A Vision-Based Fast Chinese Postal Envelope Identification System*

XUE GAO AND LIANWEN JIN
School of Electronic and Information Engineering
South China University of Technology
Guangzhou, 510640 P.R. China

This paper describes a method of real-time processing of machine-printed Chinese postage envelopes in an automated postal system. We propose a vision-based postal envelope identification method, and a system that can be applied to rapidly moving machine-printed Chinese postage envelopes through the postal process. Our system uses a high-speed camera to capture images of envelopes running on a conveying device, and then automatically recognizes the postal address and postcode on each envelope. A laser sensor is used to trigger the camera to capture the images. Our system supports a vocabulary of 4,590 categories of characters, including 4,516 frequently used Chinese characters defined in the GB2312-80 standard, 62 alphanumeric characters, and 12 punctuation marks and symbols. Supported font styles include the Song, Fang Song, Kai, and Hei fonts, among others, at a printed font size of 7.5 points and above. The results of an experimental trial of the system with 761 envelope images representing 25,060 characters revealed that an envelope with an average of 32.9 characters could be processed and recognized within 81.38 milliseconds. The character recognition rate of postal addresses is 98.72%. Furthermore, our system also provides a method for the real-time storage of envelope images and recognition results into a database, which can be used in subsequent envelope querying, tracking and management. The results of an experimental trial with real live mails in a postal center indicated that our system could achieve a speed of 21,000 envelopes per hour, with the character recognition rate of postal addresses as high as 98.92%.

Keywords: character recognition, real-time vision system, postal envelope identification, document image processing, pattern recognition

1. INTRODUCTION

Automatic address reading is one of the most important components of postal automation. In postal service centers, an enormous number of mail items must be processed and dispatched from sender to recipient every day. The processing of mail items quickly and reliably requires the destination addresses to be read quickly and accurately. However, traditional manual processing methods typically consume large amounts of labor and time. The emergence of methods using optical character recognition (OCR) has provided a solution to automatic postal address reading. In these methods, a camera is used to capture images of mail items, and then OCR is applied for recognition of the postal address. Automatic mail sorting is the most common application of automatic address reading in the automation of postal systems. Traditional automatic mail sorting systems have typically focused on the identification of the postcode on the envelope, since the

Received February 9, 2011; revised August 24, 2011; accepted August 31, 2011.
Communicated by Yuh-Jye Lee.
* This work was supported in part by NSFC (No. U0735004, 60772116, 61075021) and GDSFC (No. S201102-0000541, 2010B090400397, 2010A090100016).

correct recognition of the postcode determines the city and state of the destination address. However, recent research shows that the recognition of the actual postal address can improve the accuracy of automatic sorting [1-3] because the information contained in the postal address can be used to verify whether the postcode has been correctly recognized. For example, the names of the city and street in a postal address can be matched to those indicated by the postcode.

With the development of postal automation, deeper understanding of addresses and more precise recognition of postal addresses are urgently required for the following functions: (a) electronic archiving of mail: the recognition results of postal addresses associated with the corresponding envelope images, and storage in a database for subsequent querying and management; (b) construction of a postal address database; (c) automatic assigning of barcodes: in postal service centers, each mail item is usually assigned a barcode before dispatch [4, 5]. This barcode can then be read for the needs of subsequent mail sorting, tracking and management in postal automation. An accurately recognized address is one of the most important sources of information for barcode encoding.

Postal addresses on envelopes are typically either handwritten or machine-printed. With the development of office automation (OA), machine-printed postal addresses are becoming increasingly popular, especially for commercial letters. It has been estimated that an average of more than 300,000 outward machine-printed envelopes are processed in a day for a few of the main postal processing centers of China. To ensure the timely processing of mails, any automatic postal address reading system must have an extremely rapid processing speed for practical applications. Our preliminary investigations suggest that this required speed is around 20,000-30,000 envelopes per hour. Although address reading has been extensively exploited during the last three decades in the field of OCR [6] and several systems are already available [2, 5, 7-12], no simple system that can satisfy the speed requirement has been reported in the literature to date. Only one system based on a complex hardware structure has been able to read addresses with fair good accuracy and satisfying speed for machine-printed/handwritten English postal envelopes [5]. This is because that most present systems have primarily focused on the recognition of handwritten addresses, which generally require complex algorithms to deal with shape variation. As a result, current handwritten address recognition techniques cannot be directly applied to machine-printed address recognition, and recognition speed is unsatisfactory.

In the current study, we aim to provide a solution that enables high-speed and high-precision machine-printed Chinese postal address recognition at a relatively low cost. To this end, we propose a vision-based fast Chinese postal envelope identification method and system. Our system uses a high-speed camera to capture images of envelopes running on a conveying device, and then recognizes machine-printed postal addresses and postcodes on the envelopes. A laser sensor is used to trigger the camera to capture the images. To enable the high-speed processing of the postage envelopes, a fast destination address block location algorithm and a document image binarization algorithm are proposed at first. And then, the novel Chinese character feature representation methods and a two-stage classification strategy are applied to achieve high-speed and high-precision recognition of postal address. Finally, a post-processing algorithm of the recognized results is proposed to improve the recognition performance further. In addition, in character segmentation, we use a dynamic programming algorithm to find optimal segmentation path in term of the scores given by classifier and the constraint information about the Chinese characters and the envelope layout. Based on our collected database of character samples, a real-time processing and recognition system of the Chinese postage envelopes is constructed. The results of an experimental trial with 761 mail images revealed that our system could process and recognize the machine-printed postcodes and postal addresses rapidly and accurately, indicating the effectiveness of the proposed method.

Furthermore, besides automatic address recognition, our system also enables the real-time storage of envelope images and address recognition results into a database, which can be conveniently used in subsequent envelope querying, tracking and management. This enables our system to be widely used in postal automation.

The rest of the paper is organized as follows: Section 1.1 reviews related research. Section 2 gives an overview of our proposed postal envelope identification method and system. Section 3 describes in detail the techniques used in our system, followed by the experimental results and analysis in section 4. Section 5 outlines our conclusions.

