Manuscript for short paper submitted to *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Version 1, May 2, 2002

Handwritten Chinese Radical Recognition Using Nonlinear Active Shape Models

D. Shi, S. R. Gunn and R. I. Damper*, *Senior Member, IEEE*

Image, Speech and Intelligent System (ISIS) Research Group

Department of Electronics and Computer Science

University of Southampton

Southampton SO17 1BJ

United Kingdom

*Corresponding author (Email: rid@ecs.soton.ac.uk. FAX: +44 (0) 23 8059 4498)

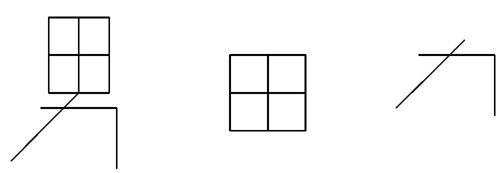
Abstract

Chinese charater recognition is one of the most difficult pattern recognition problems because it concerns complex structure, serious interconnection among the components, numerous pattern variations, and a large number of characters. Fortunately, a significant characteristic is that many thousands of Chinese characters can be made up from only a set of basic shapes, or *radicals*. Radical approaches decompose Chinese characters into a small set of radicals, then, the complex character recognition problem is converted to a simpler problem of radical extraction and optimization of combination of the radical sequences. In this paper, nonlinear active shape modeling is applied to handwritten Chinese radical recognition, in which kernel principal component analysis is employed to capture nonlinear handwriting variations, the chamfer distance transform is used to get satisfactory basins of attraction, and a dynamic tunneling algorithm is incorporated with the gradient descent to search for the optimal shape parameters. The radical recognition rate is 96.5%, which is superior to the existing approaches. We expect that it will also prove an important method for addressing other skeleton-based pattern recognition problems.

Keywords: Handwritten Chinese Radical recognition, active shape model, kernel principal component analysis, chamfer distance transform, dynamic tunneling algorithm.

1 Introduction

Chinese character recognition is one of the most difficult pattern recognition problems because it concerns complex structure, serious interconnection among the components, numerous pattern variations, and a large number of characters. Fortunately, a significant characteristic is that many thousands of Chinese characters can be made up from only a set of basic shapes, or *radicals* [1]. Radical approaches decompose Chinese characters into a small set of radicals, then, the complex character recognition problem is converted to a simpler problem of radical extraction and optimization of combination of the radical sequences. Figure 1 illustrates an example of Chinese character which can be decomposed into two radicals. The upper radical, Figure 1(b), means *land* (which looks like 2×2 grid); the lower radical, Figure 1(c), means *labor* (which looks like a bending person working), and Figure 1(a) consisting of the above two radicals, exposes how our ancestors described a man: *man is land labor*!



(a) (b) (c)

Figure 1: Illustration of a Chinese character decomposed into its radicals. (a) The Chinese character for *man*. (c) The upper radical for *land*. (d) The lower radical for *labor*.

Lam and Suen [2] applied deformable templates to handwriting digit recognition. Their method included two stages, i.e., a tree classifier working on the basis of physical structure and a relaxation matching phase. Those characters which can not be satisfactorily assigned to a class by physical structural analysis are passed to a slower relaxation matching algorithm which uses deformation to match the digit to each template.

Since radical approaches provide a way to decompose complex Chinese characters into simpler radicals, radical recognition can be treated as shape extraction from a given image. Moreover, deformable modeling is required to capture handwriting variation as well as to handle individuals.

Jain [3] investigated the application of deformable templates to handwriting recognition. This deformation system represents a binary image in terms of its contour, and then iteratively computes parameters of a continuous displacement function in order to map the contour template as closely as possible onto the edges of the target image. They reported a 99.25% recognition rate on a 2000 character subset of NIST Special Database 1.

