RECOGNITION OF HANDWRITTEN CHINESE POSTAL ADDRESS USING NEURAL NETWORKS

Yih-Ming Su[1,2] and Jhing-Fa Wang[2]

[1]Department of Electronic Engineering, I-Shou University, Kaohsiung County, Taiwan.

[2]Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan, Taiwan.

ABSTRACT

An automatic mail-sorting machine based on the development of a real time OCR system is proposed. The system being the first of its kind in handling Chinese mail is capable of distinguishing the city/county names in Taiwan, on handwritten/machine-printed standard Chinese style envelopes. The identification of the Chinese postal addresses is achieved by implementing an address recognition strategy that consist of a number of stages, including preprocessing, address block location, address segmentation, character recognition, contextual postprocessing, and result arbitration stages. Some empirical approaches are adopted to improve both the speed and effectiveness of postal address processing, in order to put the system to work in a real-time situation. The system architecture and related strategies are being reported. Experimental results demonstrated that the system is capable of sorting 5400 mail pieces per hour with a correct rate of 75.6% and an error rate of 0.92%.

1. INTRODUCTION

Mail sorting automation has become an urgent matter to improve the postal service, because traditional manual sorting consumes large amount of labor and time to handle high volume of mails every day. The goal of this study was to accomplish the automation of mail sorting by integration of a mechanized sorting machine, computer vision, and the development of OCR. This research, the first being reported on the application of such integration in the sorting of Chinese postal system, has concentrated on achieving high precision (i.e., lower error rate) and implementing the recognition system in real time.

Similar works in the US [1,2], UK [3], and Japan [4] have been reported in the past few years, but the postal system and procedure for each country differs, and so do their problems. Srihari [1,2], described a handwritten address interpretation (HWAI) system, which included digit string segmentation, digit recognition, character recognition, word recognition, and control structure in 1984. The system has a throughput rate of 12 letters per second with correct rates at about 75.8% and error rates of less than 1.2%. Many HWAI systems have been deployed throughout the United States since then. In the UK [3], the address recognition system being implemented was based on the design of a fast postcode dictionary and verification of full postcode against town and street name. Preliminary results indicated a verification rate of 57% for outward postcodes and 32% for overall postcodes, with an error rate of 4.9%. Tokunaga [4] described the developmental history of Japanese mail handling system such as automatic postal code reading, culling, facing, canceling, and sorting machine. The system has been installed at post offices to accelerate delivery-related work. However, in Taiwan, due to the difference in mail formats and writing styles with other countries, together with poor quality of the envelope images (lowest resolution of about 72 dpi) and higher skew distortion. The development of high performance real time mail sorting system remains a challenging task.

The proposed address recognition system focused on the recognition of the city/county name, because the mails in our postal system are first classified by such and then send to corresponding local postal sorting center.

In the rest of this paper, Section 2 describes the preprocessing and address block location process. In Section 3, a new address segmentation method is presented, followed by the descriptions of component analysis, fine segmentation, and component combination procedures. Section 4 presents feature extraction method and their classification in the character recognition process together with postprocessing operations. In Section 5, experimental results and discussions are reported, with concluding remarks and suggestions for future researches being mentioned in the final section.

2. PREPROCESSING AND ADDRESS BLOCK LOCATION

2.1 Envelope Detection and Displacement Modification

The image caught by an interlaced CCD camera is composed of even frame and odd frame image. However, there is a displacement between even frame and odd frame image, caused by the camera and motion of the envelope. In order to preserve the quality of the envelope image, via the formation of a static image, it is necessary to calculate the displacement between even frame and odd frame image. The concepts of gradient and LMSE fitting are adopted to evaluate the displacement. The vertical boundary points in the even and odd image frame are detected by scanning horizontal direction from right to left and finding maximum difference in gray scale between two adjacent pixels. The LMSE fitting is respectively used to transform boundary points into a line for two image frames. The displacement between two image frames is calculated by the difference of two intercepts of linear equations from two boundary lines.