1.1 Related Research

Typical automatic postal address recognition systems can generally be divided into five components: (a) envelope image preprocessing; (b) destination address block location; (c) line/character segmentation; (d) character recognition; (e) post-processing of recognized address. Intensive research over the last three decades has achieved substantial improvements in each of these components, and several systems for automatic postal address recognition are already available. Srihari and Keubert proposed a remote computer read system for machine-printed/handwritten English postal addresses, using six recognition processors and some other control devices to reach a processing speed of 90,000 envelopes per hour [5]. For handwritten Chinese postal address recognition, the results reported in the literature vary according to the particular recognition strategy adopted, and the different hardware performance [2, 8, 13, 14]. In a more recent report, Jiang *et al*. described a system achieving an accuracy rate of 85.3% with a speed of 3 seconds per mail using a 3 GHz personal computer [2]. Su *et al.* reported a system that achieved a speed of 0.65 seconds per mail [8]. Liu *et al.* described a lexicon-driven Japanese address recognition system where the processing of one envelope could be performed within approximately 110 milliseconds [9]. Japanese address recognition is a similar task in the field of OCR. However, the above-mentioned processing time does not include that of address block location and text line segmentation. Furthermore, this system needs to construct an address lexicon database, which is not a trivial task. On the other hand, the reported character correct rates for handwritten Chinese address recognition are only around 76-86% [2, 8, 9]. And it should be noted that some of these reports did not consider address block location errors. Wang *et al*. [15] presented a handwritten Chinese address recognition method with a character correct rate of 96.45%, based on their collected test database of handwritten address string images, not on real data from actual envelopes. In [14], Han *et al*. published a correct sorting rate of 96.2% for handwritten Chinese mails. However, considering that present mail sorting in China may only be exact in terms of province/city name, but not street name, this result is far from satisfactory. In addition, Jeong *et al*. [16] reported an address block location method based on connected components (CCs) for the images of machine-printed Korean envelopes with a success rate of 96%, although this system did not deal with sequent segmentation and recognition. Our experiments indicate that CC-based location methods are relatively time consuming and not suitable for real-time mail image processing systems. Table 1 summarizes some of the postal address recognition results related to the Chinese language reported in literature.

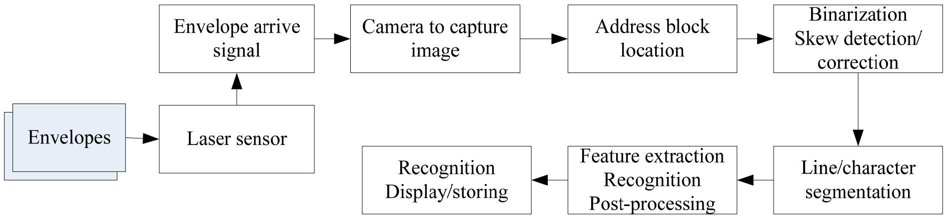
Table 1. Previous results of Chinese related automatic postal address recognition methods reported in the literature.

Method	Address Font Style	Testing Condition (CPU)	Recognition Rate (%)	Processing Speed (mails/hour)	Description
Jiang *et al.*, 2007 [2]	Handwritten Chinese	3.0 GHz	85.3	1,200	Input: envelope image Output: sorting result
Su *et al.*, 1998 [8]	Handwritten/machine printed Chinese	Pentium 133	75.6	5,400	Input: envelope image Output: sorting result
Liu *et al.*, 2002 [9]	Handwritten Chinese	Pentium III 600 Hz	83.68	Around 110 ms	Input: manually segmented address line Output: recognized address string
Han *et al.*, 2000 [14]	Handwritten Chinese	–	96.26	–	Input: envelope image Output: sorting result
Kato *et al.*, 2000 [10]	Handwritten Japanese	–	89.2	–	Input: envelope image Output: sorting result
Wang *et al.*, 2004 [15]	Handwritten Chinese	–	96.45 (character correct rate)	–	Input: address string image Output: recognized address string
Huang *et al.*, 2003 [12]	Handwritten Chinese	–	70	–	Input: envelope image Output: sorting result
Our method	Machine-printed Chinese	2.8 GHz	98.72 (address) 99.74 (postcode)	44,237	Input: envelope image Output: recognized address

"–" indicates that no detailed information was reported in the paper.

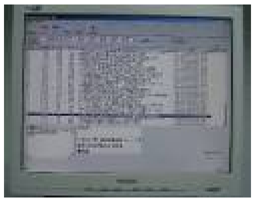
2. SYSTEM OVERVIEW

The proposed postal address recognition system uses a high-speed camera to capture images of the envelopes running on the conveyer belt of a mail-processing device, and the camera is triggered by a laser sensor. When an envelope arrives at the exact preset position under the camera, it is detected by a laser sensor, which then sends an envelope-arrival signal and triggers the camera to capture the image. The envelope image is then fed to a personal computer for image processing and recognition. The processing and recognition procedure includes the destination address block location, image binarization, skew detection/correction, line/character segmentation, feature extraction, character recognition, and post-processing of the recognized address. After processing and recognition are finished, the recognition results along with the corresponding envelope images are sent to a database on another personal computer in real-time. The database can then be used in subsequent envelope querying, tracking and management. Fig. 1 (a) shows a block diagram of the proposed system.



(a) Block diagram of the system.

(b) System setup.



(c) Interface of recognition results database.

Fig. 1. Overview of the proposed system.

In our system, the laser sensor is located a short distance in front of the camera along the envelope conveyer belt, which is shown in black (Fig. 1 (b)). When the front edge of the envelope reaches the laser sensor, its brighter color causes the changes in the reflected light of the sensor. A single-chip controller then computes the arrival time of the envelope under the camera in terms of the speed of conveyer belt and the distance between the sensor and the camera, which then triggers the camera. To accurately capture the image of the destination address block on the fast moving envelopes, the envelope arriving signal must be as exact as possible. In our system, the position error given by the laser sensor is less than 5mm, thus the envelope image can be obtained exactly. The captured image has a maximum resolution of 1280×1024 pixels and the default resolution used in our system is 1280×700. Meanwhile, to avoid the motion blur of the image for quickly moving envelopes, light sources are needed for image enhancement. In our system, a strobe light source with a maximum trigger frequency of 200 Hz is used. Our experimental results show that this type of light source can effectively improve the quality of the captured envelope images.

