Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition 40 (2007) 99–107
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/patcog

Handwritten Bangla numeral recognition system and its application to postal automation

Ying Wen[a,b,*], Yue Lu[b,c], Pengfei Shi[a]

[a]Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, Shanghai 200030, China
[b]Department of Computer Science and Technology, East China Normal University, Shanghai 200062, China
[c]Shanghai Research Institute of Postal Science, China State Post Bureau, Shanghai 200062, China

Received 23 November 2005; received in revised form 11 June 2006; accepted 4 July 2006

Abstract

A recognition system for handwritten Bangla numerals and its application to automatic letter sorting machine for Bangladesh Post is presented. The system consists of preprocessing, feature extraction, recognition and integration. Based on the theories of principal component analysis (PCA), two novel approaches are proposed for recognizing handwritten Bangla numerals. One is the image reconstruction recognition approach, and the other is the direction feature extraction approach combined with PCA and SVM. By examining the handwritten Bangla numeral data captured from real Bangladesh letters, the experimental results show that our proposed approaches are effective. To meet performance requirements of automatic letter sorting machine, we integrate the results of the two proposed approaches with one conventional PCA approach. It has been found that the recognition result achieved by the integrated system is more reliable than that by one method alone. The average recognition rate, error rate and reliability achieved by the integrated system are 95.05%, 0.93% and 99.03%, respectively. Experiments demonstrate that the integrated system also meets speed requirement.

Keywords: Bangla numeral recognition; Support vector machine; Principal component analysis; Feature extraction

1. Introduction

Based on the requirements of Bangladesh Post, China Post designs and manufactures letter sorting machines with the ability of automatically recognizing postal numerals on the envelope images. To meet the requirements of a practical usage, we need take into account not only the recognition rate as usually reported in most character recognition literature, but also the error rate, the reliability and the response time.

Bangla is the second-most popular language in the Indian subcontinent. Unfortunately, researches on Bangla character recognition are not sufficient so far, in particular on handwritten Bangla character recognition issue. Some papers on printed Bangla numeral recognition have been reported in past years [1–3], but there is few research on handwritten Bangla numeral recognition. Professor Pal et al. have done some exploring work for the issue of recognizing handwritten Bangla numerals [4–6]. Their proposed scheme is mainly based on features obtained from a concept called water reservoir. Reservoir is obtained by considering accumulation of water poured from the top or from the bottom of the numerals. They presented a system towards Indian postal automation. The accuracy of the handwritten Bangla and English numeral classifier is 94.13% and 93%, respectively. However, they did not report in their papers the recognition reliability and the response time, which are very important evaluation factors for a practical automatic letter sorting machine. Reliability reflects the relation between the error rate and the recognition rate. The higher is the reliability, the lower is the error rate. We expect to get the higher reliability and the acceptable response time which indicates the efficiency of the letter sorting machine. It is necessary to do the further researches on the handwritten Bangla numeral recognition.

* Corresponding author. Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, Shanghai 200030, China. Tel.: 86 21 62602899; fax: 86 21 62437035.
E-mail address: yingwen@sjtu.edu.cn (Y. Wen).

0031-3203/$30.00
doi:10.1016/j.patcog.2006.07.001

In the past decades, researchers have proposed two feature extraction for handwritten numeral recognition. One is the manual feature extraction approach. It indicates that the raw input is presented to a learning object to discover whatever features are inherent in the domain [7–11]. The other is the automated feature learning approach which is feasible only when there are a large number of samples available for each class [12–15]. Each approach has its own merits and weaknesses. In the former, the main difficulty lies in determining appropriate features, as well as in extracting those features in a robust and reliable way, while the latter is often affected by the number of samples.

Generally, automated feature learning approach mainly includes principal component analysis, Kernel principal component analysis (KPCA) and independent component analysis (ICA).

Feature extraction acts as a vital role for a recognition system. Principal component analysis is a well-known method for feature extraction. By calculating the eigenvectors of the covariance matrix of the original inputs, PCA linearly transforms a high-dimensional input vector into a low-dimensional one whose components are uncorrelated. KPCA is one type of nonlinear PCA developed by generalizing the kernel method into PCA [16–18]. Specially, KPCA firstly maps the original inputs into a high-dimensional feature space using the kernel method and then calculates PCA in the high-dimensional feature space.

Recently, support vector machine (SVM) has become a popular tool for numeral recognition due to its remarkable characteristics such as good generalization performance, the absence of local minima. SVM implements the structural risk minimization principle which seeks to minimize an upper bound of the generalization error rather than minimizing the training error [19].

In this paper, we present a recognition system for handwritten Bangla numerals. The system consists of four main parts: preprocessing, feature extraction, recognition, and integration. The input of the system is the image of a letter and the output is the corresponding recognition result. In this paper, we propose two recognition approaches. One is the image reconstruction recognition approach based on PCA and dissimilarity between the reconstructed image and the original image; the other is the approach of direction feature combined with PCA and SVM. The experimental results show that the performances of the two approaches are superior to other conventional recognition approaches. In order to build a system that is able to operate efficiently in recognition mode and achieve desirable accuracy, an integration scheme which combines two proposed approaches with one conventional approach may be necessary. In order for a recognition system to be acceptable in practice, the response time of the system needs to be considered. Experimental results demonstrate that our system performs very well. It meets the response time as well as the accuracy requirements.