Recently, Chung and Ip [4] applied snakes [5] to handwritten Chinese character recognition with energy functional minimization. The external energy in their work consists of two different functionals, i.e., displacement and intersection functionals. The displacement functional is employed to avoid the snake deviating too much from the original template. The intersection functional is used to avoid the intersection of character strokes and the template. However, snakes are forced onto the image by smoothness and some salient features. In their work, they did not discuss how to deal with the important problems of false salient features resulting from broken strokes and thinning algorithms.

Cootes et al. [6] proposed active shape models (ASMs) to capture the shape variation in an iterative search procedure, capable of locating the modeled structures in noisy, cluttered images—even if they are partially occluded. Active shape models have similarities to snakes, in which a contour is fitted to the image evidence by minimizing an energy functional. However, a snake only has generic prior knowledge, such as smoothness. A much greater amount of prior information can be recovered from training sets and encoded within an ASM.

Active shape modeling extracts the eigenvectors, $\mathbf{U}$, of the training examples by principal component analysis (PCA). Thereafter, models can be generated by adjusting shape parameters, $\mathbf{b}$, corresponding to the principal modes:

$$\mathbf{\Gamma} = \mathbf{\Psi} + \mathbf{U}\mathbf{b} \tag{1}$$

where $\mathbf{\Psi}$ is the mean vector of the training examples.

From equation (1), it is seen that the original ASMs are only suitable for representing linear variations within the point distribution models. However, nonlinear shape variations are common in handwriting, such as different writing styles from person-to-person, and time-varying distortion.

To generalize ASMs to the nonlinear case, Sozou et al. [7] introduced polynomial regression. The basis for the polynomial regression model is to reduce further the residuals once a linear mode has been extracted, by fitting a polynomial along the direction of the principal components. However, the technique requires that the second eigenvector can be modeled as a function of the first, otherwise implausible shapes may be generated. To solve this problem, they also applied a multilayer perceptron (MLP) [8] to find the nonlinear functionals among the shape parameters, $\mathbf{b}$. The many well-known disadvantages of the MLP method include the possibilities of over- or under-fitting the training data, and the sensitivity to the choice of perceptron architecture and to the initial start point(s) for training.

Heap and Hogg [9] described applications where the training examples form a highly nonlinear space, which can be successfully represented by piece-wise linear sub-models. They adopted a two-level hierarchical approach: an initial global PCA is carried out to produce a lower dimensional space, and then the linear subregions are constructed in this new lower space. They used the simplest way to constrain the valid shape region, i.e., Euclidean-distance-based clustering, which, however, cannot properly address the probabilistic nature of the problem.

Cootes et al. [10] represented shape variations with probability density function, which can be used to determine if a generated shape is plausible. They introduced kernel-based density estimation technique, which is constructed from a large number of kernels. To provide a cheaper way, they used a mixture of a mixture of small number of gaussians, which can be fit to the kernel estimate using a modification of the expectation maximization algorithm. The problem with this this method is how to specify the optimal number of mixture components.

Romdhani et al. [11] is the first work which introduced kernel PCA to active shape models. Original input shape are mapped to a feature space through kernel functions, and then active shape modeling in feature space is completely the same as the linear case. The expected active shape models are pre-images of the generated active shape models in feature space. They presented this approach to modeling nonlinear 2D shapes of non-rigid 3D objects and simultaneous recovering of object pose at multiple views and across the view sphere.

To address the problem of determination of valid shape regions (by upper bound on the modulus in feature space), Twining and Taylor [12] designed a 'proximity to data' measure functional. They demonstrated the performance of this function on both artificial and real-world nonlinear data examples. However, such a function may not be continuous, which calls for special treatment in searching shape parameters.

In this paper, nonlinear ASMs are applied to handwritten Chinese radical recognition. Section 2 describes how ASMs are improved by kernel PCA to capture nonlinear handwriting variations. In Section 3, the use of nonlinear ASMs in handwritten character recognition is detailed. The chamfer distance transform is applied to get a satisfactory basin of attraction, and the dynamic tunneling algorithm is employed to overcome problems of local minima when searching for the optimal shape parameters. In Section 4, we describe some experiments to determine the performance of these techniques on handwritten Chinese radical recognition, before concluding in Section 5.

