Int. Computer Symposium, Dec. 15-17, 2004, Taipei, Taiwan.

ONLINE RECOGNITION OF CHINESE HANDWRITTEN CHARACTERS BASED ON THE POINT DISTRIBUTION MODEL

Ming-Yen Tsai and *Leu-Shing Lan*[*]

Department of Electronics Engineering
National Yunlin University of Science and Technology, Taiwan

ABSTRACT

High variability is the main difficulty in online handwriting recognition. A wise mechanism should be devised to alleviate this problem. Among others, character deformation is one particular promising approach. Through shape deformation, some of the variations in handwriting is absorbed. In this paper, we present a point distribution model (PDM)-based method for online recognition of handwritten Chinese characters. The PDM technique is a powerful statistical tool that learns from a group of sample shapes to extract their mean shape and analyze their principal variation modes. These variation modes describe the main ways in which the sample shapes tend to deform from their mean shape. In applying the PDM technique to online Chinese character recognition, we treat deformations within the range of two times of the standard deviation of each variation mode as acceptable. For each test character, the variation parameters are strictly limited to be within this range by utilizing projection. The difference of the variation vectors is then used for character classification. We have conducted some experiments to demonstrate the applicability of the proposed approach. The experimental results indicate that the proposed PDM-based approach is superior compared to method 1 in [9] and the SAT approach in [6]. The main reason for better performance should be attributed to the structural deformation capability that the PDM technique owns.

1. INTRODUCTION

Online character recognition is a very important technique for portable computing and communication devices. Due to the variability of handwritten characters and differences among the writing styles of various writers, online handwritten Chinese character recognition manifests itself as a difficult task. To deal with handwriting variation, efficient distortion-tolerant recognition methods should be devised. Some earlier methods include dynamic programming [1], nonlinear shape normalization [2, 3], biomedical handwriting models [4], local affine transformation (LAT) [5], stroke-based affine transformation (SAT) [6], etc. In this paper, we present a point distribution model (PDM)-based approach [7, 8] for the recognition of online handwritten Chinese characters. The PDM is a statistical technique that learns from a group of sample shapes to extract their mean shape and analyze their principal variation modes. By placing limits on the shape parameters, realistic and plausible shapes satisfying certain physical constraints can be synthesized. To apply the PDM technique, first we need to find the covariance matrix **C** from a group of training characters. Performing eigenalalysis on **C** yields the eigenmatrix **P**. Each column **P** represents a certain variation mode. In character recognition, we adopt the viewpoint that certain amounts of shape variations are acceptable due to instability in handwriting. This instability (or variations) in handwriting is reflected through the use of Eqs. (6) and (7) given in Section 2. Under our formulation, the valid variation of each characteristic mode is confined to be within $\pm 2\sigma_k$, where σ_k is the standard deviation. Within this range, the deformed characters are all considered to be variations of the same character and will then be recognized as the same character. A recognition algorithm using this principle is proposed. Given an input test character, it is first analyzed through a principal component analysis (PCA)-like procedure. If the magnitudes of variation parameters are above the predefined thresholds, i.e., $2\sigma_k$, they are projected back to the boundaries (namely, $2\sigma_k$). The deformed templates are then matched against the test character. The recognition result comes from the minimum distance decision. Some experiments were conducted to demonstrate the applicability of the proposed approach. The experimental results reveal a definite superiority of the proposed PDM-based approach as compared to method 1 in [9] and the SAT approach in [6]. The main reason for better performance should be attributed to the structural deformation capability that the PDM technique exhibits.

The rest of this paper is organized as follows. In Section 2, the PDM technique is described in detail. In

*This work was supported by the National Science Council of Taiwan, R.O.C., under the grant no. NSC93-2213-E-224-012.
*email of the corresponding author: lanls@yuntech.edu.tw

Section 3, application of the PDM method to the recognition of online Chinese handwriting is described. Detailed experimental results are shown in Section 4, which is followed by the conclusions section.

