A Courtesy Amount Recognition System for Chinese Bank Checks

Dong Liu
Graduate School at Shenzhen
Tsinghua University
Shenzhen, China
liu-d05@mails.tsinghua.edu.cn

Youbin Chen
Graduate School at Shenzhen
Tsinghua University
Shenzhen, China
chenyb@sz.tsinghua.edu.cn

***Abstract*—In this paper, we present a complete courtesy amount recognition system for Chinese bank checks. The system takes color bank check images as input and consists of three main processing steps: numeral string extraction, segmentation & recognition, and post-processing. They focus sequentially on: detection and extraction of numeral string; segmentation and recognition of the string; and further analysis of recognition results for acceptance or rejection. Information fusion, method complementarity, multi-hypotheses generation then evaluation are three principles employed for designing algorithms in the first two modules. And logistic regression is used for post-processing. A large number of real checks collected from different banks are used for testing the system. Read rate around 82% is observed when the substitution rate is set to 1%, which corresponds to that of a human operator. The performance can also be tuned further toward a suitable balance between inaccuracy and rejection, in accordance with user preference.**

Keywords-automatic bank check processing; courtesy amount recognition; document image analysis; optical character recognition (OCR); confidence computation

I. INTRODUCTION

Bank checks are widely used all over the world. Despite the rapid growth of payment by credit cards and many other electronic means, the paper check is still one of the most popular forms for noncash payment. Since a large number of checks need to be processed manually every day, there is a great demand for systems to read checks automatically. And the courtesy amount recognition (CAR) system is one of the most important parts.

Systems for reading checks automatically have been reported in France, US, Brazil, and many other countries [1]. Some of them have already been put into practical use in different banks, such as the A2iA CheckReader [2]. However, as shown in Fig.1, the courtesy amount area in Chinese checks is so different from the other countries' in both image quality and layout that none of the currently available systems mentioned above are able to process very well. The courtesy amount in Chinese checks is written in pre-printed grids and sometimes covered by red seals. Seals bring a lot of challenges in binarization while the overlap between character strokes and preprinted guidelines makes extraction extremely difficult. And segmentation for those grid-restricted strings should be quite different from unconstrained ones. New algorithms should be designed in order to process checks in China. Several experimental CAR systems for Chinese bank checks have already been reported in [3][4], but their performances are far from satisfaction for practical use.

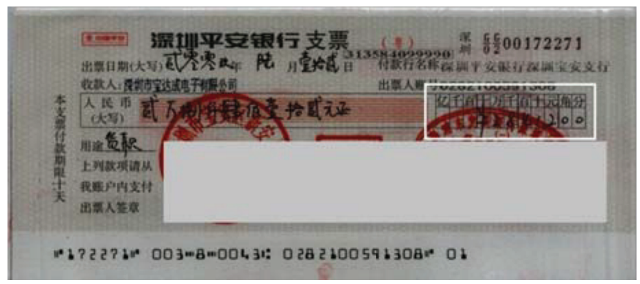

Figure 1. A typical image of Chinese bank check

In this paper, we present a complete CAR system for Chinese bank checks. String extraction, segmentation & recognition, and post-processing are sequentially implemented with color check images. Information from color, gray and binary images is fused and complementary methods are combined for binarization. Multi-hypotheses are generated by keeping all possible candidates when uncertainties exist in string extraction and segmentation. Then all candidates are evaluated by recognition and the corresponding probabilities are given. Confidence for the best recognized candidate is calculated by logistic regression model and the final decision is made on whether to accept or reject the recognition result.

The rest of this paper is organized as follows. Section 2 presents the framework of our system. Implementation details are described in section 3 and experimental results are given in section 4. Section 5 contains our conclusion and future work.

II. FRAMEWORK

The whole system consists of three main processing steps: string extraction, segmentation & recognition, and post processing. Each of them is constructed by one or more modules, as shown in Fig.2.