The remainder of the paper is structured as follows: Section 2 presents the preprocessing of the system, which mainly discusses postcode location, numeral characters segmentation and normalization. Section 3 describes the basic theories of PCA and SVM. Sections 4 and 5 propose the recognition approach based on the image reconstruction with PCA and the recognition approach based on Kirsch combined PCA and SVM, respectively. Section 6 presents the experimental results achieved by the proposed approaches. Section 7 discusses integration part of the recognition system. Finally, in Section 8 we draw out conclusions.

2. Preprocessing

The goal of this part is to extract numeral images of postcode from a letter image for the subsequent recognition. Fig. 1 presents two images captured from real Bangladesh letters.

The numerals of postcode are located either in the postcode frame or in the destination address region. We detect the position of postcode and then segment the numeral characters. All of the handwritten numeral samples tested in our experiments are extracted from real Bangladesh letters provided by Bangladesh Post. Fig. 2 gives some examples.

Size normalization is implemented on each character image. The size is normalized to be 16×16 pixels. Then, we employ the smooth filter to remove the noise and then transform the gray image into the binary image. We notice

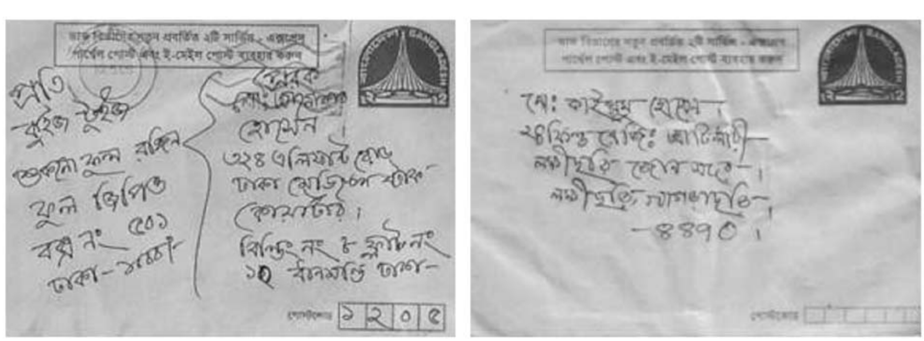

Fig. 1. Images of real Bangladesh letters.

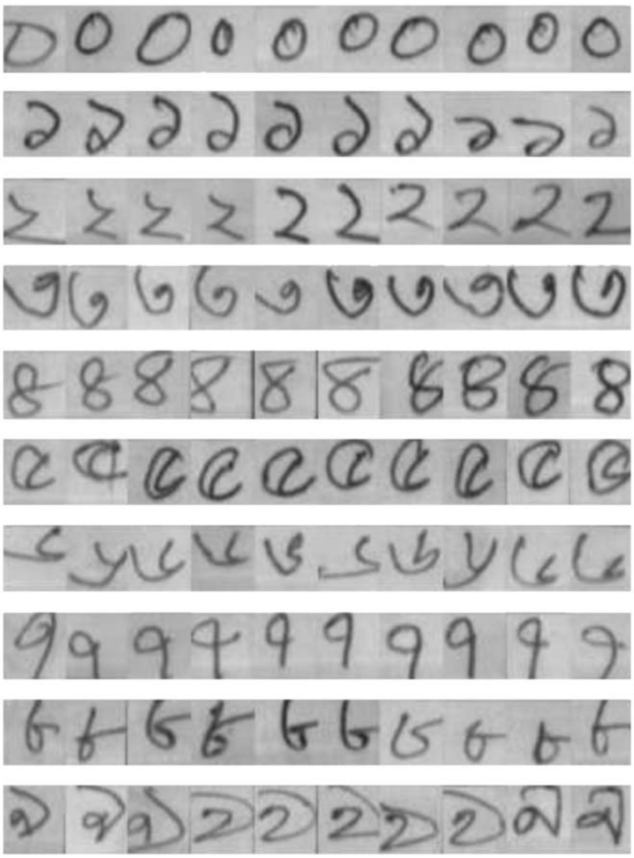

Fig. 2. Samples of handwritten Bangla numeral images (0–9).

Fig. 3. (a) Segmented numeral images; (b) normalized binary images of (a).

that the thickness of the stroke is varying because different people use different writing tools. Since the thickness of the stroke influences the recognition result, we use the thinning and dilatation algorithm to process the normalized numerals and get the same thickness of the stroke (see Fig. 3).