2. THE POINT DISTRIBUTION MODEL

The point distribution model (PDM) [7] is a statistical tool that can be used to describe characteristics of shapes with deforming capabilities. The PDM approach assumes that a set of training sample shapes are available from which the statistical description of this class of shapes is determined. In principle, landmark points are sampled around the boundaries of shapes. Variations in the positions of these landmark points would then be attributable to natural variations among each shape. The PDM technique provides a means to describe the variations in shapes in a parametric form. The PDM approach comprises two main steps: aligning of training data, and principal component analysis (PCA). The alignment step is necessary to remove the shape variations due to affine transformation (translation, rotation, and scaling). The PCA step is commonly used in signal processing to reduce the size of a feature set, whereas here it is employed for the analysis of characteristic modes in shape variations.

2.1 Alignment of the Training Shapes

The alignment step is briefly summarized as follows [7]:

(1) In a pairwise fashion, align each shape with the first shape, denoted as $\mathbf{x_1}$. The alignment of two shapes are based on the optimal affine transformation regarding translation, rotation, and scaling, which can be formally expressed as

$$\mathbf{x}_{\text{aligned}} = \mathbf{A}\mathbf{x}_{\text{old}} + \gamma \tag{1}$$

where $\mathbf{A}$ is the optimal transformation matrix taking care of rotation and scaling, and $\boldsymbol{\gamma}$ represents the optimal translation vector.

(2) Calculate the mean of the transformed shapes.

(3) Align the mean shape with the first shape $\mathbf{x_1}$.

(4) Align the other shapes (except the first shape) to match to the adjusted mean shape.

(5) If the mean shape has not converged, go to step 2; otherwise, stop.

2.2 Principal Component Analysis [7]

Let the landmarks of shape k be denoted by a vector

$$\mathbf{x_k} \triangleq (X_1^{(k)}, Y_1^{(k)}, X_2^{(k)}, Y_2^{(k)}, \cdots, X_N^{(k)}, Y_N^{(k)})^T \tag{2}$$

and let the mean shape be denoted by $\mathbf{x}_{\text{mean}}$. Defining $\Delta\mathbf{x}_k = \mathbf{x}_k - \mathbf{x}_{\text{mean}}$, the covariance matrix of the training shapes can be estimated by

$$\mathbf{C} = \frac{1}{M}\sum_{k=1}^{M} \Delta\mathbf{x_k}\Delta\mathbf{x_k}^T \tag{3}$$

where $\mathbf{C}$ is a 2N x 2N matrix. From the eigen-decomposition of $\mathbf{C}$, we obtain its eigenvectors $\mathbf{p_k}$ and eigenvalues λ_k. Denoting the orthogonal eigenmatrix by $\mathbf{P} = (\mathbf{p_1}\ \mathbf{p_2}\ \mathbf{p_3} \ldots \mathbf{p_{2N}})$, then any sample shape in the training set can be expressed as

$$\mathbf{x} = \mathbf{x}_{\text{mean}} + \mathbf{Pb} \tag{4}$$

where $\mathbf{p_k}$ represents the kth mode in shape variation and the components of $\mathbf{b}$ indicate the corresponding strength of each mode. In fact, the set

$$\{\mathbf{x} \mid \mathbf{x} = \mathbf{x}_{\text{mean}} + \mathbf{Pb}\} \tag{5}$$

spans a linear space of characters with the basis consisting of the columns of $\mathbf{P}$. It is known that the variance of $\mathbf{b_k}$ can be shown to be equal to the associated eigenvalues λ_k. If the values of $\mathbf{b}$ are restricted to be within a certain range to obtain a new $\tilde{\mathbf{b}}$, Eq. (5) then has the form

$$\mathbf{x}_{\text{def}} = \mathbf{x}_{\text{mean}} + \mathbf{P}\tilde{\mathbf{b}} \tag{6}$$

where $\mathbf{x}_{\text{def}}$ may be viewed as deformed shapes that are generated through the PDM approach. The restriction on $\mathbf{b}$ should be based on physical constraints of the shapes under consideration. A simplified scheme is to let $\tilde{\mathbf{b}}$ satisfy

$$-\beta_k\sqrt{\lambda_k} \le \tilde{b}_k \le \beta_k\sqrt{\lambda_k} \tag{7}$$

where β_k is empirically determined and $\sqrt{\lambda_k}$ corresponds to the standard deviation of b_k.