The input check images are got from different banks by different scanners and operators. Some of them are skewed so the skew angle is detected then corrected in the first module. According to the unique layout standard for all Chinese bank checks, courtesy amount area is roughly extracted from a pre-defined region and sent to the next

978-0-7695-4774-9/12 $26.00 © 2012 IEEE
DOI 10.1109/ICFHR.2012.154

module. Uniform background and red seals which may cover the characters are removed in the binarization module. In the extraction module, guidelines in the restricted grids are exactly localized then removed to get the numeral string extracted. Multi-hypotheses are made by keeping all possible results when it is not sure whether one part of guideline should be removed or not. For segmentation, all the connected components (CCs) of a numeral string are analyzed first and the ones which correspond to connected characters are identified according to their sizes and positions. Then the cut points for segmentation are determined and multiple candidates are kept if uncertainties exist. All candidates generated previously are sent to the recognition module, and a list of possible results with their corresponding probabilities is given. Confidence for the most probable candidate in the list is calculated in the post processing module, then decision is made on whether to accept or reject the recognized amount.

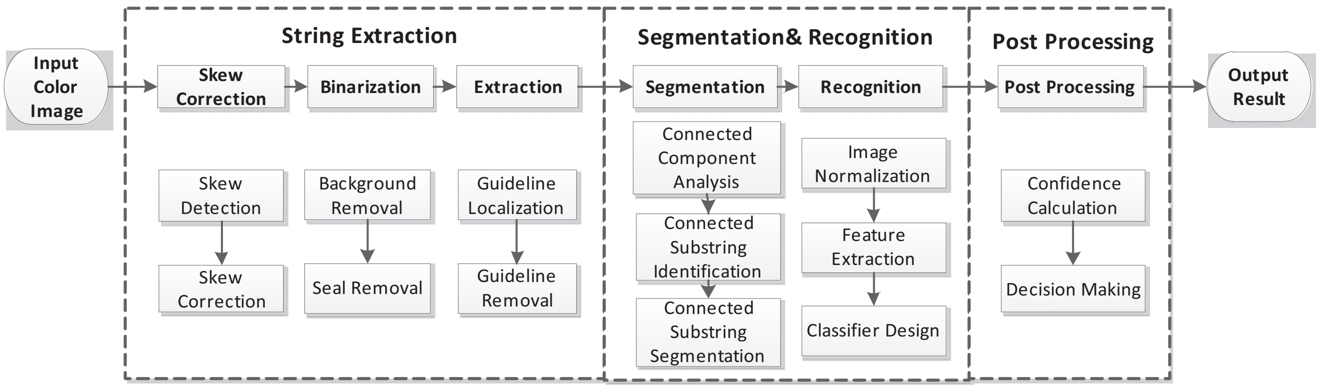

Figure 2. Flow chart of proposed system

III. Implementation Details

The prototype of our CAR system which contains the first two processing steps has been presented previously. So we just give a high-level description for modules in the first two steps in order to leave space for other parts. More information about the first two processing steps can be found in our previous work [5][6].

A. Skew Correction

Skew for the whole image is detected then corrected in this module. The input image is converted from RGB color space to gray level first and edges are extracted using Canny method. Then long horizontal guidelines are detected using the standard Hough transformation. Skew angle is calculated by averaging the angles of top few significant guidelines and the original image is rotated to the standard position. In our system, skew angle never exceeds 20 degrees. So we just detect lines within 20 degrees in the Hough transformation for acceleration. According to the unique layout standard for Chinese bank checks, the courtesy amount area is roughly extracted from a pre-defined region, which locates at the upper right position of a check, and sent to the next module for further processing.

B. Binarization

Uniform background and red-seals are sequentially removed in this module. Otsu method [7] is applied with gray image first, followed by Niblack method [8] which is used only on the left foreground pixels. With the combination of these two complementary methods, character strokes and red seals which may appear in the courtesy amount area are sharply binarized from uniform background.

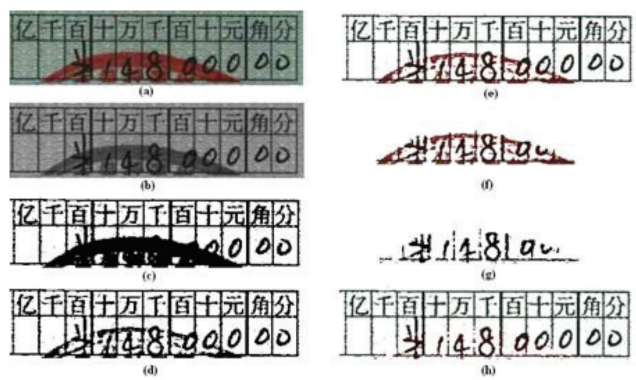

Figure 3. (a)Original color image (b)Gray image (c)Result by Otsu method (d)Result after employing Niblack method (e)Color image with background removed (f)Seal-like pixels extracted (g)Stroke pixels got by threshold T (h)Final result after stroke recovery