3. Basic theories of PCA and SVM

3.1. Principal component analysis

PCA is a popular technique for feature extraction and dimensionality reduction. Suppose that $X = \{x_i; i = 1, 2, \ldots, N\}$ is a set of centered observations of an m-dimensional zero-mean variable. Let

$$\sum_{i=1}^{N} x_i = 0. \tag{1}$$

The covariance matrix of the variable can be estimated as follows:

$$C = \frac{1}{N} \sum_{i=1}^{N} x_i x_i^{\mathrm{T}}. \tag{2}$$

PCA aims at making the covariance matrix C in Eq. (2) be diagonal. It leads to an eigenvalue problem:

$$\lambda_j \upsilon = C\upsilon_j, \quad j = 1, \ldots, m, \tag{3}$$

where λ_j is one of the eigenvalues of C and υ_j is the corresponding eigenvectors.

PCA linearly transforms each vector x_i into a new one y_i. Based on PCA, the components of y_i are then calculated as the orthogonal transformations of x_i:

$$y_i = \upsilon_j^{\mathrm{T}} x_i, \quad i = 1, \ldots, N; \quad j = 1, \ldots, m. \tag{4}$$

The new components are called principal components. By using only the first several eigenvectors sorted in descending order of the eigenvalues, the number of principal components in y_i can be reduced. So PCA has the dimensional reduction characteristic and the principal components are uncorrelated.

3.2. Support vector machine

The appeal of SVM lies in their strong connection to the underlying statistical learning theory. According to the structural risk minimization principle, a function that can classify training data accurately and which belongs to a set of functions with the lowest capacity (particularly in the VC-dimension) will generalize best, regardless of the dimensionality of the input space. In the case of a canonical hyperplane, minimizing the VC-dimension corresponds to maximizing the margin. As a result, for many applications, SVM have been shown to provide a better generalization performance than conventional techniques.

The training data points can be expressed as $(x1, y1)$, $(x2, y2), \ldots, (xl, yl)(xi \in R^n$; *$yi \in R$ is the target value*), SVM approximates the function using the following form:

$$f(x) = sgn(w \cdot \Phi(x) + b), \tag{5}$$

where $\Phi(x)$ represents a high-dimensional feature space which is nonlinearly mapped from the input space x. The coefficients w and b are estimated by minimizing the regularized risk function [20]. Support vector machine for a pattern recognition problem can be formulated as the quadratic optimization problem [21]:

$$\begin{aligned} &\textit{maximize}: && \sum_{i=1}^{l} \alpha_i - \frac{1}{2}\alpha^{\mathrm{T}} Q \alpha \\ &\textit{subject to}: && 0 \leqslant \alpha_i \leqslant C, \quad i = 1, \ldots, l \\ &&& \sum_{i=1}^{l} y_i \alpha_i = 0, \end{aligned} \tag{6}$$

where α is a vector of length l and its component α_i corresponds to a training sample (x_i, y_i), Q is an $l \times l$ semidefinite kernel matrix, and C is a parameter chosen by the user. A larger C assigns a higher penalty to the training errors. The training vector x_i whose corresponding α_i is nonzero is called support vector. Support vector machine maps training vector x_i into a high-dimensional feature space by the function $\Phi(x)$ and $K(x_i, x_j) = \Phi^{\mathrm{T}}(x_i)\Phi(x_j)$. When the above optimization problem is solved, we can obtain an optimal hyperplane in a high-dimensional feature space to separate the two-class samples. The decision function is given by

$$f(x) = sgn\left(\sum_{i=1}^{l} y_i \alpha_i K(x_i, x) - b\right), \tag{7}$$

where

$$sgn(u) = \begin{cases} 1 & \text{for } u > 0, \\ -1 & \text{for } u < 0. \end{cases} \tag{8}$$

4. Recognition approach of the image reconstruction based on PCA

In the paper, we propose the recognition approach of the image reconstruction based on PCA (we call it IRPCA approach hereafter). The space of handwritten numeral is a typical high-dimension space. The original higher-dimensional inputs space should be transformed into the lower-dimensional feature space by PCA.

For an image of a character, it is represented by an m-dimensional data set $A = \{a_i; i = 1, 2, \ldots, m\}$. For one class training set, the corresponding covariance matrix is C and m eigenvalues can be gotten. We sort the eigenvalues (λ) (and corresponding eigenvectors (υ)) so that $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_m$. We can select the first d eigenvectors $(\upsilon)(i = 1, 2, \ldots, d; d < m)$ and generate the data set in the new (usually compressed) representation to build a corresponding PCA model for every class of numerals. Handwritten Bangla numerals have 10 classes so we can get ten image reconstruction models, respectively. For a test image A, its PCA reconstructed images $\widetilde{A}_j (j = 0, 1, \ldots, 9)$ can be calculated by

$$\widetilde{A}_j = \sum_{i=1}^{d} \upsilon_{ij}^{\mathrm{T}}(A - \overline{A}_j)\upsilon_{ij} + \overline{A}_j, \quad j = 0, 1, \ldots, 9. \tag{9}$$

The corresponding dissimilarities between the reconstructed images (0–9) and the test image are computed by

$$E^{(j)}(A) = \|A - \widetilde{A}_j\|^2 = \left\|\sum_{i=1}^{d} \upsilon_{ij}^{\mathrm{T}}(A - \overline{A}_j)\upsilon_{ij} - (A - \overline{A}_j)\right\|^2, \quad j = 0, 1, \ldots, 9, \tag{10}$$

where $\overline{A}_j$ is the mean value of each class training set and υ_{ij} is the corresponding eigenvector of each class.