3. ONLINE RECOGNITION OF HANDWRITTEN CHINESE CHARACTERS USING THE PDM

Typically a Chinese character is composed of a certain number of strokes where each stroke in online writing is a pentip trajectory between the pen-down and pen-up statuses. The pentip trajectory is recorded as a sequence of spatial points that are uniformly sampled in time through an input device such as a stylus pen or a mouse. In a general online character recognition system, features are extracted from the input writing and compared to the standard templates stored in the character database. Major difficulties of recognition lie in the high variations in character shapes as well as varying writing styles among different writers. To resolve the character variability problem, a promising family of approaches is to apply deformable models to absorb shape and structure variations in online handwriting so that only the invariant information is exploited in pattern matching. The stroke-based affine transformation (SAT) [6] is one such approach. In this work, we follow the streamline of character deformation; however, we adopt a different character deformation technique that is based on the PDM. The PDM is a generic statistical facility utilizing shape aligning and principal component analysis for shape representation, generation, and detection. Under the formulation of principal component analysis, the variations in shape and structure are represented through the combinations of varying characteristic modes where each mode is characterized by an eigenvector. The characteristic variation mode thus obtained can then produce shape deformation under certain inherent structural constraints. Usually the PDM is used to represent closed contours; however, it can also be used to model online handwriting trajectories as well. For the purpose of shape deformation, the landmarks (or feature points) on each stroke are resampled uniformly in the spatial domain. For simplicity, each stroke has the same number of landmark points. The landmarks of these strokes are then concatenated to produce a long vector that has the same meaning of *shape vector* used in the conventional PDM approach. Eqs. (6) and (7) described in the preceding section provide exactly the tool for our character deformation. To apply the PDM-based deformation model, first we need to find the eigenmatrix **P** in Eq. (6). Since each Chinese character represents a shape class, it is necessary to collect enough handwritten samples for each character category so that variation information can be effectively recorded in the statistical model. After collecting the training samples for a specific character category, Eq. (3) is used to estimate the covariance matrix **C**. Performing eigenalalysis on **C** gives the eigenmatrix **P**. In character recognition, we adopt the viewpoint that certain amounts of shape deformations are possible due to instability in handwriting. This instability (or variations) in handwriting is reflected through the use of Eqs. (6) and (7). The allowable range of b_k is Eq. (7) is empirically determined to be $2\sqrt{\lambda_k}$, i.e., $\beta_k = 2$ for all k in Eq. (7). Therefore, under our formulation, the valid variation of each characteristic mode is confined to be within $\pm 2\sigma_k$, where σ_k is the standard deviation. Within this range, the deformed characters are all considered to be variations of the same character and will then be recognized as the same character. To summarize, the proposed PDM-based online Chinese handwriting recognition scheme is given as follows:

(1) Record the pen trajectory of the test character. Denote the uniformly resampled point sequence by **y**.

(2) For each character category, find the deformation vector $\mathbf{b_1}$ by

$$\mathbf{b_1} = \mathbf{P}^T(\mathbf{y} - \mathbf{x}_{\text{mean}}) \tag{8}$$

(3) Project $\mathbf{b_1}$ onto the constraint sets to obtain $\tilde{\mathbf{b}}$ so that each component of $\tilde{\mathbf{b}}$ satisfies Eq. (7) with β_k=2. By this projection, the components of $\tilde{\mathbf{b}}$ larger than 2 and less than –2 will be set to be exactly equal to 2 and –2, respectively.

(4) Generate the deformed character using Eq. (6).