2 Shape Representation by Kernel PCA

Principal component analysis is a technique for extracting structure from possibly high-dimensional datasets. It is readily performed by solving an eigenvalue problem, or by using iterative algorithms which estimate principal components. PCA is useful when the original pattern space can be accurately described by a subspace spanned by the first several principal eigenvectors. Often, the data lie in a subspace, and if this is linear then a small number of principal components is sufficient to account for most of the variation in the data.

Kernel PCA [13] provides a way to extend linear PCA to nonlinear subspaces of the data. Here, linear PCA is performed in some high-dimensional feature space $\mathcal{F}$ which is related to the input space by a nonlinear map $\Phi : \mathbb{R}^N \rightarrow \mathbb{R}^M$. The number of nonlinear components obtained by kernel PCA can be greater than the original input dimension; the dimension size of the feature space is specified by the number of training examples. However, the method will confer no advantage if the data lie in a linear subspace. The main challenge with this method is how to choose an appropriate nonlinear transformation.

We start with a set of examples for a particular radical $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_M\}$, which are represented by N landmark points labeled semi-automatically. We use 10 points to represent a stroke: a line is drawn from the start point to the termination point of a stroke, and automatically determines the other points. Hence: $\mathbf{e}_k = (x_{k0}, y_{k0}, \ldots, x_{k(N-1)}, y_{k(N-1)})^T$. The mean vector of the set is then defined by

$\mathbf{\Psi} = \frac{1}{M}\sum_{k=1}^{M}\mathbf{e}_k$. In the original active shape models, the $\mathbf{\Psi}$ is calculated after the training examples are aligned, and then each training example is rotated and scaled into the tangent space to $\mathbf{\Psi}$ so as to minimize the distance between them. This process is implemented iteratively until it converges ([6]; [10]). However, such an alignment strategy does not work well with handwritten Chinese radicals. One of the reasons is that the local rotation may lead to different radical classes. Hence, alignment in this case will be fulfilled in this way: 1. Characters are deskewed based on the document layout analysis; 2. Each segmented character is normalized to (64×64) dot matrix; 3. In recognition, the unknown normalized charter image is rotated to a few positions, and all the radical classes are matched with the rotated images.

The details of kernel PCA applied to active shape models can be seen in Romdhani et al. [11]. The projections of a data point $\mathbf{e}$ onto the eigenvectors $\mathbf{V}^k$ in the feature space $\mathcal{F}$ can be defined as:

$$\beta_k(\mathbf{e}) = \mathbf{V}^{kT}\tilde{\mathbf{\Phi}}(\mathbf{e}) = \sum_{i=1}^{M'}\alpha_i^k\tilde{\mathbf{\Phi}}(\mathbf{e}_i)^T\tilde{\mathbf{\Phi}}(\mathbf{e}) = \sum_{i=1}^{M'}\alpha_i^k\tilde{K}(\mathbf{e}_i, \mathbf{e})$$

where M' is the number of principal components (modes). Any example in the training set can be approximated using the mean vector and a weighted sum of these deviations obtained from the first M' modes. The $\tilde{\mathbf{\Phi}}(\mathbf{e}_i)$ is the centralized point in the feature space $\mathcal{F}$ corresponding to the vector $\mathbf{\Psi}$. However, we have no centralized data in $\mathcal{F}$, so we cannot compute $\tilde{K}_{ij} = \tilde{\mathbf{\Phi}}(\mathbf{e}_i) \cdot \tilde{\mathbf{\Phi}}(\mathbf{e}_j)$ directly but relying on non-centered counterpart $\mathbf{K}$. Define the notations $1_{ij} = 1$ for all i, j, and $(1_M)_{ij} = 1/M$ to compute $\tilde{K}_{ij}$ ([14]): $\tilde{K}_{ij} = (\mathbf{K} - \mathbf{1}_M\mathbf{K} - \mathbf{K}\mathbf{1}_M + \mathbf{1}_M\mathbf{K}\mathbf{1}_M)_{ij}$.