We input a test sample image and get the 10 reconstructed images by Eq. (9). According to Eq. (10), 10 dissimilarities can be obtained by the 10 reconstructed images and the test image. If there is a minimum dissimilarity among the ten dissimilarities, it indicates that the class with the minimum dissimilarity is the recognition result.

Given the test sample , if r satisfies:

$$E^{(r)}(A) = \min_{i=0,1,\ldots,9}\{E^{(i)}(A)\}. \tag{11}$$

Then the sample belongs to the r class.

5. Recognition approach based on Kirsch mask, PCA and SVM

Based on PCA and SVM, we propose the recognition approach based on Kirsch mask combined PCA and SVM (we call it KPS approach hereafter). Generally, Kirsch edge detector, Prewitt edge detector, Sobel edge detector, and so on are the representative edge detectors. Among them, the Kirsch edge detector has been known to detect four-directional edges more accurately than others because the Kirsch edge detector considers all eight neighbors [22]. Kirsch defined a nonlinear edge enhancement algorithm as follows [23]:

$$G(i, j) = \max\left\{1, \max_{k=0}^{7}[|5S_k - 3T_k|]\right\}, \tag{12}$$

where

$$S_K = A_K + A_{K+1} + A_{K+2}, \tag{13}$$

$$T_K = A_{K+3} + A_{K+4} + A_{K+5} + A_{K+6} + A_{K+7}. \tag{14}$$

In Eqs. (12)–(14), $G(i, j)$ is the gradient of pixel (i, j), $A_k (k = 0, 1, \ldots, 7)$ is eight neighbors of pixel (i, j) defined as shown in Fig. 4.

In this paper, we calculate directional feature vectors for horizontal (H), vertical (V), right-diagonal (R), and left-diagonal (L) directions as follows:

$$G_H(i, j) = \max(|5S_0 - 3T_0|, |5S_4 - 3T_4|),$$

$$G_V(i, j) = \max(|5S_2 - 3T_2|, |5S_6 - 3T_6|),$$

$$G_R(i, j) = \max(|5S_1 - 3T_1|, |5S_5 - 3T_5|),$$

$$G_L(i, j) = \max(|5S_3 - 3T_3|, |5S_7 - 3T_7|). \tag{15}$$

A_0	A_1	A_2
A_7	(i, j)	A_3
A_6	A_5	A_4

Fig. 4. Definition of eight neighbors of pixel (i, j).

$$\begin{array}{|c|c|c|}\hline 5&5&5\\\hline -3&0&-3\\\hline -3&-3&-3\\\hline\end{array}\quad\begin{array}{|c|c|c|}\hline 5&5&5\\\hline -3&0&-3\\\hline -3&-3&-3\\\hline\end{array}\ \text{horizontal}\qquad\begin{array}{|c|c|c|}\hline 5&5&5\\\hline -3&0&-3\\\hline -3&-3&-3\\\hline\end{array}\quad\begin{array}{|c|c|c|}\hline 5&5&5\\\hline -3&0&-3\\\hline -3&-3&-3\\\hline\end{array}\ \text{vertical}$$

$$\begin{array}{|c|c|c|}\hline 5&5&5\\\hline -3&0&-3\\\hline -3&-3&-3\\\hline\end{array}\quad\begin{array}{|c|c|c|}\hline 5&5&5\\\hline -3&0&-3\\\hline -3&-3&-3\\\hline\end{array}\ \text{diagonal}\backslash\qquad\begin{array}{|c|c|c|}\hline 5&5&5\\\hline -3&0&-3\\\hline -3&-3&-3\\\hline\end{array}\quad\begin{array}{|c|c|c|}\hline 5&5&5\\\hline -3&0&-3\\\hline -3&-3&-3\\\hline\end{array}\ \text{diagonal}/$$

Fig. 5. Kirsch masks.

Fig. 5 shows the Kirsch masks used for calculating directional feature vectors. The normalized image is processed by the pairs of Kirsch masks (3×3 pixels), resulting in four graded feature maps, coding for the presence of horizontal, vertical or diagonal edges. In this way, we get four direction feature maps. We also take the original image as the fifth map in order to compensate the disadvantages of four direction feature maps. The final data representation corresponds to $5 \times 16 \times 16 (=1280)$ dimensional inputs. The 1280-dimension is too much for input of a classifier so we employ PCA to extract feature and decrease the dimension. Finally, we obtain the new feature presentation as the input of SVM.

6. Experimental results

To meet the requirements of a practical usage of the letter sorting machine for Bangladesh Post, we should consider not only the recognition rate and the error rate but also the response time. For our experiments, 16 000 numerals were obtained from the real letters written by Bangladesh. These numerals are acquired by a letter sorting machine. There are various writing styles. We randomly select 6000 numerals as the training set and the rest as the test set.