(5) Calculate the inter-character distance by

$$\| \mathbf{y} - \mathbf{x}_{\text{def}} \|_2^2 = \| \mathbf{P}(\mathbf{b_1} - \tilde{\mathbf{b}}) \|_2^2 = \| (\mathbf{b_1} - \tilde{\mathbf{b}}) \|_2^2 = \| \Delta\mathbf{b} \|_2^2 \tag{9}$$

(6) Determine the character class by

$$\arg\min_k \{ \| \mathbf{y} - \mathbf{x}_{\text{def}} \|_2^2 \}_k \tag{10}$$

4. EXPERIMENTAL RESULTS

Normally, a two-step procedure is used in online Chinese character recognition, including coarse classification and fine classification. The main reason for this two-step design is mainly due to the concern of recognition speed. In the first step, the coarse classifier generates a small set of possible candidates. Then in the second step, the fine classifier gives the conclusive class recognition result among the candidates. Since the candidates may in certain sense resemble one another, the recognition in this stage will encounter more difficulties. One method to improve the recognition accuracy is then to apply a deformation model to remove possible unnecessary variations in shape and structure of the test character due to handwriting instability. Although the deformation concept seems to be directly applicable to the overall character recognizing scheme, it is seldom employed this way because of the expensive computational cost that is needed to perform deformations on all character categories. Following this reasoning, the proposed PDM-based character recognition method is also applied in the fine classification stage only.

4.1 Training Set and Mean Shape

In Figure 1, we show the mean shapes of two characters with their corresponding training samples overlaid. Each training set is composed of 30 character samples, with various forms of shape variations. The two different characters in Figures 1(a) and 1(b) represent *see* and *but* in Chinese, respectively. The training samples were first aligned using the procedure described in Section 2.1, where each stroke of the characters has 8 landmark points and both characters have 7 strokes. It is seen that the sample points of the training sets are scattered around those of the mean shapes, with different area of scattering neighborhood.

4.2 Deformed Templates vs. Test Character

As described earlier, the proposed PDM-based character recognition scheme should be used in the fine classification stage. Suppose there are 10 characters in the candidate set. To recognize the input test character, the feature vector of the input test character is matched against those of all the deformed templates that are generated from the ten candidates using Eq. (6). Figure 2 shows the pattern matching results. In Figure 2(a) through 2(c), the input test characters are “但”, “位”, and “見”, respectively. In all the three cases, the same candidate set that is composed of “伸”, “但”, “佐”, “住”, “佗”, “伺”, “位”, “佇”, “貝”, and “見” is used. From all these three cases, it is clearly seen that the deformed templates attempt to change their shapes to approximate the input test character. In Figure 2(a), after pattern deformation, the deformed mean shape of "但" matches the input character "但" very closely, while the other nine deformed mean shapes are still *far away* from the input character. Thus the deformation tool exhibits its capability to decrease the inter-character distance of the correct character class, while at the same time effectively increases the inter-character distances of those incorrect character classes. These observations can also be found in Figures 2(b) and 2(c).

4.3 Comparison of Recognition Performance

To demonstrate the improvement of recognition performance that might be gained using the PDM-based approach, we compare the behavior of inter-character distance among three different schemes. Scheme 1 is the method 1 described in [9], which is a simple template matching method without exploiting any shape deformation technique. Scheme 2 is the SAT approach proposed in [6], where two subsequent strokes are used as a unit subject to an affine transformation. Scheme 3 is the proposed PDM-based approach. In the comparison, we define the normalized distance μ as the distance ratio

$$\mu \triangleq \frac{\text{actual inter-character distance}}{\min\{\text{all inter-character distances in the candidate set}\}} \quad (11)$$

With larger values of μ, the incorrect character classes will be further away from the input test character, thereby allowing easier recognition to yield the correct result. In Tables 1 through 3, we show the results of these three different schemes where the input test characters are "但", "位", and "見", respectively. Tables 1 through 3 list the results of Scheme1, Scheme 2, and Scheme 3, respectively. For the purpose of easy comparison, we inspect the second smallest value of μ in each case, denoted as μ_2. With a smaller value of μ_2, there are more chances that the recognition result will be incorrect in a probabilistic sense. For the input test character "但", the μ_2 values for the three schemes are 2.99, 3.02, and 3.87, respectively. For the input test character "位", the μ_2 values for the three schemes are 1.87, 1.90, and 1.95, respectively. For the input test character "見", the μ_2 values for the three schemes are 1.36, 1.35, and 1.39, respectively. These results seem to indicate the superiority of the PDM-based scheme. A physical explanation is in order. Scheme 1 does not provide any facility to absorb character variations, while Scheme 2 does perform shape deformation; however, the deformation does not fully exploit the inherent properties inside the character structure.