Our purpose is to build up models for each handwriting class, which requires approximate representations of the data in input space rather than in feature space. To this end, by introducing shape parameters, we can also generate ASMs on the basis of kernel PCA with the following two steps:

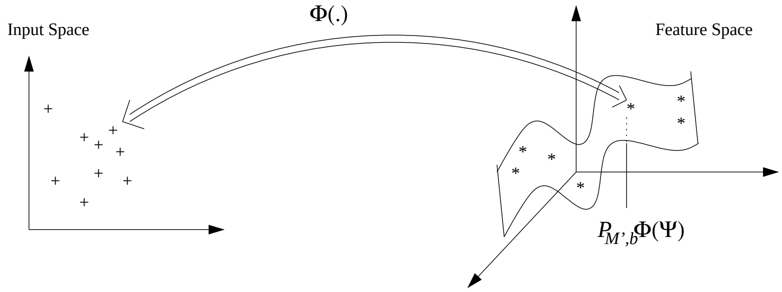


Figure 2: Illustration of pre-image in input space.

1. Generate active shape models in feature space with the mean vectors $\mathbf{\Psi}$. We define an operator $P_{M',b}$ by:

$$P_{M',b}\tilde{\mathbf{\Phi}}(\mathbf{\Psi}) = \sum_{k=1}^{M'} \beta_k(\mathbf{\Psi}) b_k \mathbf{V}^k.$$

2. Find the active model $\mathbf{\Gamma}$ which is a *pre-image* in the feature space so as to minimize:

$$\begin{aligned} \rho(\mathbf{\Gamma}) &= ||P_{M',b}\tilde{\mathbf{\Phi}}(\mathbf{\Psi}) - \tilde{\mathbf{\Phi}}(\mathbf{\Gamma})||^2 \\ &= \tilde{\mathbf{K}}(\mathbf{\Gamma},\mathbf{\Gamma}) - 2\sum_{k=1}^{M'} b_k \beta_k(\mathbf{\Psi}) \sum_{i=1}^{M} \alpha_i^k \tilde{\mathbf{K}}(\mathbf{e}_i,\mathbf{\Gamma}) + \tilde{\mathbf{K}}(\mathbf{\Psi},\mathbf{\Psi}). \end{aligned} \tag{2}$$

The following description may help understanding Equation (2). All the points in the input space will be mapped to a hyperplane in the feature space. As seen in Figure 2, a point in the feature space will be away from this hyperplane under linear operations (e.g. component truncation and shape parameter). Hence, its pre-image will be approximated by the one corresponding to its nearest hyperplane point in the feature space.

3 Radical Recognition with Nonlinear ASMs

With models generated by adjusting shape parameters, from equation (2), the radical is matched to the observed character by minimizing the distance between the radical model and the character skeleton.

3.1 Chamfer Distance Transform

The ASMs can be fitted to the target image by adjusting shape parameters. However, a satisfactory basin of attraction is needed to find the optimal shape parameters efficiently. The chamfer distance transform [15] is applied to touch this goal in this research. The most significant property of the transform is its ability to handle noisy and distorted data, as the edge points of one image are transformed by a set of parametric transformations, which describe how the images can be geometrically distorted in relation to one another. We have previously applied the chamfer distance transform to character recognition with some success [16] but for the linear PCA case.