In this paper, our experiments are tested on Windows 2000, Pentium ¢ô 2.8 G and 512 RAM. We employ Visual C++ to realize the program. To evaluate the recognition results of different approaches, the following measures are employed: recognition rate (Recog.), error rate (Err.), rejection rate (Reject.), and reliability (Reliab.). They are defined as follows:

$$Recog. = \frac{N1(correct\ recognized)}{N4(tested)} \times 100\% \tag{16}$$

$$Err. = \frac{N2(error\ recognized)}{N4(tested)} \times 100\% \tag{17}$$

$$Reject. = \frac{N3(rejected)}{N4(tested)} \times 100\% \tag{18}$$

$$Reliab. = \frac{Recog.}{100\% - Reject.}, \tag{19}$$

where $N1$ (correct recognized) is defined as the number of numerals correctly classified, $N2$ (error recognized) is defined as the number of numerals misclassified, $N3$ (rejected) is defined as the number of numerals rejected by the classifier, and $N4$ (tested) is the number of input numerals for test.

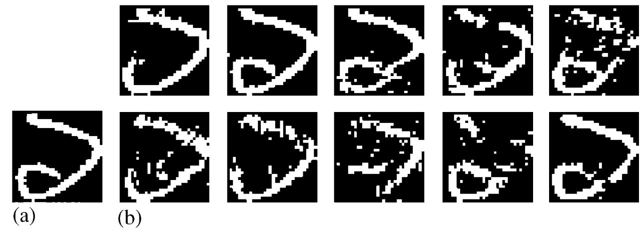

Fig. 6. (a) A Bangla numeral "1" test image; (b) ten numeral "1" reconstructed images.

6.1. Experiment with the IRPCA approach

In this experiment, we obtain the eigenvalues and eigenvectors of every class according to the covariance of the training set. As discussed in Section 4, in order to reconstruct image by the feature space model, we need to get the first d eigenvectors ($d < m$) to construct the feature space model. The parameter d can be obtained as follows:

$$\frac{\sum_{i=1}^{d} \lambda_i}{\sum_{i=1}^{m} \lambda_i} \geqslant \alpha, \tag{20}$$

where α is defined as 0.95 in our experiments, which means the ratio of the energy of the sample in the first d-axis compared with that of the whole energy.

Fig. 6(a) demonstrates a Bangla numeral "1" test sample image and Fig. 6(b) shows 10 numeral "1" reconstructed images obtained by 10 classes PCA models (0–9 classes), respectively. It can be seen from Fig. 6 that the reconstructed image by "1" class PCA model is the most similar with the test image among 10 reconstructed images. We obtain 10 dissimilarities according to Eq. (10). By Eq. (11), the dissimilarity of "1" class compared with others is the minimum one so we recognize the test image as "1".

6.2. Experiment with the KPS approach

In this experiment, we first obtain four direction feature maps of a numeral image by the Kirsch mask as described in Section 5. The size of each map is 16×16 pixels. The original image is considered as the fifth feature map. Then we arrange the five images to get the feature vectors ($5 \times 16 \times 16 = 1280$ dimensions). PCA is also used to decrease the dimension and extract main features. In this way, recognition time and training time can be reduced. Fig. 7 gives the corresponding procedure.

6.3. Performance comparison of different recognition approaches

To evaluate the performance of our proposed approaches, three conventional PCA and SVM approaches are also employed. They are:

(1) Original image without any feature extraction.
(2) PCA features extraction.
(3) KPCA features extraction and polynomial function as kernel function.

We take the above three feature representations as the input of SVM classifier. Table 1 shows the training and recognition time of different approaches. Table 2 presents the recognition results of five approaches.

Original+SVM is the only approach without PCA process, and IRPCA is the only approach without SVM. Tables 1 and 2 show that the recognition rate and time of Original+SVM is not good. KPCA+SVM is better than PCA+SVM in the recognition rate but not in time. Apparently, the time spent in KPCA+SVM is much larger than those of PCA+SVM and Original+SVM. The reason lies in that the number of training data points is much larger than the input dimension in all the data sets studied; therefore, KPCA needs to estimate a larger number of principal components than others. In three conventional approaches, the recognition rates of Original+SVM, PCA+SVM, KPCA+SVM are similar but that of PCA+SVM is superior to others in the recognition time. IRPCA is better than Original+SVM, PCA+SVM and KPCA+SVM at the recognition rate and time. According to these five approaches, the recognition and training time of KPS may not be better than others, but the recognition rate is the best. The reason lies in the feature extraction. In order to maximize the performance of numeral recognition, KPS extracts four directional local feature vectors with the Kirsch masks and one global feature vector from a normalized input image. Hence, the important features are retained and the nonimportant pixels are minimized. The recognition result of KPS outperforms the others. However, KPS has extracted more principal components than PCA+SVM and Original+SVM, so it needs much more time to estimate than the other two approaches. As far as the recognition rate concerned, it can be seen that the two recognition approaches we proposed in this paper are superior to other conventional approaches.