2.2 Envelope Skew Adjustment and Translation

The envelope image is slanted. Such skew may be caused by many factors including the strength of vacuum pump, the moving speed of conveyor belt, and envelop quality. If the image is slanted, difficulties will arise for subsequent processes such as address block location and address segmentation. Therefore, the skew angle of an envelope should be evaluated in order to adjust a slanted image at this stage. The skew angle is calculated by transforming the slope of a boundary line. The skew angle is calculated by transforming the slope of a boundary line. If the

0-7803-4455-3/98/$10.00 © 1998 IEEE

difference of two skew angles from horizontal and vertical direction is more than 5 degrees, the processing of the envelope image is abandoned in order to avoid too much distortion in the adjusted image. The intersection of the vertical and horizontal boundary line is at the upper-right hand corner of the envelope. According to this intersection point and the vertical skew angle of the vertical boundary line, the slanted image can be translated and rotated to coordinate (511,0).

2.3 Binarization and Address Block Location

The image captured by the CCD camera is a gray-scale image. For convenience of the further processing, this image should be transformed into a binary image. The histogram of the image pixel shows two peak values existed in the overlapped static image. These two peak values represented the foreground and background of the gray-scale image. Using the following formulae, a binary image can be obtained.

$$Cut\_Threshold = \begin{cases} P_x * 0.85, & H_{fp} \geq 150, \\ P_x * 0.81, & 100 \leq H_{fp} < 150, \\ P_x * 0.78, & H_{fp} < 100, \end{cases}$$

where Px represents the total pixel of the foreground image and H_{fp} is the peak value of the foreground image. The factors used in each formula are arbitrarily selected according to experimental results for optimum binarization. Since the shutter speed of the CCD camera is very fast, and the intensity of the illuminating light (as connected to AC source) changes continuously, the images caught by the camera will differ in grey levels. This will be true even for the image of the same envelope being captured at different time.

On standard Chinese envelopes, a number of simple geometrical features can assist in the identification of address block on the envelopes. The geometrical features include a bounding rectangle located on the central part of an envelope with sides parallel to the edges of the envelope and five small bounding rectangles on the top right hand corner of the envelope. The name of the consignee is filled within the central bounding rectangle, and the ZIP code filled within the first three of the five small rectangles. In addition, the address of the consignee is filled in the space at the right side of the central rectangle. The address block is identified by the following. (1) The right vertical frame line of the central rectangle is extracted by seeking maximum gradient in the vertical projection of the given bottom block. (2) The upper horizontal frame line of the central rectangle is extracted by seeking maximum gradient in the horizontal projection of the strip. (3) The upper-right corner of the central rectangle is extracted by calculating the intersection of the vertical and horizontal frame line of the central rectangle. (4) The upper horizontal starting position of address block is located by selecting a small block at the right side of the corner as shown in the top binarized block. The small printed character postal tag can be removed by identifying their properties such as their larger aspect ratio and more narrow width.

3. ADDRESS SEGMENTATION

A new address segmentation method based on three characteristics of Chinese handwriting style is used to perform component analysis, fine segmentation, and component combination. The characteristics are as follows. (a) Most Chinese characters have a certain range in their aspect ratio. (b) The size of each handwritten character for an individual does not vary too much. (c) The gaps between individual characters are larger than those between radicals of a character. In this stage, to obtain best performance in address segmentation, in addition to three characteristics of Chinese writing being used, it was further assumed that a segmentation point is located on a ligature between characters. In most cases, the assumptions of Chinese writing are sufficient.

The address segmentation process is broken down into several steps. Firstly, the address characters are analyzed and processed by statistic estimation, which evaluates character characteristics including stroke width, likely character height, and its variance. Then, the three characteristics of Chinese characters written by an individual are used to separate over long character blocks within connected components (i.e., touching characters) for fine segmentation, using curve-mapping function [5] to partition connected components. Finally, component combination, based on the operation of a decision table is used to determine whether adjacent components should be merged or not.