5. CONCLUSIONS

The main difficulty in handwritten character recognition lies in the high variability or instability in handwriting. A wise mechanism should be devised to alleviate this problem. One recent popular approach in shape detection is to use a deformable model. Through shape deformation, a flexible matching can be performed to detect the shape under investigation. Owing to the wide-accepted popularity of the deformation techniques, a variety of deformable models have been proposed. The point distribution model (PDM) or the related active shape model (ASM) can be viewed as one category of these deformable models. In this paper, we present a PDM-based method for online recognition of handwritten Chinese characters. The PDM technique was originally a useful statistical tool for image contour description and detection. It can learn from a group of sample shapes to extract their mean shape and analyze their principal variation modes. These variation modes describe the main ways in which the sample shapes tend to deform from their mean shape. Although not closed contours, handwritten Chinese characters can also be modeled by the PDM technique. Strokes of a handwritten Chinese character are represented by sequences of landmark points, which are concatenated together to form a long vector. Samples of handwritten characters of the same category are then used to train the PDM. In applying the PDM technique to online Chinese character recognition, we adopt the viewpoint that certain amounts of shape deformations are possible due to instability in handwriting. This instability in handwriting is reflected through the use of Eqs. (6) and (7). The allowable range of b_k is Eq. (7) is empirically determined to be $2\sqrt{\lambda_k}$. Within this range, the deformed characters are all considered to be variations of the same character and will then be recognized as the same character.

We have conducted some experiments to demonstrate the applicability of the proposed approach. From the experimental results, it is seen that definite superiority of the proposed PDM-based approach is observed compared to method 1 in [9] and the SAT approach in [6]. The performance merit was illustrated through inspection of the normalized inter-character distances. The primary reason for better performance is very likely due to the structural deformation capability that the PDM technique owns.

6. REFERENCES

[1] K. Yoshida and H. Sakoe, "online Handwritten Character Recognition for a Personal Computer System," *IEEE Tr. Consumer Electronics*, Vol. 28, No. 3, pp. 202-208, 1982.

[2] H. Yamada, K. Yamamoto, and T. Saito, "A Nonlinear Normalization Method for Handprinted Kanji Character Recognition—Line Density Equalization," *Pattern Recognition*, Vol. 23, No. 9, pp. 1023-1029, 1990.

[3] S.-W. Lee and J.-S. Park, "Nonlinear Shape Normalization Methods for the Recognition of Large-Set Handwritten Characters," *Pattern Recognition*, Vol. 27, No. 7, pp. 895-902, 1994.

[4] R. Plamondon and F. J. Maarse, "An Evaluation of Motor Models of Handwriting," *IEEE Tr. Systems, Man, and Cybernetics*, Vol. 19, No. 5, pp. 1060-1072, 1989.

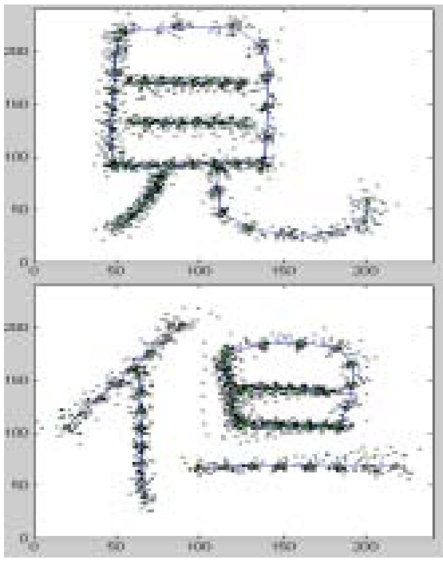
[5] T. Wakahara, "Shape Matching Using LAT and Its Application to Handwritten Numeral Recognition," *IEEE Tr. Pattern Analysis and Machine Intelligence*, Vol. 16, No. 6, pp. 618-629, June 1994.

[6] T. Wakahara and K. Odaka, "On-Line Cursive Kanji Character Recognition Using Stroke-Based Affine Transformation," *IEEE Tr. Pattern Analysis and Machine Intelligence*, Vol. 19, No. 12, pp. 1381-1385, Dec. 1997.