Then color information is used for removing seals. The left foreground elements in the original image are converted from RGB to HSI color space and all the red-like pixels are extracted if their hue values are around zero. So all the seal pixels are extracted, as well as stroke pixels covered by seals, which are also red-like although their R values in RGB color space are lower than pure seal pixels. Then a threshold is automatically calculated for the R value of the extracted pixels in RGB color space. Pixels with a lower R value are regarded as strokes covered by seals and brought back to the original image, while others are removed. Finally, all the remaining pixels are regarded as foreground and the binarization result without background and seal is got. An example for the proposed binarization algorithm is shown in Fig.3.

C. Extraction

In this module, guidelines in the restricted grids are localized then removed in order to get the numeral string extracted.

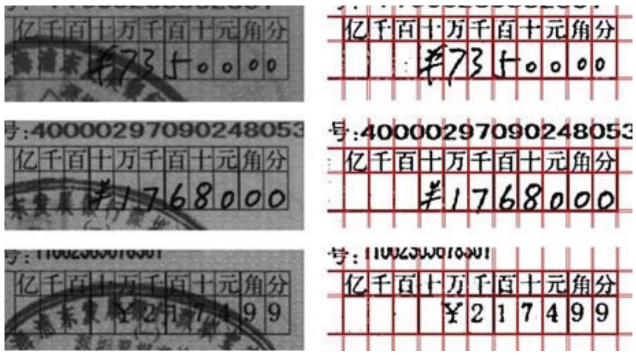

Figure 4. Examples of localization results

Projection in both horizontal and vertical directions are carried out with the binarized image first. Peaks which correspond to the guidelines in the projection results are identified with the help of the prior known layout of Chinese checks, and positions for the guidelines are got. Then double edges of the guidelines are further localized. Gradient is computed within a small neighborhood of the guidelines in the original gray image and positions for the double edges can be determined precisely by finding maximum (or minimum) in the projection histogram. Some localization results by our algorithm are shown in Fig.4.

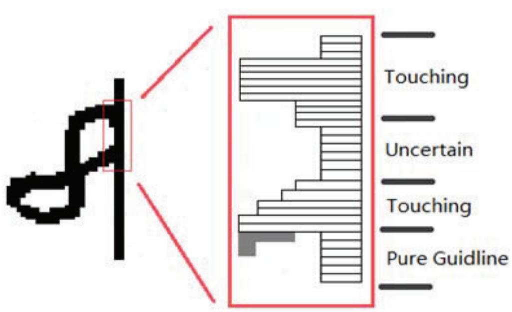

Figure 5. Example of run-length analysis for vertical line

After all the edge positions are got, run-length analysis is performed for every guideline in order to remove the restricted grids. A set of connected black pixels in the same row for vertical line (or column for horizontal line) is regarded as a run-length, and a line is represented by a group of run-lengths from different rows (columns), as illustrated in Fig.5. If one run-length exceeds the area defined by the double edges of the guideline, it is likely that strokes are connected with guidelines at this position. Guidelines at connected positions are preserved, while others are removed. If more than one connected position exits on the same line of a restricted box, it is not sure whether the gaps between them are pure guidelines or strokes covered by guidelines. So multi-hypotheses are generated by both keeping and removing the uncertain parts in different candidate results. Some examples of the extracted candidates are given in Fig.6.

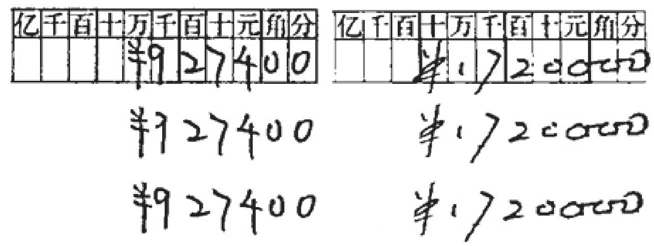

Figure 6. Examples of extracted candidates

D. Segmentation

Numeral string is segmented into isolated digits in this module. All the connected components (CCs) are found first by contour tracing, and several groups of CCs are obtained by merging small fragments into big CCs nearby. Because the numeral string is written in pre-printed grids, whose positions have been got in previous module, it is easy to tell isolated digits from connected ones. If the main body of one group lies in a single grid, it is extracted as an isolated digit. And the groups containing connected substrings are left for further segmentation. Number of connected digits in a group is estimated by how many grids it covers, and vertical boundaries of the grids are used for reference in finding the segmentation points. Contour difference algorithm, which is proposed by Fujisawa [9] for segmenting connected digits, is performed and cutting points as well as the ligatures which connect the digits sometimes are found. Sometimes ligature is one part of the digits and cannot be removed directly. In order to make sure the correct segmentation result can be got, two candidates are kept: one without ligature and the other one with part of ligature.