After the preprocessing and address block location, the binarized image of the address block is partitioned into separated character blocks by scanning the vertical projection of the address block, in order to resolve parallel character strings. The vertical projection of the address block is composed of simple running count of the black pixels in each column. Moreover, the projection indicates positions of connected strings in the address block. When the character strings touch or overlapped, the projection often contains a minimum at the proper segmentation point. The block width of a vertical character string is calculated by counting successive columns with non-zero vertical projection value. If the width of a block is too small, the block is then regarded as the small print postal tag on the envelope and discarded. Fig. 1 shows an example of the extracted individual characters after preprocessing and address segmentation.

4. CHARACTER RECOGNITION

In the present study, after address segmentation, the individual characters are extracted for character recognition, the process of which is broken down into several steps. First, a normalization process was only used to stretch smaller size characters by linear interpolation technique. Then, a feature extraction technique based on non-linear division method was used to extract important features of a single character. Again, a classification process that uses multi-layer perceptron as classifier was employed to identify input characters. Finally, contextual postprocessing based on word matching and dictionary look-up was used to further enhance the system performance.

4.1 Feature Extraction

Feature extraction is introduced to speed up data processing time. However, for Chinese characters, the wide variation in handwriting style together with the large number of individual characters poses severe difficulties in locating optimum features in each character for feature extraction. Fortunately, in the present application, due to the relatively small number of characters involved, viz. 32 altogether, a simple and effective technique of feature extraction is still possible. To reserve the structural information between the strokes within a character, the image of each character is partitioned into 7 by 7 grid frames, as a 49-dimension feature vector, using a nonlinear division method [5]. The extracted features for each character are invariant with respect to translation and scaling.

4.2 Character Classification

Neural network models have been successfully applied in the field of OCR as a pattern classifier. However, initial learning of numerous data (also known as training the network) is required to facilitate satisfactory operations. The objective of training the network is to adjust the weightings of different parameters, so that input feature vector can produce the desired output through the network. The present multi-layer perceptron (MLP) is trained with back-propagation algorithm [6] as classifier. Such algorithm can adjust the weighings of different parameters between layers and finds a suitable hyper-region to cluster each character category. The neural network model is fully connected between adjacent layers with the number of nodes in the input layer corresponding directly to the dimensions of feature vector, and the number of nodes in the output layer corresponding directly to the number of the character categories.

In the present system, three different multi-layer perceptron models are used individually, to recognize the three related characters that comprise city/county names. During the network training stage, a learning rate (η) of 0.05, and the total error tolerance, (*error*) of 0.005, are used for each network structure. In order to enhance the system capability, rejected envelopes were also used for on-line training through manual operations.

4.3 Contextual Postprocessing

When the first two related characters of city/county names were recognized from character classification process, contextual postprocessing is used to further enhance the performance of address recognition. The relevant characters combined to form a word that is passed forward for further analysis in which the candidate words are ranked by their outcomes. Each recognized character selects *n* highest output of the character classifier as candidate characters. Then word matching process is employed to decide whether all of the candidate characters have contextual relationship by the help of a prior knowledge consisting of a number of city/county names. If two or more word candidates are found in the word matching process, a word candidate with the maximum criterion outcome (*CO*) is selected as the best result of address recognition. The *CO* for the *j*th word is defined as:

$$CO(j) = sum(output_j^1 + output_j^2), \quad \text{for } j=1,..,M,$$

where $output_j^1$ indicates the outcome of the output node in the first character classifier for the first character in the *j*th word, $output_j^2$ indicates the outcome of the output node in the second character classifier for the second character in the *j*th word, and M indicates the number of word candidates. Finally, result arbitration is then performed by choosing an arbitration threshold (*AT*) value above which the best result is considered acceptable. A function of the result arbitration (*RA*) is described as follows:

$$RA = \begin{cases} accetped & \text{if } CO \text{ of the best result} \geq AT, \\ rejected & \text{otherwise.} \end{cases}$$

A real result for address recognition is checked with a three-digit ZIP code, through the mapping of the best result, before the ZIP code is sent to the sorting machine through an RS232 communication.