A shape model is denoted by $\mathbf{\Gamma}(\mathbf{b})$, which can be generated by changing parameters $\mathbf{b}$. In the matching phase, the shape models for all the classes are superimposed upon the target image. Treating the target image as a functional of shape models, i.e., $I(\boldsymbol{\gamma}_j(\mathbf{b}))$, the energy of a shape model $\mathbf{\Gamma}$ is given by:

$$E(\mathbf{b}) = \sum_{j=1}^{N} D_{\text{chamfer}}(I(\boldsymbol{\gamma}_j(\mathbf{b}))) \tag{3}$$

where $\boldsymbol{\gamma}_j$ is the jth point of $\mathbf{\Gamma}$, which is located in a 2-dimensional position (x_j, y_j).

Hence, searching for the optimal shape is equivalent to minimizing the energy $E(\mathbf{b})$.

3.2 Gradient Descent with Dynamic Tunneling Algorithm

One solution to minimize the energy $E(\mathbf{b})$ of equation (3) is to use gradient descent. Given initial shape parameters $\mathbf{b}$, an initial model $\mathbf{\Gamma}$ is obtained. Given a test character image, it is matched against each model and the corresponding minimum energy $E(\mathbf{b})$, then the test image is classified into the class with the overall minimum energy.

To overcome problems with local minima we use a dynamic tunneling algorithm (DTA) [17]. The algorithm is based on a physical analogy to the quantum-mechanical tunneling of a particle through a potential barrier. The degree of tunneling allowed is a function of time, increasing during iterative search. The dynamic tunneling procedure can jump to another basin of attraction where the new, initial search point is even lower in energy. From this new starting point, gradient descent can again be used to find a lower minimum. The algorithm for searching the kth shape parameter b_k by gradient descent with DTA was given by [16]:

Step 1 $b_k^* = 0$, $b_k^+ = 0$.

Step 2 Search for local optimum with gradient descent from point b_k^* to obtain new current global minimum b_k^*.

Step 3 Dynamic tunneling phase begins, setting the tunneling time $t = 0$.

Step 4 $t = t + 1$.

Step 5 In the case of positive direction, select the point $b_k = b_k^* + \sqrt{(\frac{2t}{3\rho})^3}$;

In the case of negative direction, select the point $b_k = b_k^* - (b_k^+ + \sqrt{(\frac{2t}{3\rho})^3})$, i.e. treating b_k^+ as an offset)

Here ρ represents the strength of the repeller, and $\rho = 7000$ in our current experiments.

Step 6 In the case of positive direction, if $b_k > 3\sqrt{\lambda_k}$, then $b_k^+ = b_k^*$, and go to **Step 3**;

In the case of negative direction, if $b_k < -3\sqrt{\lambda_k}$, go to **Step 10**.

Step 7 Calculate the energy function $E_\Gamma$in b_k.

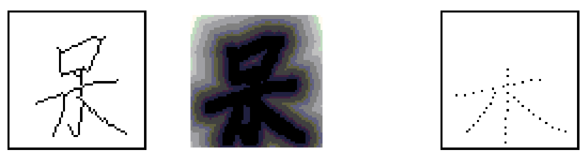
Step 8 If $E_\Gamma(b_k) > E_\Gamma(b_k^*)$ go to **Step 3** to continue dynamic tunneling phase.

Step 9 Otherwise, go to **Step 2** to start a gradient descent phase.

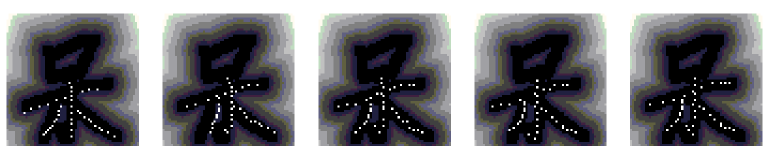
Step 10 End.