After the processing and recognition of the envelope images, the recognition results along with the corresponding envelope images will be sent to a database as shown in Fig. 1 (c). The recognition results database contains the fields of postcode, province name, town/county name, postal address, recognition confidence and assigned barcode for each envelope. This information can conveniently be used in querying, sorting and reediting of recognition results. In addition, to reduce storage, only the located block of postcode and address are stored in the database.

3. MATERIALS AND METHODS

To satisfy the need for real-time processing of moving envelopes in our system, fast processing and recognition algorithms must cooperate with the hardware system. In this section, we give the detail of the methods used in our system.

3.1 Destination Address Block Location

The detection and location of the destination address block (DAB) in each envelope image is the foundation of subsequent processing and recognition in the postal address reading system. The incorrect location of DAB will greatly affect the overall performance of the system. In this section, a fast DAB location algorithm is proposed to cooperate with our hardware system. Our algorithm is based on the following reasonable assumption: For each batch of envelopes to be processed, the distance that the envelopes passed from the laser sensor to the position need to trigger the camera to capture the DAB image, are roughly the same. Although different batches of envelopes can have the different position of DAB on the envelopes, the preset parameter of distance that determines the triggering time in the single-chip controller can be easily adjusted. Therefore, we are able to guarantee that the vast majority of captured images of envelopes will contain the address block for envelopes of different layouts. To locate the edges of DAB, we adopted the following algorithm, based on several observations. First, in the region of address block, the stroke density is generally larger than that of the background region. Then, if we scan the image horizontally (or vertically), the gray value distribution of pixels for the lines across the stroke pixels will be quite different from that of the background pixels, and the stroke edge pixels will generally have larger gradients, as shown in Fig. 2.

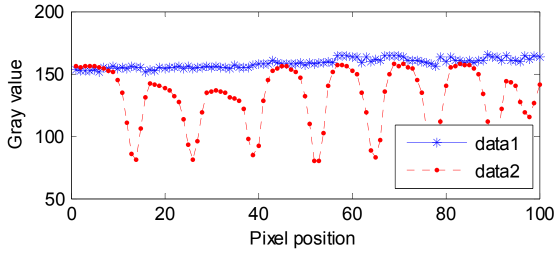


Fig. 2. Gray value distribution of pixels of portions of horizontal scan lines (data1 refers to the scan line across the background pixels, and data2 refers to the scan line across the stroke pixels).

For simplicity, in the illustration we only give the procedure of locating the upper and lower edges here and locating the left and right edges follows a similar procedure. Let $f(i, j)$ be the gray value of pixel (i, j). The horizontal gradient $gradh(i, j)$ is defined as Eq. (1). If $gradh(i, j)$ is larger than a threshold T_{gradh}, then the pixel (i, j) is considered as a stroke pixel. T_{graph} is determined by experimental testing. For each row i of envelope image, we scan each pixel and count the number of the stroke pixels (denoted as $snh(i)$).

$$gradh(i, j) = (abs(i, j-1) - f(i, j+1)) \tag{1}$$

where $abs(.)$ represents the absolute value.

Let i_{maxsnh} be the row with maximum $snh(i)$, we scan from i_{maxsnh} upward/downward, respectively. If there exist T_{bl} continuous rows whose $snh(.)$ is less than a threshold, the current row is set as the upper (or lower) edge of the address block. In our experiments,

T_{bl} is set to 50. The experimental results with a large number of real mail indicated that our algorithm can locate the address block effectively and extremely quickly.

3.2 Binarization and Skew Correction

Before line/character segmentation can be performed in our system, fast algorithms for binarization and skew detection/correction of envelope images are needed. Binarization methods for document images have been explored extensively in the last several decades, and can be roughly grouped into global methods and locally adaptive methods [17, 18]. Generally, the global methods are faster; however, they are easily affected by sources of noise such as deficiency in illumination and exposure. On the other hand, locally adaptive methods are more effective, especially for resisting image noise. Ye proposed a fast binarization method based on a local extreme value, which can be regarded as a modified version of Bernsen's locally adaptive method [18]. Our experimental results show that this method gives the best result when both processing speed and recognition performance are considered. However, this algorithm takes 134.7 milliseconds for an image with a resolution of 640 × 320, which is not sufficient for our system. In this paper, we propose a binarization algorithm that combines features of the global method and the locally adaptive method. The basic idea is that, if the envelope image is divided into a few local sub-regions, and one global threshold is found for each sub-region, the problems of deficiency in illumination and exposure can be avoided to some extent. The algorithm is as follows:

Suppose that the envelope image $f(i, j)$ is evenly divided into N sub-regions $R(k)$, $k = 1, \ldots, N$. For each sub-region $R(k)$, its binarization threshold is denoted as $T_{reg}(k)$, and its maximum and minimum gray values of pixels is $fmax(k)$ and $fmin(k)$, respectively. For each sub-region $R(k)$, it is first labeled as foreground or background region in terms of difference between $fmax(k)$ and $fmin(k)$. And then, the median of gray vales is adopted as the binarization threshold for foreground region. Thus, the threshold $T_{reg}(k)$ follows the Eq. (2). The size of sub-regions can be used to control the locality of algorithm. If the sub-region size is set equal to the image size ($N = 1$), our algorithm is degraded to a global approach. If the sub-region size is set too small, for example, less than the width of strokes, some sub-region of stroke can be mis-labelled as background region. In our experiments, the size of sub-regions is set to 30 × 30.

$$T_{reg}(k) = \begin{cases} (fmax(k) + fmin(k))/2, & \text{if } fmax(k) - fmin(k) > T_1 \\ fmin(k), & \text{otherwise} \end{cases} \tag{2}$$

where T_1 is a preset constant determined by experiments, we use $T_1 = 30$ in our experiments.