[7] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Training Models of Shape from Sets of Examples," *Proceedings of the British Machine Vision Conference*, pp. 9-18, 1992.

[8] T. F. Cootes, A. Hill, C. J. Taylor, and J. Haslam, "The Use of Active Shape Models for Locating Structures in Medical Images," *Image and Vision Computing*, Vol. 12, No. 6, pp.355-366, July 1994.

[9] T. Wakahara, H. Murase, and K. Odaka, "On-Line Handwriting Recognition," *Proc. IEEE*, Vol. 80, No. 7, pp. 1181-1194, July 1992.



(a)

(b)

Figure 1. Scatter-points plot of the aligned landmarks, with the mean character shape overlaid. In both figures, each stroke has 8 landmark points.

Table 1. Normalized distance (μ) for Scheme 1 using method 1 in [1]. The top row shows the members in the candidate set. The leftmost column shows the input test characters to be recognized.

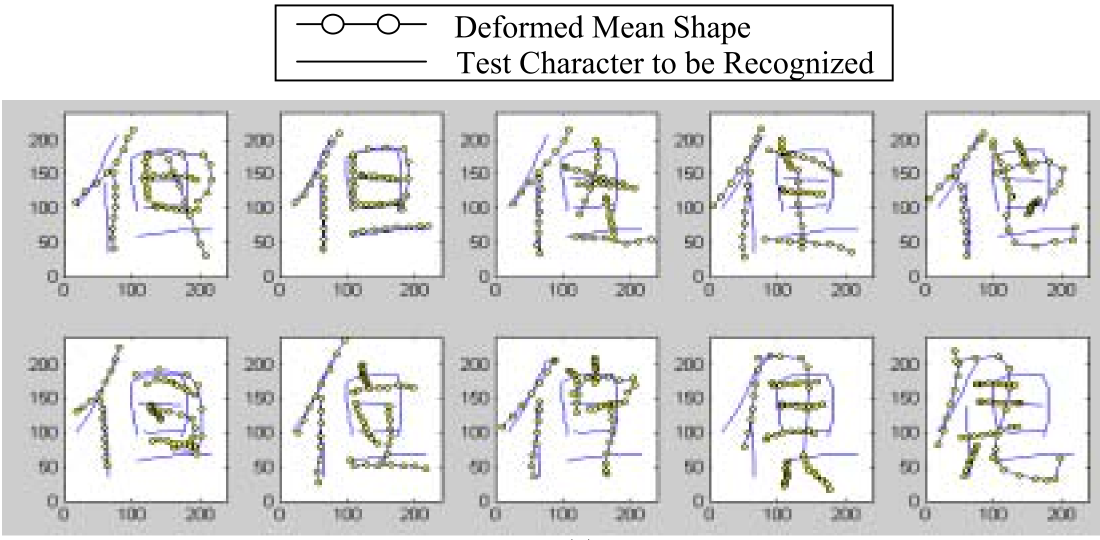
	伸	但	佐	住	佗	伺	位	佇	貝	見
但	2.99	1	3.37	3.52	3.85	4	3.52	4.66	6.19	6.42
位	3.62	2.63	2.1	1.87	2.57	2.98	1	3.39	4	4.72
見	4.71	3.93	3.94	3.36	4.05	4.52	3.63	4.8	1.36	1

Table 2. Normalized distance (μ) for Scheme 2 using the SAT approach. The top row shows the members in the candidate set. The leftmost column shows the input test characters to be recognized.