4 Experiments and Results

In the experiments to be described, nonlinear active shape modeling was applied into handwritten Chinese radical recognition. All experiments reported were carried out with Gaussian kernels, $k(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{2\sigma^2}\right)$, which gave better results than polynomial or spline kernels. Figure 3(a) shows an example character skeleton and its chamfer transform, while Figure 3(b) shows a typical mean model for characters from one of the classes. Figure 3(c) shows the ASMs generated as the number of principal components varies from 1 to 5, whereupon the chamfer distance between the models and the input skeleton is 198, 104, 43, 32 and 26 respectively. Our purpose is not accurate reconstruction of the image but pattern classification, the number of principal components can be truncated such that $\left(\sum_{i=1}^{M'} \lambda_i\right) / \left(\sum_{j=1}^{M} \lambda_j\right) > 90\%$.



(a) (b)



E(b)	198	104	43	32	26
M'	1	2	3	4	5

(c)

Figure 3: Active model generation by adjusting shape parameters. (a) An example of character skeleton and its chamfer transformed image. (b) Mean models. (c) Active models generated with the number of principal components from 1 to 5.

These experiments have used a subset of the database collected by Harbin Institute of Technology and Hong Kong Polytechnic University, which comprises a total of 751,000 loosely-constrained handwritten Chinese characters. There are 3755 categories written by 200 different writers [16].

We now compare some representative works on radical recognition with our proposed nonlinear ASM method.

Method 1: Nonlinear active shape models with kernel PCA. The experiments are conducted on 98 radicals covering 1400 loosely-constrained Chinese character categories written by 200 different writers (i.e., 280,000 characters). There are 60 randomly selected character examples with landmark labelling for training a particular radical. Since these 60 examples are unseen by the other radical categories, they are still put into the test set.

Method 2: Active shape models with linear PCA [16]. The experiments are conducted on the above dataset.

Method 3: Stroke-based approach by Wang and Fan [18]. Their experiments for radical extraction were conducted on just 32 radicals and 1856 test characters.

Method 4: Snake-fitting approach by Chung and Ip [4]. Their experiments were conducted on 100 character categories written by 10 people (i.e., 1000 test examples only). They considered and reported results on six most common radical combination schemes, namely, vertical, left-down, surrounding, horizontal, up-left, and cover, respectively.

From Table 1, we can see that the nonlinear ASM method is easily the best among the existing radical approaches. It deals with the largest number of radicals on a test set which is significantly larger than other workers have used, and still achieves the best correct matching rate.

Table 1: Performance comparison of different radical approaches to Chinese character recognition.

	Test set size (characters)	Number of radicals trained	% radicals correct
Method 1 (Nonlinear ASM)	280,000	98	96.5
Method 2 (ASM, linear PCA)	280,000	98	94.2
Method 3 (stroke-based)	1856	32	92.5
Method 4 (snake fitting)	1000	100	68.0 - 95.0

The poorer correct recognition rate of Method 3 results from a problem of ambiguity when strokes intersect. At the point of intersection, it is problematic as to which of the radiating lines should be grouped together so that some strokes may be spurious. As snakes are forced onto the image by smoothness and some salient features, Method 4 suffers from false salient features because of broken strokes and thinning algorithms.

The advantage of our proposed ASM approach is the capability to handle individual writer variations with only a small number of shape parameters. Our models can also avoid stroke extraction, as mentioned above, which is a source of difficulty in handwriting recognition as there will be considerable interconnection among the strokes as well as many broken strokes. However, this method requires longer matching time (0.4 seconds per character in Pentium III PC, 455MHz, 128MRAM) caused by working at the pixel level and shape-parameter searching. Importantly, radical extraction in our approach can be run in parallel, so overall computation time can be reduced.

5 Conclusions and Future Work

The paper has introduced a novel method of handwritten Chinese radical recognition based on nonlinear active shape models. Kernel principal component analysis is employed to capture the main nonlinear variation of the training examples around the mean vector. In matching, the dynamic tunneling algorithm is incorporated with gradient descent to search for the optimal shape parameters in terms of chamfer distance minimization. The experimental results show that our method achieves superior performance compared with existing similar representative works. Our proposed nonlinear ASMs are also suitable for other skeleton-based pattern recognition, such as numeral and English character recognition.