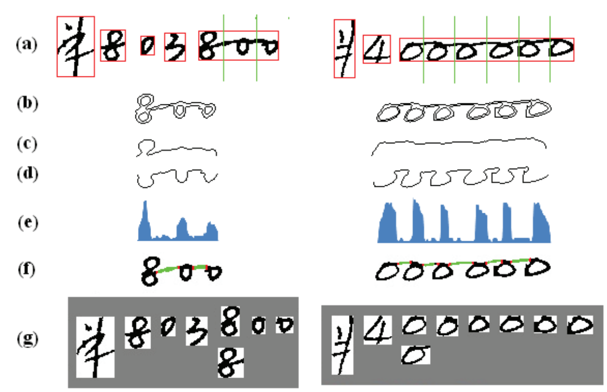

Figure 7. (a)Connected substring identification by analyzing CCs, positions of vertical lines are taken as referenced cutting points (b)Contour of connected substring (c)Upper contour (d)Lower contour (e)Contour difference (f)Candidate cutting points and ligatures (g)Segmentation results with candidates

In Chinese checks, the courtesy amount string is often ended with several zeros connected together. In this situation, ligatures between zeros should be removed directly in order to save computation effort. So in our system, if a CC to be segmented lies at the end of the string and inner loops can be detected in every grid, two candidates are only kept for the first digit in order to avoid mistake when the substring does not start with zero. And one single candidate is kept for the following digits by removing ligatures directly. Illustration for our segmentation algorithm is given in Fig.7.

E. Recognition

Our OCR engine for isolated digit is constructed with state-of-art techniques mentioned in [10][11]. Input images are first zoomed into a standard size by moment normalization. Gradient feature is then extracted and support vector machine (SVM) is utilized for classification. Part of source code from LIBSVM [12] is used for building the engine, while algorithm proposed in [13] is employed for computing the probabilities of the SVM outputs.

For each group of segmentation candidates, all the isolated digits except the first RMB sign are sequentially recognized. Because the last two digits which indicate cents in courtesy amount string of Chinese bank checks are zero in most cases, it is reasonable for us to assume that zero has a higher prior probability in the last two grids. So the recognized results for the last two digits are modified by multiplying a parameter k (1.5 in our system) with the probability of zero, then probabilities for all digits from each grid are multiplied together as the final probability of the whole recognized string.

All the groups are processed in the same way and results with corresponding probabilities are obtained. Then they are sorted by decreasing probability and sent to the next module.

F. Post Processing

In this module, the likelihood of correctness for the best recognized result in the list is calculated as confidence. And the decision is made on whether to accept the result or not.

In order to evaluate the confidence, two features are extracted from the candidate list: the probability for the top candidate and the ratio of probability between the first two candidates. Their distributions among 10,000 samples are shown in Fig.8，for convenience all the ratio values higher than 20 are regarded as 20 in the distribution histogram. It is clearly shown that these two features are very effective in distinguishing right from wrong because they have different feature values and results with high feature values are more likely to be correct.

Logistic regression model is then employed for calculating the confidence, which is ideally suited as it always predicts a value between 0 and 1. The confidence is represented as:

$$\text{Confidence} = 1/\left(1 + \exp\left(-\sum_i w_i \cdot f_i\right)\right) \qquad (1)$$

Where f_i is the value of the i^{th} feature and w_i is the pre-trained weight for that feature. 2,000 samples with equal positive and negative data are randomly selected as training set, and LIBLINEAR toolbox [14] is used for training our model. The logistic model we obtained is shown in Fig.9, where the relationship between confidence and the two features is perfectly reflected.