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this section, we analyze the experimental results for address recognition on the city/county names. In order to investigate the performance of the mail sorting system, a database set of live envelopes provided by the Postal Pavilion was used as a source of training and test pattern setting. The database set consists of 32 categories each contains 80 patterns of Chinese characters, with 50 as training data and 30 as test data. Since the 32 categories included all recognized characters from the city/county names in Taiwan, they can be partitioned into two groups that represent the first or second character of the city/county names.

In general, each city/county name consists of a three-character ideogram. Some of specific cases are abbreviated by a two-character ideogram. There are 16 distinct characters present in the first group, as shown in the first row of Table 1(a). There are 19 distinct characters present in the second group as shown in the first row of Table 1(b). Thus, the segmented characters from addresses pass through the character recognition stage in turn. The architecture of a four-layer perceptron for the first character group is constructed with 49 nodes in input layer, 30 nodes in the first hidden layer, 40 nodes in the second hidden layer, and 16 nodes in the output layer. The architecture of a four-layer perceptron for the second character group is constructed with 49 nodes in input layer, 40 nodes in the first hidden layer, 40 nodes in the second hidden layer, and 19 nodes in the output layer. The architecture of a four-layer perceptron for the third character group (i.e., "縣/県(Hsien)" and "市(City)") is constructed with 49 nodes in input layer, 15 nodes in the hidden layer, and 3 nodes in the output layer.

Table 1(a), (b), and (c) show the individual recognition rates of each character by using the test pattern data. In Table 1(a), some characters, such as "北(Pei)", "中(Chung)", "宜(I)", and "高(Kao)", are not easily differentiated in feature space, because they are structurally simple and similar with other characters. For structurally simple characters, they may yield few valid components in feature extraction process. Structurally similar characters (e.g., different characters with only vertical and horizontal strokes) may not be distinguished reliably as well. In Table 1(b), since the number of characters is more than that in Table 1(a), the average recognition rates decreased accordingly. Recognition performance for all cities and counties after contextual postprocessing is an average correction rate of 75.6% and error rate of 0.91% under an arbitration threshold value of 1.1, when the top 3 candidates for each recognized character are selected. The proposed system when applied in the automation of actual postal mails system demonstrated a very low error rate and a satisfactory correct rate. The throughput rate of the entire recognition process based on the process speed of the Pentium-133 using C language is about 1.54 envelopes per second, namely, 0.65 second per envelope.

Most of the rejections in address recognition could be attributed to four causes. Firstly, the camera sensitivity to light changes is not very stable, resulting in undesirable changes of gray level. Secondly, the address segmentation for very irregular writing style poses severe challenges and resulting in broken characters extraction. Thirdly, the stroke structures for scripted characters are not very clear, resulting in the extraction of incomplete characters features for recognition. Finally, poor contrast due to inadequate envelope quality or lightly written characters also can be lower recognition performance. Should these problems be overcome, we believe the recognition rate will be enhanced further.

6. CONCLUSIONS AND SUGGESTIONS

A new mail sorting system based on real-time recognition of hand-written/machine-printed Chinese postal addresses has been developed for automatic mail handling in Taiwan. The recognition strategy for the postal addresses including preprocessing, address block location, address segmentation, character recognition, contextual postprocessing, and result arbitration stages is proposed for the system. Some simple and efficient approaches are used in each stage in order to achieve a real-time operation requirement. Experimental results of running the system gave an average correct rate of 75.6% and error rate of 0.92% for 1468 test envelops. This clearly demonstrated the feasibility of automating the sorting of Chinese postal addresses.