To keep the smoothness of the binarization thresholds of pixels in neighboring sub-regions, a linear threshold interpolation technique is applied as follows. Let $T_{img}(i, j)$ be the binarization threshold of pixel (i, j) which belongs to sub-region $R(k_1)$ with the center coordinate (i_1, j_1), $R(k_2)$ and $R(k_3)$ are the closest sub-regions of pixel (i, j) in vertical and horizontal directions, respectively, and the corresponding sub-region centers are (i_2, j_2) and (i_3, j_3). Then, we have Eqs. (3)-(5):

$$T_{img}(i, j) = (TH_1(i, j) + TH_2(i, j))/2 \tag{3}$$

$$TH_1(i, j) = T_{reg}(k_1) + \frac{T_{reg}(k_2) - T_{reg}(k_1)}{i_2 - i_1}(i - i_1) \tag{4}$$

$$TH_2(i, j) = T_{reg}(k_1) + \frac{T_{reg}(k_3) - T_{reg}(k_1)}{j_3 - j_1}(j - j_1) \tag{5}$$

The experimental results show that our binarization algorithm takes only 1.18 milliseconds for an envelope image of 640 × 320, with satisfactory character recognition results. Thus, this technique was found to be substantially faster than Ye's method [18].

For skew detection, we use a projection profile based method. The range of this method is limited to ± 15°, and the precision is set to 0.1°. The skew correction method proposed in [19], referred to as the linear whole block moving method, is used in our system. Further details of this method can be found in [19].

3.3 Line/character Segmentation

After skew correction, we use a projection profile analysis based line segmentation method to locate the address lines in the address block of the envelope image. Once an address line is located, an over-segmentation strategy is adopted for character segmentation. An address line is first segmented into a sequence of sub-segments using projection profiles. For sub-segments that are too small or too large, the aspect ratio, interval, maximum and minimum width constraints, and other heuristics about envelope layout are used to re-segment the sub-segments. From these sub-segments, a candidate segmentation path network is constructed with each node representing a potential segmentation point, each arc representing the cost for that segmentation point in terms of scores given by a classifier. Finally, a dynamic programming algorithm is used for finding the best segmentation path together with its recognition results [20].

3.4 Feature Extraction

Feature extraction is one of the most important components affecting processing speed and recognition performance. In our system, we adopt two different feature extraction methods for performance comparison. One is the Gabor feature method, which is either based on binary character images (denoted as **GaborBin**), or gray character images (denoted as **GaborGray**). The other method is the gradient feature method, based on gray character images (denoted as **GradGray**). Gray image-based character feature methods can avoid the noise introduced by the binarization procedure, and thus may produce better recognition performance. The experimental results indicated that gradient features showed a better recognition rate with comparative processing speed. In the following section, we briefly explain the feature extraction procedures for Gabor features and gradient features.

3.4.1 Gabor features

Gabor features have been widely used in many pattern recognition fields such face recognition and character recognition. An important property of Gabor features is that they can achieve a joint optimal resolution in both the spatial and spatial-frequency domains. In our system, we adopt the 2-D Gabor filter reported in [21], as shown in Eqs. (6)-(8).

$$G(x, y; \kappa, \theta_k) = G_1(x, y)[\cos(R) - \exp(-\frac{\sigma^2}{2})] + iG_1(x, y)\sin(R) \tag{6}$$

$$G_1(x, y) = \frac{\kappa^2}{\sigma^2}\exp[-\frac{\kappa^2(x^2 + y^2)}{2\sigma^2}] \tag{7}$$

$$R = \kappa(x\cos\theta_k + y\sin\theta_k) \tag{8}$$

where $\kappa = \frac{2\pi}{\lambda}$, $\theta_k = \frac{\pi k}{M}$, $k = 0, \ldots, M - 1$, $\sigma = \pi$, and the parameters λ and θ_k are the wavelength and orientation, respectively. M is the number of orientation, and it is set to 4 in our experiments.

Let $f(x, y)$ be the binary/gray character image. Then, we have Gabor feature for the sampling point (x_0, y_0) as shown in Eq. (9), where S is a constant parameter. We use S = 10 in our experiments.

$$F_{\lambda,k}(x_0, y_0) = \left|\sum_{x=-S}^{S}\sum_{y=-S}^{S} f(x_0 + x, y_0 + y)G(x, y; \kappa, \theta_k)\right| \tag{9}$$

3.4.2 Gradient features

Gradient features have been proven to be the most effective and time efficient features for use in character recognition [22]. In our system, gradient feature extraction is directly applied on gray character images. Thus, the noises introduced by binarization procedure can be avoided.

To extract gradient features, a 3 × 3 Sobel operator (as shown in Fig. 2) is used to obtain the horizontal and vertical gradient at each pixel, respectively. The character image is then decomposed into a number of regions corresponding to L directions with an equal interval $2\pi/L$ (L is set to 8 in our experiments), and the gradient vector of each pixel is assigned to its two nearest directions in a parallelogram manner, as illustrated in Fig. 3. In this way, an L-dimensional gradient code can be formed at each pixel, and character image is decomposed into L directional pattern images.

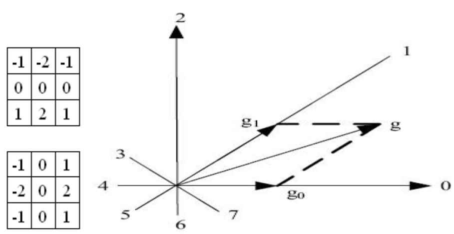


Fig. 3. Sobel operators and decompsing of the gradient vector.

Let $f_d(x, y)$, $d = 0, \ldots, L - 1$ be the directional pattern image, then we have gradient feature for the sampling point (x_0, y_0) as shown in Eq. (10).

$$F_d(x_0, y_0) = \sum_{x=-S}^{S} \sum_{y=-S}^{S} f_d(x_0 + x, y_0 + y) G_1(x, y) \tag{10}$$

To extract the character feature, the character image is uniformly divided into $n \times n$ cells (8×8 in our experiments), and the center of each cell is adopted as the sampling point. Therefore we obtain a 256 ($4 \times 8 \times 8 = 256$)-dimension Gabor feature and a 512 ($8 \times 8 \times 8 = 512$)-dimension gradient feature, respectively.