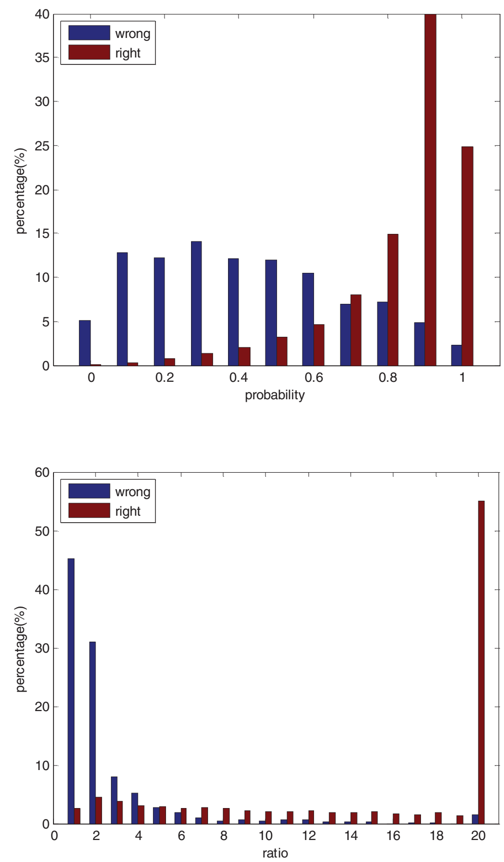

Figure 8. Feature distribution histograms

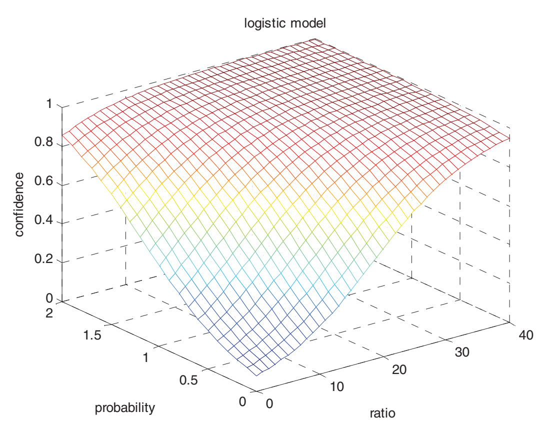

Figure 9. Logistic model for confidence

After the confidence is got, whether to accept the recognized result is naturally decided by a pre-defined threshold. If the top candidate in the list has a higher confidence than the threshold, it is accepted as the final result, otherwise it is rejected. Due to the particular requirements of this application, the threshold can be tuned to yield low level of incorrect reading with high rejection rate or vice versa. And the system can always get a suitable balance between inaccuracy and rejection.

IV. EXPERIMENTS

In China, banks are nearly paranoiac about the confidentiality of the checks they provide for research. It is almost impossible to get a public check database for testing our system, so all the experiments are carried out with data collected by ourselves. In our experiments, more than 20,000 real checks are collected from different banks and scanned into color images with 200DPI. Some of them are used for training our system and the other 19,144 images are left for testing.

Several different measurements used for evaluating the performance of our system are defined as follows:

$$\text{Read} = \frac{\text{number of accepted checks}}{\text{total number of processed checks}} \tag{2}$$

$$\text{Recognition} = \frac{\begin{array}{c}\text{number of checks accepted}\\ \text{with correct amount}\end{array}}{\text{total number of processed checks}} \tag{3}$$

$$\text{Substitution} = \frac{\begin{array}{c}\text{number of checks accepted}\\ \text{with incorrect amount}\end{array}}{\text{total number of accepted checks}} \tag{4}$$

It is notable that Substitution is used here, which is more meaningful than the frequently used Error for it only considers errors occurred within the accepted checks.

TABLE I. PERFORMANCE OF DIFFERENT SYSTEMS

System	No. of test samples	Recognition rate		
		1st candidate	*Top 2 candidates*	*Top 3 candidates*
System in [3]	84	89.3%	/	/
System in [4]	1053	86.5%	/	/
Our system	19144	91.3%	93.8%	94.5%

The system is first tested without rejection by setting a zero threshold in the post processing module. And its global recognition accuracy is shown in Table 1. Results of two other CAR systems for Chinese bank checks are also listed for comparison. As shown in the table, our system can get high recognition accuracy on a large data set. And higher accuracy can be obtained by considering more candidates. For security reasons, the three systems listed in Table 1 are tested with different data. It is not fair to compare the results directly. However, our system is tested with a large number of real check images and a considerable high accuracy can be obtained, it is still reasonable for us to announce that the results of our system are very satisfactory.