We have been doing research on character composition, including finding the optimal radical combination on the basis of a pre-compiled lexicon and the dynamic programming. Our future work also includes (fully) automatic landmark point labeling. One feasible approach is stroke extraction method associated with labeled reference radicals. A reference radical is created for each class, in which all the strokes involved are labeled. An alternative is solve reparameterisation/alignment/clustering problem simultaneously ([19]). However, there has been no strong report to support this method applied to skeleton shape with no connectivity information. Further investigation is required to deal with touching/broken strokes in our case.

References

[1] S. K. Chang. An interactive system for Chinese character generation and retrieval. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(3):257–265, 1973.

[2] L. Lam and C. Y. Suen. Structural classification and relaxation matching of totally unconstrained handwritten zip-code numbers. *Pattern Recognition*, 21(1):19–31, 1988.

[3] A. K. Jain and D. Zonker. Representation and recognition of handwritten digits using deformable templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(12):1386–1391, 1997.

[4] F. L. Chung and W. W. S. Ip. Complex character decomposition using deformable model. *IEEE Transactions on Systems, Man, and Cybernetics–Part C: Applications and Reviews*, 31(1):126–132, 2001.

[5] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1987.

[6] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models—their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.

[7] P. D. Sozou, T. F. Cootes, C. J. Taylor, and E. C. Di Mauro. Non-linear generalization of distribution models using polynomial regression. *Image and Vision Computing*, 13(5):451–457, 1995.

[8] P. D. Sozou, T. F. Cootes, C. J. Taylor, and E. C. Di Mauro. Non-linear point distribution modelling using a multi-layer perceptron. In *Proceedings of British Machine Vision Conference, BMVC'95*, pages 107–116, Birmingham, UK, 1995.

[9] T. Heap and D. Hogg. Improving specificity in PDMs using a hierarchical approach. In *Proceedings of British Machine Vision Conference, BMVC'97*, pages 80–89, Colchester, UK, 1997.

[10] T. F. Cootes and C. J. Taylor. A mixture model for representing shape variation. *Image and Vision Computing*, 17(8):567–574, 1999.

[11] S. Romdhani, S. Gong, and A. Psarrou. A multi-view nonlinear active shape model using kernel PCA. In *Proceedings of British Machine Vision Conference*, pages 483–492, Nottingham, UK, 1999.

[12] C. J. Twining and C. J. Taylor. Kernel principal component analysis and the construction of non-Linear active shape models. In *Proceedings of British Machine Vision Conference II*, pages 23–32. Manchester, UK, 1982.

[13] B. Schölkopf, A. J. Smola, and K. Müller. Kernel principal component analysis. In B. Schölkopf, C. J. C. Burges, and A. J. Smola, editors, *Advances in Kernel Methods*, pages 327–352. MIT Press, Cambridge, MA, 1998.

[14] B. Schölkopf, S. Mika, A. Smola, G. Rätsch, and K. Müller. Kernel PCA pattern reconstruction via approximate pre-images. In L. Niklasson, M. Bodén, and T. Ziemke, editors, *Proceedings of the 8th International Conference on Artificial Neural Networks*, pages 147–152. Springer Verlag, Berlin, Germany, 1998.

[15] G. Borgefors. Hierarchical chamfer matching: A parametric edge matching algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(6):849–865, 1988.

[16] D. Shi, S. R. Gunn, and R. I. Damper. A radical approach to handwritten Chinese character recognition using active handwriting models. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 670–675, Hawaii, USA, 2001.

[17] Y. Yao. Dynamic tunneling algorithm for global optimization. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(5):1222–1230, 1989.

[18] A. B. Wang and K. C. Fan. Optical recognition of handwritten Chinese characters by hierarchical radical matching method. *Pattern Recognition*, 34(1):15–35, 2001.

[19] N. Duta, A. K. Jain, and M. P. Dubuisson-Jolly. Automatic construction of 2D shape models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(5):433–446, 2001.