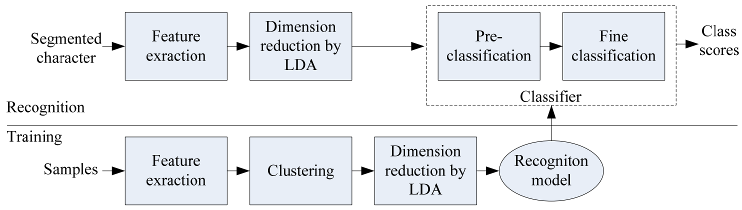


Fig. 4. Block diagram of the training and recognition process used in our system.

3.5 Recognition

Fig. 4 shows a block diagram of the training and recognition process in our system. For recognition, we adopt a two-stage classification strategy, which includes pre-classification and fine classification. Fine classification is limited within the candidate classes given by pre-classification. A minimum Euclidean distance classifier is used for each stage. At the pre-classification stage, one recognition prototype is used for each class of characters to speed up the recognition. In fine classification, four recognition prototypes for each class is formed by *k*-means clustering method. To improve recognition performance, linear discriminative analysis (LDA) is used to reduce the number of feature dimensions. In our experiments, the number of dimensions of raw features was reduced to 30 in pre-classification, and to 96 in fine classification.

Post-processing is another technique adopted in our system to improve recognition performance. For each envelope, we can obtain a postcode string and an address string after recognition, respectively. In China, postcodes are 6-digit codes that correspond to the mail address on the envelope. For example, the first three digits “510” in postcode refer to the city of GuangZhou in GuangDong province. Thus, we can use the correspondence between the postcode and address to correct the recognition results. In our system, a postcode database is constructed which contains 2,262 items of pairs of postcodes and their corresponding addresses for the provinces and main towns/counties in China. And the following algorithm is used to process the recognition results, where the edit distance is used for measuring the matching between the two name strings.

Step 1: Search for the keywords “sheng (省)” and “shi/xian (市/县)” in the recognition results, meaning the province and town/county in Chinese, respectively. If the keywords can be located, then run step 2, otherwise run step 3.

Step 2: Match the located key names of province and town/county in postcode database, and then run step 4.

Step 3: If the keywords could not be located, we search for the names of province and town/county from the beginning of the recognized address string directly, and run step 4.

Step 4: If the best matched item of located town/county name in postcode database is fully matched, then modify the recognized province name and postcode according to the postcode database, and the algorithm terminates. Otherwise, run step 5.

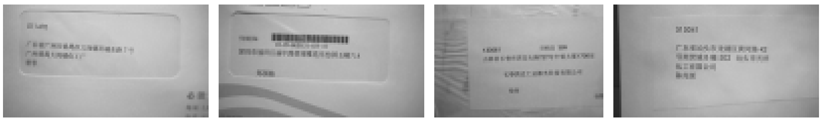
Step 5: If the best matched item of located town/county name in postcode database is partially matched and the corresponding postcode is fully matched, then we will modify the recognized province name and town/county name according to the postcode database. Otherwise the algorithm terminates.

3.6 Data Collection

Since a large difference exists between handwritten samples of Chinese characters and their machine-printed counterparts, the present handwritten Chinese character databases cannot be used for training our recognition engine. To evaluate the performance of our system and construct a recognition model in the training phase, a database of machine-printed Chinese characters and a live mail image database were constructed. The database of machine-printed Chinese characters contains 4,516 frequently used Chinese characters defined in GB2312-80 standard, 62 alphanumeric characters, and 12 punctuation marks and symbols. Each character had 100 samples collected with different printed font size and illumination using the same high-speed camera as in our envelope identification system. Each character was printed onto paper similar to that of ordinary envelopes with a laser printer, and the envelope image processing algorithms such as binarization and skew detection/correction described in section 3 were applied to obtain isolated character samples. The font size of printed characters ranged from 7.5-point to 14-point. It should be noted that some manual editing is needed to correct possible mis-segmentation. The real mail image database contains 761 mail images with a total of 25,060 postcode and address characters. The postal addresses used were collected randomly from some Chinese websites, then printed onto envelopes of different formats with a laser printer. The font size of printed characters also ranges from 7.5-point to 14-point. The envelope images were captured using our system operating in image acquisition mode. The image resolution used is 1280 × 700 pixels. The ground truth for the postcode and address of each mail image is edited manually as a benchmark for recognition test. Some samples from our databases are shown in Fig. 5.

(a) Printed samples of isolated characters. (b) Samples of isolated characters database.

Fig. 5. Samples from the databases used in our experiments.



(c) The samples of the real mail image database.
Fig. 5. (Cont'd) Samples from the databases used in our experiments.

4. EXPERIMENTAL RESULTS

4.1 Recognition Test

To test the performance of the proposed system, three categories of experiments were carried out on a personal computer running Microsoft Windows XP with a 2.8 GHz CPU. The first category is recognition experiments performed on our collected database of isolated machine printed characters to test the performance of feature extraction and recognition methods. Some character samples of the database are shown in Fig. 5. The database contained 100 samples with different printed font sizes and illuminations for each character, 70 samples of which are used as training data to create the recognition models, and the other 30 samples as the testing data.

In these experiments, three kinds of character features were tested, as listed in section 3.4: the Gabor feature based on the binary character image (**GaborBin**), the Gabor feature based on the gray character image (**GaborGray**), the gradient feature based on the gray character image (**GradGray**). We adopted a recognition strategy, as shown in section 3.5. The performance, in terms of recognition error rates and processing speed on the isolated machine printed characters database are summarized in Table 2, where the feature time refers to the average feature extraction time of one character sample, and the recognition time is the average recognition time of one sample including the dimension reduction, pre-classification and fine classification operations.

Table 2. Performance on the isolated characters database.

Feature	Feature time (ms)	Recognition time (ms)	Recognition rate (%)
GaborBin	0.27	1.13	98.95
GaborGray	0.38	1.13	98.18
GradGray	0.45	1.28	99.49

From Table 2, it can be seen that the Gabor feature based on the binary character image was the fastest. The gradient feature based on the gray character image exhibited the best recognition performance, but a slower speed. Comparatively, the Gabor feature based on the gray character image had no advantage, either in processing speed or recognition rate. Thus, it is not used in our system.