The relationship between rejection and inaccuracy is then tested by varying the threshold used in the post processing module. And the correlation curve is shown in Fig.10. About 82% of checks can be accepted when substitution is fixed at 1%, which is a typical level for human operators. Compared with the A2iA CheckReader mentioned in [2], whose read rate in practical use varies from 65% to 85% when the substitution is fixed at 1%, the performance of our system is very promising.

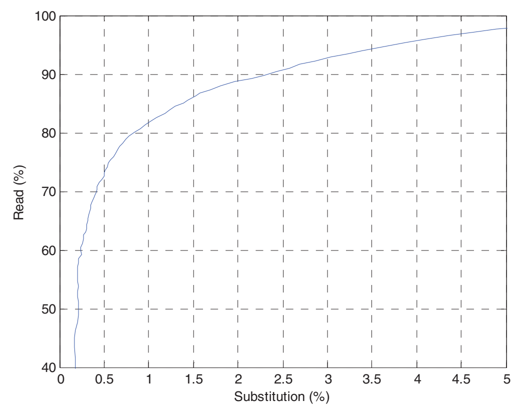

Figure 10. Correlation curve between Read and Substitution

All the experiments are implemented on a laptop with Intel i5 2540m @2.6GHz, 8GB RAM, and running Windows 7 64bit. The CAR system uses a single CPU thread and the average time for processing one check image is 280ms, which is efficient enough for practical use.

V. CONCLUSION AND FUTURE WORK

In this paper, we present a complete courtesy amount recognition system for Chinese bank checks. The system sequentially performs automatic extraction by skew correction, binarization and grid removal; segmentation of extracted string into isolated digits; recognition of digits; confidence computation and decision making to accept or reject the recognized amount. An 82% read rate is observed on a large test set with an error rate fixed at the level of a human operator, which is very promising for practical use.

Despite the achieved performance, new algorithms which are complementary to the current system are still under development in order to improve the accuracy. Research on the fusion and cross-validation with legal amount will be carried out in the near future.

REFERENCES

[1] G.Dimauro, S.Impedovo, G.Pirlo, A.Salzo, Automatic bankcheck processing: A new engineered system, World Scientific, Singapore, SINGAPOUR, 1997.

[2] N. Gorski, V. Anisimov, E. Augustin, O. Baret, S. Maximov, Industrial bank check processing: the A2iA CheckReader; TM, International Journal on Document Analysis and Recognition, 3 (2001) 196-206.

[3] Hongdong Song, Yali Wang, Shaowei Xia, Recognition System of Handwritten Numerals on Checks Using Support Vector Machine, Computer Engineering and Applications, 3 (2003), 58-60.

[4] Junxia Gu, Xiaoqing Ding, Multi-model Segmentation and Recognition Algorithm of Courtesy Amount, Journal of Image and Graphics, 4 (2008), 696-701.

[5] Dong Liu, Youbin Chen, A Prototype System of Courtesy Amount Recognition for Chinese Bank Checks, the 10th IAPR International Workshop on Document Analysis Systems(DAS 2012), in press.

[6] Dong Liu, Youbin Chen, Binarizing the Courtesy Amount Field on Color Chinese Bank Check Images, the 11th International Conference on Document Analysis and Recognition (ICDAR2011), pp.774-778.

[7] N. Otsu, A threshold selection method from graylevel histograms, IEEE Trans. Syst., Madn. & Cybern., 9 (1979) 62-66.

[8] Wayne Nilback, An introduction to digital image processing, Prentice-Hall International N.J., 1985.

[9] H. Fujisawa, Y. Nakano, K. Kurino, Segmentation methods for character recognition: from segmentation to document structure analysis, Proceedings of the IEEE, 80 (1992) 1079-1092.

[10] C. Liu, Handwritten digit recognition: benchmarking of state-of-the-art techniques, Pattern Recognition, 36 (2003), 2271-2285.

[11] C.-L. Liu, K. Nakashima, H. Sako, H. Fujisawa, Handwritten digit recognition: investigation of normalization and feature extraction techniques, Pattern Recognition, 37 (2004) 265-279.

[12] C.-C. Chang and C.-J. Lin. LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011.

[13] T.-F. Wu, C.-J. Lin, R.C. Weng, Probability Estimates for Multi-class Classification by Pairwise Coupling, Journal of Machine Learning Research, 5 (2004), 975-1005.

[14] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin. LIBLINEAR: A Library for Large Linear Classification, Journal of Machine Learning Research, 9 (2008), 1871-1874.