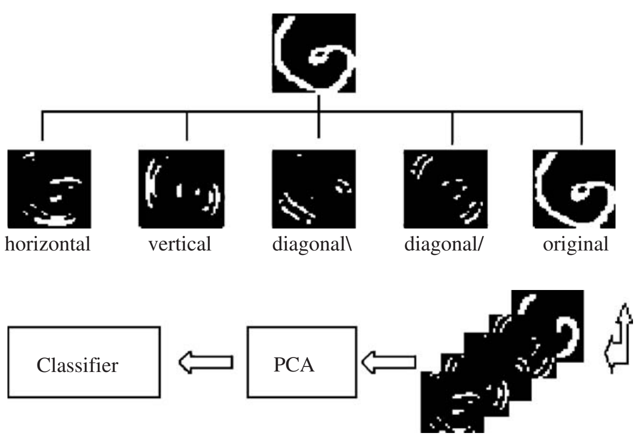

Fig. 7. Recognition process based on the KPS approach.

Table 1
The recognition and training time of five approaches

Approach	Recognition time (ms)	Training time (s)
Original+SVM	21.44	67.94
PCA+SVM	12.50	39.22
KPCA+SVM	398.45	1516.24
IRPCA	9.15	0
KPS	31.24	146.47

7. Integration of recognition system

Our work aims at applying the research to the letter sorting machine in practice. In general, an automatic letter sorting machine is capable of processing more than 10 letters per second. Actually, due to the highly practical demand of the sorting machine for the handwritten Bangla numeral recognition, two important evaluation factors, namely, the recognition reliability and the response time should be considered. To take into account these requirements, we employ an integrated system combining three different approaches.

Based on the aforementioned comparison, we choose three approaches with higher recognition rate and less recognition time, namely, PCA+SVM, IRPCA, KPS to form a practical system. The integration strategy applied in our system is majority voting. In other words, if at least two out of three recognition results are identical,

Table 2
The recognition rates (%) of five approaches

Approach/numerals	0	1	2	3	4	5	6	7	8	9	Average
Original+SVM	96.31	91.02	66.25	82.33	93.23	90.04	76.45	92.56	89.27	83.24	86.07
PCA+SVM	97.01	91.12	70.38	82.59	93.33	93.14	77.23	91.27	90.67	83.37	87.01
KPCA+SVM	97.08	92.11	71.25	81.85	93.33	92.84	76.37	91.36	90.66	84.10	87.10
IRPCA	93.55	89.17	90.53	86.27	87.70	90.08	94.20	98.79	98.02	94.44	92.27
KPS	99.06	92.58	93.18	91.95	97.88	97.25	93.14	99.51	98.42	90.64	95.36

the results are accepted as the final one. If three recognition results are different, the final result will be rejection. In this way, the correct rate can be improved and the error rate can be lowered.

Since three approaches are applied in the system, the response time of the system increases undoubtedly. Although parallel hardware will be used in the practical system, the response time for processing one letter is required to be less than 0.5 s for our testing environments.

Generally speaking, the letter preprocessing is a necessary work including postcode location, de-noising and segmentation. The postcode of a Bangladesh letter is composed of four or six numerals.The experimental result shows that the average preprocessing time of a letter is about 0.05 s. Table 3 shows that the average system response time with different approaches for a Bangladesh letter recognition (the average system response time contains the preprocessing time).

According to the testing of the system which has been integrated with PCA+SVM, IRPCA, KPS, the average response time of sorting a Bangladesh letter is about 0.36 s, and the reliability is 99.03%. However, if the system is applied with only PCA+SVM, or IRPCA, or KPS, the average response time of sorting a letter is about 0.125, 0.105, 0.23 s, and the reliability is 87.01%, 92.27%, 95.36%, respectively. Although the average response time of the system integrated with three approaches is longer than that with one approach alone, the reliability is improved significantly. In the application of the sorting machine, generally, the improvement of reliability is very important. Therefore, the integrated system is much more practical and valuable.

Table 4 shows the recognition rate, the error rate and the reliability of different numerals for the integrated system. Finally, we get the average recognition rate, the average error rate and the average reliability are 95.05%, 0.93% and 99.03%, respectively. Experimental results demonstrate that our presented system performs very well. It meets the response time as well as the accuracy requirements. According to the above description, our recognition system for handwritten Bangla numerals is depicted in Fig. 8.