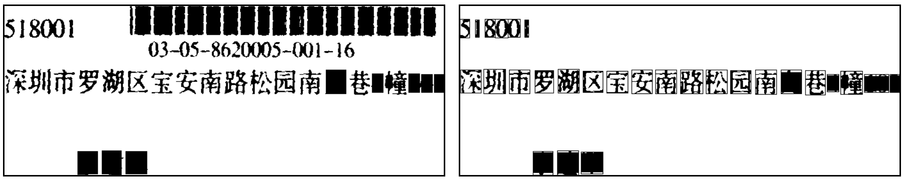
The second category of experiments was carried out on a mail images database collected with real mail, to test the overall performance of our system. The database contains 761 mail images with a total of 25,060 postcode and address characters. The ground truth for the postcode and address of each mail image is edited manually as a benchmark for the recognition test. For each mail image, it is processed and recognized with the methods described in this paper, and the recognition performance is evaluated using edit distance according to Eq. (11)

$$ErrorRate = \frac{edit\ distance}{character\ number} \times 100\% \tag{11}$$

where edit distance refers to the edit distance between the recognition result and its ground truth, and the character number refers to the character number of the corresponding ground truth.

An example of the mail image processing and recognition procedure by our system is shown in Fig. 6, and the experimental results are summarized in Tables 3 and 4. Table 3 shows the recognition performance, where no post-processing means the recognition test without using the post-processing technique, and postcode and address refer to the recognition results for postcode and address, respectively. In Table 4, the CPU times for each stage of our system are given, where preprocessing refers to the system stages that include the address block location, binarization, skew detection/correction, line segmentation and character over-segmentation, while recognition refers to the system stages that include dynamic programming based character segmentation, feature extraction, recognition and post-processing.

(a) Captured mail image. (b) Located block of the postal address and postcode. (c) Binarization result of the block of postal address and postcode.



(d) Block of postal address and postcode after skew detection/correction. (e) Segmentation results of postal address and postcode.

Postcode: 518001 & Address: 深圳市罗湖区宝安南路松园南*巷*幢******

(f) Recognition results.

Fig. 6. Example of the mail image processing and recognition procedure used in our system. (From the image, the name of recipient and some detail address have been hidden to maintain privacy. The barcode on the top of image has been omitted in the processing of segmentation and recognition.)

Table 3. Recognition performance on the machine-printed mail images database (%).

Feature	No post-processing		Post-processing	
	Postcode	*Address*	*Postcode*	*Address*
GaborBin	97.15	98.35	99.69	98.41
GradGray	97.24	98.70	99.74	98.72

Table 4. Average CPU times of our system on machine-printed mail images database (ms).

Feature	Preprocessing	Recognition	Overall System	Throughput (mails/hour)
GaborBin	19.42	40.81	60.23	59,770
GradGray	26.49	54.89	81.38	44,237

From Table 3, it can be seen that our system achieved a recognition rate of 99.74% and 98.72% for postcodes and addresses, respectively. This means that only 0.64 characters were mis-identified for the average postal address item out of an average of 50 characters, indicating that the system has a high level of effectiveness. In addition, the recognition rate of the postcode can be improved with post-processing.

From Table 4, it can be seen that our system achieved a processing speed of 59,770 and 44,237 mails per hour for the GaborBin feature and GradGray feature, respectively. This is much faster than other systems reported in the literature, indicating a high level of efficiency.

Finally, our system was implemented with a common postal envelope printer and tested on site with real mail at a postal service center. The envelope conveying system of the postal envelopes printer can guarantee that the envelopes are transmitted on the conveyer belt fast and automatically. In this system, a high-speed camera is used to capture the images of envelopes running on the conveying device and a laser sensor is equipped to detect the mail arriving signal. The captured images are fed into our system for mail identification. The recognition results and corresponding envelope images are saved in the database in real-time. The experimental results are given in Table 5. In this experiment, the Gabor feature based on the binary character image was used, and the recognition rate was computed by checking the recognition results manually. A total of 75 envelopes representing 2,470 characters of postal addresses were tested, and the 'throughput' (mails/hour) was measured by the conveying device.

Table 5. The performance of our system tested on site with live mails.

Feature	Font size (points)	Recognition rate (%)	Throughput (mails/hour)
GaborBin	no less than 10.5	98.92	21,000
GaborBin	7.5-9	94.44	21,000

From Table 5, it can be seen that the recognition rate reached 98.92% for envelopes with a printed font size of no less than 10.5 points, and 94.44% for envelopes with printed font size ranging from 7.5 to 9 points. The processing speed was 21,000 mails per hour. It should be noted that the processing speed in Table 4 was determined for envelopes with a maximum of 50 characters of postal address for the need of real-time processing, and that the processing time consists of not only the CPU times reported in Table 3, but also the times for obtaining the envelope images by the high-speed camera and transferring and saving in real-time the recognition results into the database located on another personal computer. In addition, the recognition performance of envelopes with font sizes ranging from 7.5 to 9 points can be further improved by zooming in the lens for practical applications.

4.2 Analysis of Mis-recognition

To identify the causes of mis-recognition by our system and further improve the system performance, we randomly selected a number of test samples of envelope images that were mis-recognized and checked the causes of each recognition error manually. According to the processing stages applied in our algorithm, the causes of mis-recognition can be classified into four error types. **Type A** refers to the mis-recognition caused by the processing of the address block location and line/character segmentation, **Type B** refers to the mis-recognition by binarization, **Type C** refers to the recognition confusion caused by similar characters such as “1” and “I”, and **Type D** refers to other classification errors caused by the classifier. The samples include a total of 50 recognition errors, and the statistical results are listed in Table 6. Fig. 7 shows some examples of the mis-segmentation and mis-recognition, where the first row is from Type B, the second row from Type C, the third row from Type D. The last row of Fig. 7 shows some examples of Type A, where envelope images with a dirty background caused mistakes in the location and segmentation, and consequently mis-recognition.

Table 6. The statistical results of mis-recognition.