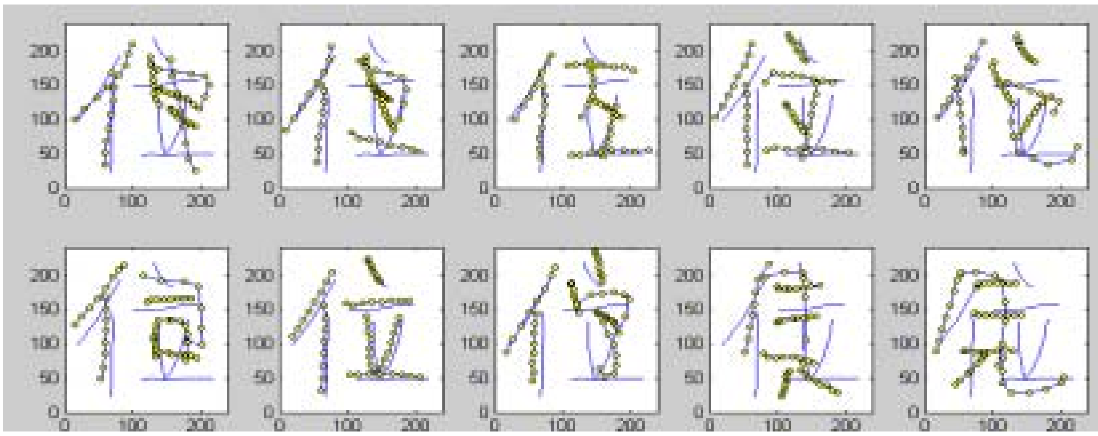
	伸	但	佐	住	佗	伺	位	佇	貝	見
但	3.02	1	4.71	5.03	5.04	6.59	5.28	5.07	5.44	5.53
位	4	4.02	2.45	1.90	2.44	2.92	1	2.32	4.56	4.7
見	3.94	3.33	3.97	3.17	3.64	4.04	3.66	3.57	1.35	1

Table 3. Normalized distance (μ) for Scheme 3 using the PDM-based approach. The top row shows the members in the candidate set. The leftmost column shows the input test characters to be recognized.

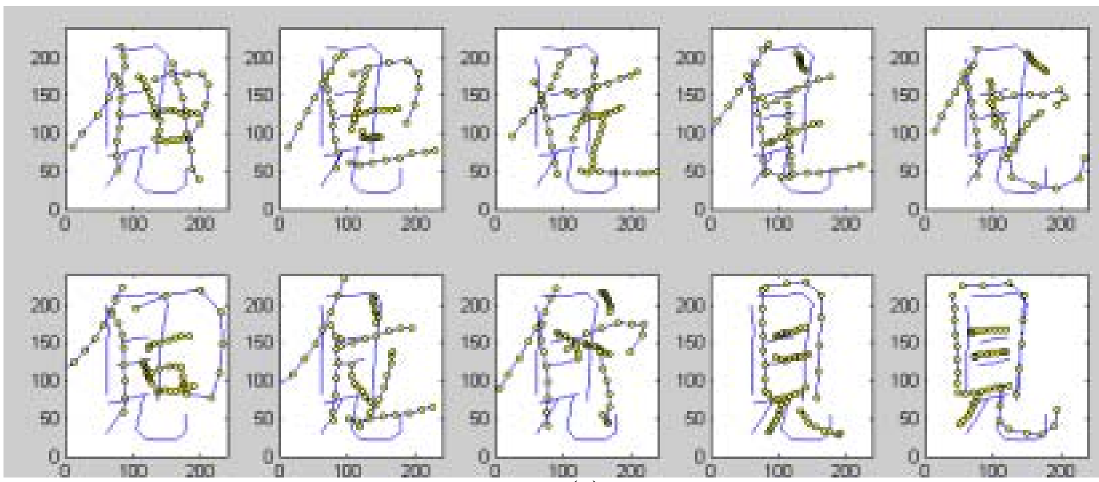
Ratio	伸	但	佐	住	佗	伺	位	佇	貝	見
但	3.87	1	4.02	3.96	4.53	4.47	4.6	5.93	8	7.98
位	4.27	2.87	2.48	1.95	2.75	3.14	1	4.12	5.12	5.39
見	5.06	4.28	4.31	3.39	4.49	4.51	3.93	5.62	1.39	1



(a)



(b)



(c)

Figure 2. Deformed mean template shapes vs. the test character. (a) The input character “但” is matched against the ten deformed templates contained in the candidate set. (b) The input character “位” is matched against the ten deformed templates contained in the candidate set. (c) The input character “見” is matched against the ten deformed templates contained in the candidate set. The same candidate characters “伸”, “但”, “佐”, “住”, “佗”, “伺”, “位”, “佇”, “貝”, and “見” are used in (a) through (c).