Table 4 shows the recognition performance for different numerals. It has been found that the highest reliability is 99.54% obtained by Bangla numeral ‘seven’. The mistaken recognition or rejection is due to variability of handwritings as well as bad writing. From the experiments we note that the most confusing numeral pair is “zero” and “three” (as shown in Fig. 9(a). Second confusing pair is “one” and “nine”

Table 3
The average system response time with different approaches for a Bangladeshi letter recognition

PCA+SVM (s)	IRPCA (s)	KPS (s)	Integration of three approaches (s)
0.125	0.105	0.23	0.36

Table 4
The final recognition results by the integrated system

Recognition results	0	1	2	3	4	5	6	7	8	9	Average
Recog.(%)	98.87	94.41	92.56	91.25	95.67	95.25	93.02	98.36	97.89	93.25	95.05
Err.(%)	0.48	0.98	1.56	1.36	0.87	0.86	1.02	0.45	0.65	1.03	0.93
Reliab.(%)	99.52	98.98	98.34	98.53	99.10	99.11	98.92	99.54	99.34	98.91	99.03

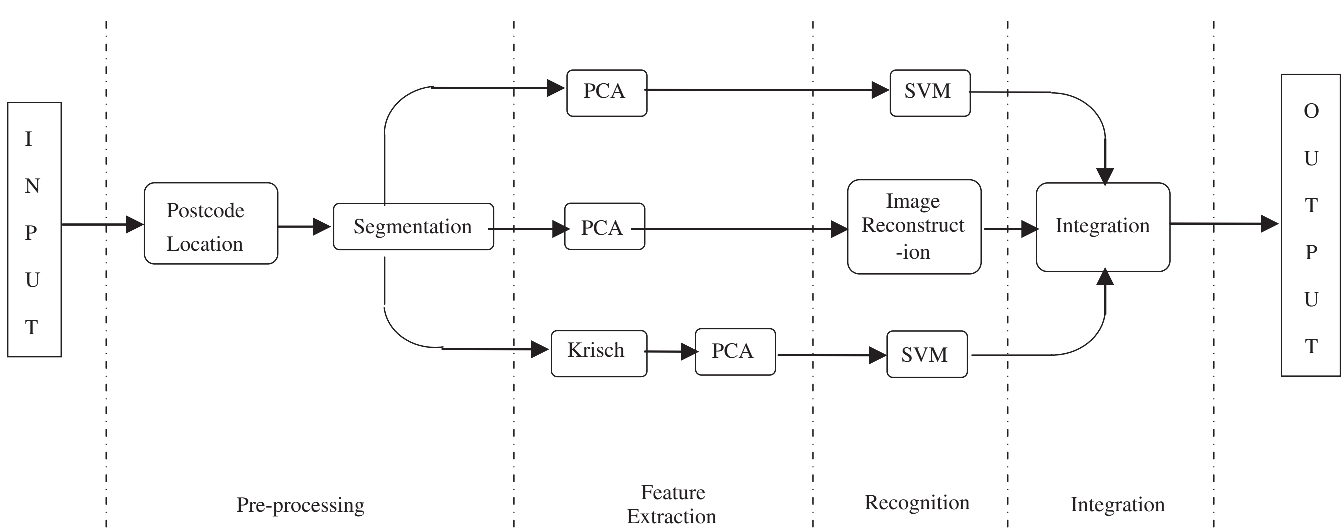

Fig. 8. Block diagram of Bangladeshi letter recognition system.

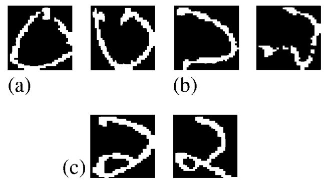

Fig. 9. Examples of some confused handwritten Bangla numeral pairs: (a) “0” and “3”; (b) “1” and “9”; (c) “1” and “2”.

(as shown in Fig. 9(b). Third confusing pair is “one” and “two” (as shown in Fig. 9(c). Therefore, error rate or rejection rare for these numerals is higher.

8. Conclusion

An efficient recognition system for handwritten Bangla numerals has been developed. In the proposed system, we first discuss two recognition algorithms based on usage of PCA, and experimental results confirm the effectiveness of the proposed approaches. Then we employ the integration strategy in the system for improve recognition performance. The recognition result achieved by the integrated system is more reliable than that by one method alone. Experiments also demonstrate that the response time of the integrated system is acceptable.

It is obvious that integration of different approaches will improve the performance of the whole recognition system. In our future work, we will include more recognition approaches in our integrated system for achieving better recognition ability in the case that the response time is acceptable.

Acknowledgments

This work is jointly supported by the contract project of Bangladesh Post no. PARI-2/2-16/2003/2004, and the National Natural Science Foundation of China under Grant no. 60475006, and Program for New Century Excellent Talents in University (NCET-05-0430). The authors would like to thank Bangladesh Post for preparing the envelop images for us.