Recognition error type	Type A	Type B	Type C	Type D
Error number	6	16	19	9
Rate	12%	32%	38%	18%

From Table 6, it can be seen that, among all of the recognition errors, Type A errors accounted only for 12%, while more than half of the errors were related directly to the classifier (Type C and D). According to our investigation, Type A errors are mainly due to dirty backgrounds of envelope images, which can be caused by unexpected ink and creases as shown in Fig. 7. The dirty background can lead to falsely segmented characters in subsequent processing stages. To reduce this type of false segmentation, we can adopt a rejection strategy or verification models at the classification stage, since false segmentation will generally have lower scores from the classifier. Type B errors arise from the binarization process. Although we usually obtained correct segmentation in these cases, as shown in the first row of Fig. 7, the segmented characters may include unwanted strokes or lose some strokes in binarization, due to noise or degradation of the images. For example, some strokes in the machine-printed Chinese character “王” in the Song font can be very fine, and if image blurring occurs some strokes can be lost in binarization. As our experiments have shown, the Type B errors can be reduced to some extent by using the gray image based character feature such as the GradGray. Meanwhile, image enhancement techniques may be needed in some cases. Type C errors are mainly due to the confusion among similar characters. Recognition of similar characters is a difficult task in the field of character recognition, since the differences in shape between some similar characters is very small, such as the character pairs “G” and “C”, “1” and “I”, *etc*. To ameliorate this type of recognition confusion, a language model may be needed in recognition post-processing. In addition, we also need to improve our recognition engine to identify Type C and D mis-recognition errors.

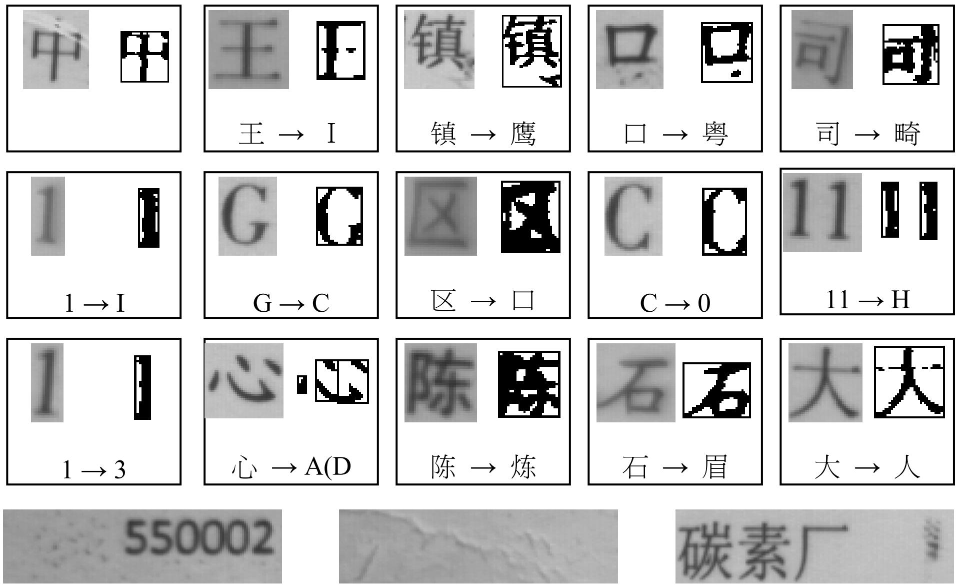


Fig. 7. Some typical examples of mis-segmentation and mis-recognition by our system. In each box, the top left shows the input character image (cropped from the envelope image for display), the top right is the segmented character produced by our system, and below are the ground truth and recognition results, respectively. The last row shows some examples of envelope images with dirty backgrounds.

5. CONCLUSIONS

Automatic postal address recognition is one of the most important applications of offline OCR, which aims to sort mail automatically, or to assign barcodes based on recognition of addresses for subsequent mail sorting, tracking and management. This paper focuses on developing a method of real-time processing of machine-printed Chinese postal envelopes. We propose a system in which a laser sensor is used to detect a signal of the envelopes' arrival, enabling a vision-based fast postal envelope identification system for moving machine-printed Chinese postal envelopes. Our system uses a high-speed camera to capture images of envelopes running on the conveying device, and then recognizes the postal address and postcode on each envelope image. The experimental results revealed that the proposed system achieved high-speed processing and accurate recognition of characters on moving envelopes. Envelopes with an average of 32.9 characters were processed and recognized within 81.38 milliseconds, with a character recognition rate of 98.72% for postal addresses.

Furthermore, our system also provides a method for storing the envelope images and their recognition results into a database in real-time, which is convenient for subsequent envelope tracking and management. The results of an experimental trial with real mail on site at a postal center indicated that our system could reach a speed of approximately 21,000 envelopes per hour, with a character recognition rate as high as 98.92% for postal addresses. In addition, our system can be conveniently implemented with pre-existing envelope processing devices commonly used in postal service centers. The proposed system is also applicable to other OCR applications besides postal automation, such as automatic recognition of machine-printed labels in logistics systems, electronic archiving of document materials, *etc*.

Although our system has achieved more promising results in the recognition of machine-printed postage envelopes, the recognition performance can be further improved in several aspects. One strategy under consideration is to adopt more effective classification techniques in the stage of fine classification. For example the modified quadratic discriminant function (MQDF) classifier [23, 24] which has been successfully applied to recognition of handwritten Chinese characters to achieve very good performance. Since fine classification is limited within the candidate classes given by pre-classification, more complicated algorithms will not lead to too much loss in recognition speed. On the other hand, applying the language model in post-processing can also be expected to be able to improve the recognition performance. Chinese address is a type of language expression with a special grammar format, which usually takes a hierarchical multi-layer structure from the province/city to the street and house number, and each layer has a specific keyword to identify. Thus, if we can use this information to design an effective language model, a higher recognition rate can be expected.

