)
2. Li, W., Tian, W., Cao, X., Gao, Z.: Application of support vector machine (SVM) on serial number identification of RMB. In: Proc. World Congress on Intelligent Control and Automation, pp. 6262–6266 (2010)
3. Feng, B.Y., Ren, M., Zhang, X.Y., Suen, C.Y.: Extraction of serial numbers on bank notes. In: Proc. International Conference on Document Analysis and Recognition, pp. 698–702 (2013)
4. Feng, B.Y., Ren, M., Zhang, X.Y., Suen, C.Y.: Automatic recognition of serial numbers in bank notes. *Pattern Recognition* 47(8), 2621–2634 (2014)
5. Nieves, J., Ruiz-Agundez, I., Bringas, P.G.: Recognizing banknote patterns for protecting economic transactions. In: Proc. International Workshop on Database and Expert Systems Applications, pp. 247–249 (2010)
6. Lowe, D.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60(2), 91–110 (2004)
7. Cortes, C., Vapnik, V.: Support-vector networks. *Machine Learning* 20(3), 273–297 (1995)
8. Suen, C.Y., Guo, J., Li, Z.: Analysis and recognition of alphanumeric handprints by parts. *IEEE Trans. Syst., Man, Cybern.* 24(4), 614–631 (1994)
9. Suen, C.Y., Kim, J., Kim, K., Xu, Q., Lam, L.: Handwriting recognition—the last frontiers. In: Proc. ICPR, pp. 1–10 (2000)
10. Uchida, S., Liwicki, M.: Part-based recognition of handwritten characters. In: Proc. International Conference on Frontiers in Handwriting Recognition, pp. 545–550 (2010)
11. Wang, S., Uchida, S., Liwicki, M.: Comparative study of part-based handwritten character recognition methods. In: Proc. International Conference on Document Analysis and Recognition, pp. 814–818 (2011)
12. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: Speeded up robust features. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006, Part I. LNCS, vol. 3951, pp. 404–417. Springer, Heidelberg (2006)
13. Hartigan, J.A., Wong, M.A.: A k-means clustering algorithm. *Applied Statistics* 28(1), 100–108 (1979)
14. Bai, Z.L., Huo, Q.: A study on the use of 8-directional features for online handwritten chinese character recognition. In: Proc. International Conference on Document Analysis and Recognition, pp. 262–266 (2005)
15. Mika, S., Ratsch, G., Weston, J., Scholkopf, B., Muller, K.: Fisher discriminant analysis with kernels. In: Proc. Signal Processing Society Workshop, pp. 41–48 (1999)
16. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. and Technol.* 2(3), 27:1–27:27 (2011)
17. Suen, C.Y., Lam, L.: Multiple classifier combination methodologies for different output levels. In: Proc. International Workshop on Multiple Classifier Systems, pp. 52–66 (2000)
18. Simard, P., Steinkraus, D., Platt, J.C.: Best practices for convolutional neural networks applied to visual document analysis. In: Proc. International Conference on Document Analysis and Recognition, pp. 958–963 (2003)
19. Wang, W.-N., Zhang, X.-Y., Suen, C.Y.: A novel pattern rejection criterion based on multiple classifiers. In: Zhou, Z.-H., Roli, F., Kittler, J. (eds.) MCS 2013. LNCS, vol. 7872, pp. 331–342. Springer, Heidelberg (2013)