Future research efforts will emphasize on improving the correction rate of address recognition and further recognition of the other portion of a handwritten address such as street name and section numbers. In addition, some issues on character segmentation and feature extraction shall be enhanced such that the mail sorting system can achieve better performance.

7. REFERENCES

[1] P.W. Palumbo and S. N. Srihari, "Postal Address Reading in Real Time," Inter. J. of Imaging Science and Technology, 1996.

[2] S.N. Srihari and E.J. Kuebert, "Integration of Hand-written Address Interpretation Technology in the United States Postal Service Remote Computer Reader System," CEDAR-TR-97-1, 1997.

[3] A.C. Downton, R.W.S. Tregidgo and C.G. Leedham, and Hendrawan, "Recognition of Handwritten British postal addresses," From Pixels to Features III: Frontiers in Handwriting Recognition, pp. 129-143, 1992.

[4] Y. Tokunaga, "History and current state of postal mechanization in Japan," Pattern Recognition Letters, vol. 14, no. 4, pp. 277-280, April 1993.

[5] Y. M. Su and J. F. Wang. "Recognition of Handwritten Chinese Postal Addresses Using Dual-Expert Classification Scheme," Proceedings of International Conference on Image Processing and Character Recognition (part of International Computer Symposium), Kaohsiung, Taiwan, Republic of China, pp. 213-219, 1996.

[6] Simon Haykin, "Neural networks- A Comprehensive Foundation", New York, Macmillan College Publishing Company, 1994.

Table 1(a). The recognition rate for the first character group

	台	北	中	南	基	宜	新	桃
Recognition	88	66	73	80	83	73	86	93
Rate (%)	苗	彰	雲	嘉	高	澎	屏	花
	83	83	90	83	73	80	90	80

Table 1(b). The recognition rate for the second character group

	投	隆	蘭	竹	園	栗	化	林	義	雄
Recognition	80	73	83	76	80	93	70	76	83	76
Rate (%)	湖	東	蓮	縣	県	市	北	中	南	
	80	73	90	96	60	70	66	83	73	

Table 1(c). The recognition rate for the third character group

	縣	県	市
Recognition Rate	96.6%	87.5%	90.6%

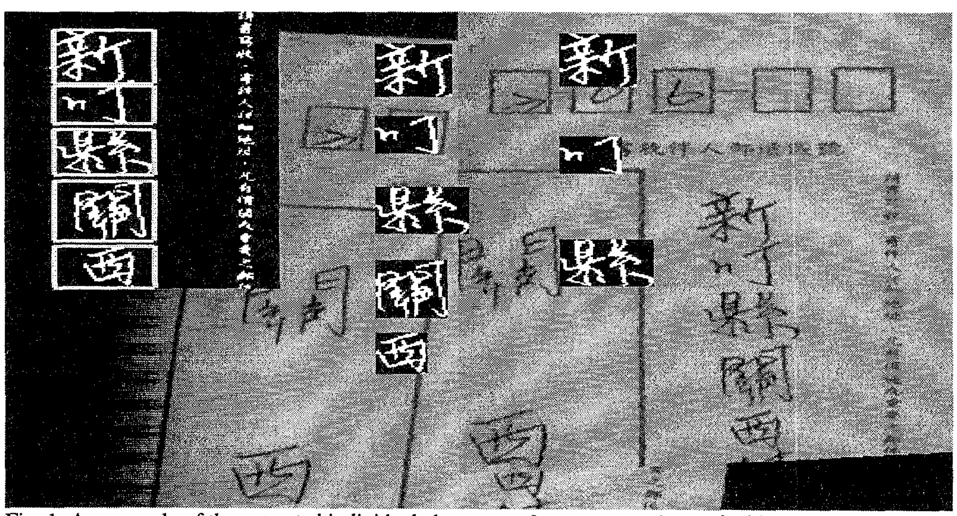

Fig. 1. An example of the extracted individual characters after preprocessing and address segmentation.