References

[1] B.B. Chaudhuri, U. Pal, A complete printed Bangla OCR system, Pattern Recognition 31 (1998) 531–549.
[2] U. Pal, On the development of an optical character recognition (OCR) system for printed Bangla script, Ph.D. Thesis, 1997.
[3] U. Pal, B.B. Chaudhuri, Indian script character recognition: a survey, Pattern Recognition 37 (2004) 1887–1899.
[4] U. Pal, B.B. Chaudhuri, Automatic recognition of unconstrained off-line Bangla hand-written numerals, in: T. Tan, Y. Shi, W. Gao (Eds.), Advances in Multimodal Interfaces, Lecture Notes in Computer Science, vol. 1948, Springer, Berlin, 2000, pp. 371–378.
[5] K. Roy, S. Vajda, U. Pal, B.B. Chaudhuri, A system towards Indian postal automation, in: Proceedings of the Ninth International Workshop on Frontiers in Handwritten Recognition (IWFHR-9 2004), October 2004, pp. 580–585.
[6] U. Pal, A. Belad, Ch. Choisy, Touching numeral segmentation using water reservoir concept, Pattern Recognition Lett. 24 (2003) 261–272.
[7] I.-J. Kim, J.-H. Kinm, Statistical character structure modeling and its application to handwritten Chinese character recognition, IEEE Trans. Pattern Anal. Mach. Intell. 25 (11) (2003) 1422–1436.
[8] P.-K. Wong, C. Chan, Off-line handwritten Chinese character recognition as a compound Bayes decision problem, IEEE Trans. Pattern Anal. Mach. Intell. 20 (9) (1998) 1016–1023.
[9] K.-W. Cheung, D.-Y. Yeung, R.T. Chin, A Bayesian framework for deformable pattern recognition with application to handwritten character recognition, IEEE Trans. Pattern Anal. Mach. Intell. 20 (12) (1998) 1382–1388.
[10] V. Govindaraju, P. Slavik, H. Xue, Use of Lexicon density in evaluating word recognizers, IEEE Trans. Pattern Anal. Mach. Intell. 24 (6) (2002) 789–800.
[11] S. Madhvanath, G. Kim, V. Govindaraju, Chaincode contour processing for handwritten word recognition, IEEE Trans. Pattern Anal. Mach. Intell. 21 (9) (1999) 928–932.
[12] C.J. Twining, C.J. Taylor, The use of kernel principal component analysis to model data distributions, Pattern Recognition 36 (2003) 217–227.
[13] H.-C. Kim, D. Kim, S.Y. Bang, A numeral character recognition using the PCA mixture model, Pattern Recognition Lett. 23 (2002) 103–111.
[14] M. Partridge, R.A. Calvo, Fast dimensionality reduction and simple PCA, Intell. Data Anal. 2 (1998) 203–214.
[15] J. Yang, D. Zhang, Frangi, et al., Two-dimensional PCA: a new approach to appearance-based face representation and recognition, IEEE Trans. Pattern Anal. Mach. Intell. 26 (1) (2004) 131–137.
[16] B. Scholkopf, A. Smola, K.R. Muller, Nonlinear component analysis as a kernel eigenvalue problem, Neural Comput. 10 (1998) 1299–1319.
[17] Z. Liang, P.F. Shi, Uncorrelated discriminant vector using a kernel method, Pattern Recognition 38 (2005) 307–310.
[18] Z. Liang, P.F. Shi, Kernel direct discriminant analysis and its theoretical foundation, Pattern Recognition 38 (2005) 445–447.
[19] V.N. Vapnik, The Nature of Statistical Learning Theory, Springer, New York, 1995.
[20] I.J. JolliHe, Principal Component Analysis, Springer, New York, 1986.
[21] J.-X. Dong, A. Krzyzak, C.Y. Suen, Fast SVM training algorithm with decomposition on very large data sets, IEEE Trans. Pattern Anal. Mach. Intell. 27 (4) (2005) 603–618.
[22] S.W. Lee, Off-line recognition of totally unconstrained handwritten numerals using multilayer cluster neural network, IEEE Trans. Pattern Anal. Mach. Intell. 18 (6) (1996) 648–652.
[23] W.K. Pratt, Digital Image Processing, Wiley, New York, 1978.

About the Author—YING WEN received her B.S. degree from Hefei University of Technology in 1997, and M.S. degree from Shanghai University, China in 2002. From 2002, she is an engineer at the Shanghai Research Institute of China State Post Bureau. Now, she is Ph.D. candidate in image processing and pattern recognition at Shanghai Jiao Tong University, Shanghai, China. Her research interests include computer vision, digital image processing, pattern recognition and artificial intelligence.

About the Author—YUE LU received the B.E. and M.E. degrees in telecommunications and electronic engineering from Zhejiang University, Zhejiang, China, in 1990 and 1993, respectively, and his Ph.D. degree in pattern recognition and intelligence system from Shanghai Jiao Tong University, Shanghai, China, in 2000. He is a professor and head in the Department of Computer Science and Technology, East China Normal University. From 2000 to 2004,

he was a research fellow with the Department of Computer Science, National University of Singapore. From 1993 to 2000, he was an engineer at the Shanghai Research Institute of China State Post Bureau. Because of his outstanding contributions to the research and development of the automatic mail sorting system OVCS, he won the Science and Technology Progress Awards issued by the Ministry of Posts and Telecommunications in 1995, 1997, 1998, respectively. In 2005, he was honored New Century Excellent Talent in University and Shanghai Shuguang Scholar. His research interests include document image processing, document retrieval, image analysis, character recognition, and intelligence system.

About the Author—PENGFEI SHI is a professor in the Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, Shanghai, China. His research interests include computer vision, digital image processing, pattern recognition and artificial intelligence.