REFERENCES

1. U. Mahadevan and S. N. Srihari, "Parsing and recognition of city, state, and ZIP codes in handwritten addresses," in *Proceedings of the 5th International Conference on Document Analysis and Recognition*, 1999, pp. 325-328.
2. Y. Jiang, X. Q. Ding, and Z. Ren, "A suffix tree based handwritten Chinese address recognition system," in *Proceedings of the 9th International Conference on Document Analysis and Recognition*, 2007, pp. 292-296.
3. D. Gaceb, V. Eglin, F. LeBourgeois, and H. Emptoz, "Physical layout segmentation of mail application dedicated to automatic postal sorting system," in *Proceedings of the 8th IAPR International Workshop on Document Analysis Systems*, 2008, pp. 408-414.
4. P. S. R. Prasanna, S. Balaji, T. H. Khezhie, C. Vasanthanayaki, *et al*., "Destination address interpretation for automating the sorting process of Indian postal system," in *Proceedings of International Conference on Convergent Technologies for Asia-Pacific Region*, 2003, pp. 858-862.
5. S. N. Srihari and E. J. Keubert, "Integration of hand-written address interpretation technology into the United States postal service remote computer reader system," in *Proceedings of the 4th International Conference on Document Analysis and Recognition*, 1997, pp. 892-896.
6. R. Plamondon and S. N. Srihari, "On-line and off-line handwritten recognition: a comprehensive survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, 2000, pp. 63-84.
7. T. Akiyama, D. Nishiwaki, and E. Ishidera, "Handwritten address interpretation system allowing for non-use of postal codes and omission of address elements," in *Proceedings of the 9th International Workshop on Frontiers in Handwriting Recognition*, 2004, pp. 527-532.

8. Y. M. Su and J. F. Wang, "Recognition of handwritten Chinese postal address using neural networks," in *Proceedings of IEEE International Symposium on Circuits and Systems*, 1998, pp. 25-28.
9. C. L. Liu, M. Koga, and H. Fujisawa, "Lexicon-driven segmentation and recognition of handwritten character strings for Japanese address reading," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, 2002, pp. 1425-1437.
10. N. Kato, K. Tokumoto, Y. Nemoto, and M. Suzuki, "A large scale Japanese handwritten address recognition system using rough and fine classification," in *Proceedings of the 5th International Conference on Signal Processing*, 2000, pp. 1423-1426.
11. K. Roy, S. Vajda, U. Pal, B. B. Chaudhuri, and A. Belaid, "A system for Indian postal automation," in *Proceedings of the 8th International Conference on Document Analysis and Recognition*, 2005, pp. 1060-1064.
12. L. Huang and C. P. Liu, "Software system of automatic mail sorting machine," *Computer Engineering and Applications*, Vol. 39, 2003, pp. 21-24.
13. J. L. Xue, X. Q. Ding, C. S. Liu, R. Zhang, and W. W. Qian, "Location and interpretation of destination addresses on handwritten Chinese envelopes," *Pattern Recognition Letters*, Vol. 22, 2001, pp. 639-656.
14. Z. Han, C. P. Liu, and X. C. Yin, "A two-stage handwritten character segmentation approach in mail address recognition," in *Proceedings of the 8th International Conference on Document Analysis and Recognition*, 2005, pp. 111-115.
15. S. H. Jeong, S. I. Jang, and Y. S. Nam, "Locating destination address block in Korean mail images," in *Proceedings of the 17th International Conference on Pattern Recognition*, 2004, pp. 387-390.
16. C. H. Wang, Y. Hotta, M. Suwa, and N. Naoi, "Handwritten Chinese address recognition," in *Proceedings of the 9th International Workshop on Frontiers in Handwriting Recognition*, 2004, pp. 539-544.
17. O. D. Trier and A. K. Jain, "Goal-directed evaluation of binarization methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, 1995, pp. 1191-1201.
18. Y. X. Ye, F. H. Qi, and J. Y. Wu, "Fast binarization algorithm for document image," *Journal of Infrared and Millimeter Waves*, Vol. 16, 1997, pp. 344-350.
19. D. L. Ming, J. Liu, J. H. Hu, and H. T. Li, "The slanting image detecting and rectifying technology in OCR system," *Journal of Huazhong University of Science and Technology*, Vol. 28, 2000, pp. 66-68.
20. X. Gao, P. M. Lallican, and C. Viard-Gaudin, "A two-stage handwritten Chinese character segmentation algorithm based on dynamic programming," in *Proceedings of the 8th International Conference on Document Analysis and Recognition*, 2005, pp. 735-739.
21. Q. Huo, Y. Ge, and Z. D. Feng, "High performance Chinese OCR based on Gabor features, discriminative feature extraction and model training," in *Proceedings of IEEE International Conference on Acoustics*, *Speech*, *and Signal Processing*, 2001, pp. 1517-1520.
22. C. L. Liu, K. Nakashima, H. Sako, and H. Fujisawa, "Handwritten digit recognition: investigation of normalization and feature extraction techniques," *Pattern Recognition*, Vol. 37, 2004, pp. 265-279.
23. F. Kimura, K. Takashina, S. Tsuruoka, and Y. Miyake, "Modified quadratic discrimi-

nant functions and the application to Chinese character recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 9, 1987, pp. 149-1153.
24. T. Long and L. W. Jin, "Building compact MQDF classifier for large character set recognition by subspace distribution sharing," *Pattern Recognition*, Vol. 41, 2008, pp. 2916-2925.
25. X. Gao and L. Jin, "SwiftPost – A vision-based fast postal envelope identification system," in *Proceedings of 2009 IEEE International Conference on Systems*, *Man*, *and Cybernetics*, 2009, pp. 3368-3373.

Xue Gao (高学) received the Ph.D. degree in Electronic Science and Technology from South China University of Technology in 2003. From 2004 to 2005, he worked as a postdoctoral researcher at Ecole Polytechnique de l'Université de Nantes, France. He is now an associate professor at School of Electronic and Information Engineering, South China University of Technology. His current research interests include document analysis and recognition, machine learning.

Lianwen Jin (金连文) is a member of the IEEE and IEEE Computer Society. He received a B.S. degree from the University of Science and Technology of China and a Ph.D. degree from the South China University of Technology in 1991 and 1996, respectively. He visited the University of Hong Kong as a Research Fellow in 2002 and 2006 respectively. Now he is a Professor at the School of Electronic and Information Engineering, South China University of Technology. In 2006, he received the New Century Excellent Talent Program Award from the China MOE. He has published more than 90 papers in the fields of handwritten Chinese character recognition, image processing, and pattern recognition. His current research interests include character recognition, pattern analysis and recognition, image processing, machine learning and intelligent systems. He served as Program Committee member for a number of international conferences, including IWFHR2006, ICFHR2008, ICDAR 2009, ICMLC2009, ICPR2010, ICMLC2010, ICDAR2011, *et al*.