[20] A. J. Viterbi. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory*, 13:260–269, 1967.

Reply to Review A:

1. ... automatic landmark labeling ....

Actually, our current labeling method is semi-automatic. Please refer to current page 6. We use 10 points to represent a stroke: a line is drawn from the start point to the termination point of a stroke, and automatically determines the other points. In page 14, we list two possibility of our future investigation of (fully) automatic landmark point labeling. One feasible approach is stroke extraction method associated with labeled reference radicals. Another is Duta, Jain and Jolly's method. However, there has been no strong report to support this method applied to skeleton shape with no connectivity information, and it needs further investigation how to deal with touching/broken strokes in our case, which are very important prior knowledge.

2. ... processing time ....

In page 13 we have given the processing time – 0.4 seconds per character in Pentium III PC, 455MHz, 128MRAM.

3. ... training and test sets ....

In page 10: there are 60 randomly selected character examples with landmark labelling for training a particular radical. Since these 60 examples are unseen by the other radical categories, they are still put into the test set.

4. ... radical recognition or character recognition? ....

The paper title is now changed to "Handwritten Chinese Radical Recognition Using Nonlinear Active Shape Models", all the related description has been focused on radical recognition. An radical approach can convert the complex Chinese character recognition problem into a simpler problem of extracting a small number of radicals and then finding the optimal combination of these radicals. Our future work also includes finding the optimal radical combination on the basis of a pre-compiled lexicon and the Viterbi algorithm ([20]).

Reply to Review B:

1. ... missing references and novelty ....

We have briefly introduced works of Sozou et al. [7], [8], Heap and Hogg [9], Cootes et al. [10], and Romdhani et al. [11]. But none of these is discussing handwriting recognition. We have shrunk the section about KPCA, but still claim it is a novel approach to handwritten Chinese radical recognition, in which KPCA is used to capture the nonlinear variation, chamfer distance transform is used to obtain satisfactory basin of attraction, and the dynamic tunneling algorithm is incorporated with gradient descent to search for the optimal shape parameters.

2. ... description of the KPCA and DTA ....

As we submit our manuscript as a short paper, we have to further shorten the description of KPCA. But we do stengthen the issues about shape alignment (see page 5), centralisation in feature space (see page 6) and dynamic tunneling (see page 9).

Reply to Review C:

1. ... explain how to apply these techniques ....

Now, some detailed description of KPCA has been removed, whereas more explanations and illustration of KPCA and dynamic tunneling have been strengthened. The whole character recognition will be fulfilled by radical combination based on the Viterbi algorithm ([20]). We have been doing research on finding the optimal radical combination on the basis of a pre-compiled lexicon and the Viterbi algorithm.

2. ... automatic landmark labeling ....

We have explained this situation in the last paragraph. Our current labeling method is actually semi-automatic. We have review and possibly correct all the points marked by the automatic method of ([19]). There is no connectivity information in this case, and there are lots of broken or touching strokes. A feasible approach is stroke extraction method associated with labeled reference radicals. A fully automatic labeling is difficult and there is much room to investigate.

3. ... explanation of radicals ....

We have given more explanation about the radical idea in the beginning of the introduction.

4. ... difference from the authors' PRL paper ....

The PRL paper is emphasized on character recognition on the basis of radical candidates, but this paper is discussion radical recognition in detail.

5. The system is very time-consuming. How to resolve this problem.

Importantly, radical extraction in our approach can be run in parallel, so overall computation time can be reduced.

Reply to Review D:

1. ... missing references ....

More references have been added, such as, Sozou et al. [8], Heap and Hogg [9], Cootes et al. [10], and Romdhani et al. [11].

Summary of Revision

1. **Title has been changed.** The novelty of this paper is handwritten Chinese radical recognition using nonlinear ASMs, in which, KPCA is used to capture nonlinear handwriting variation, chamfer diatance transfer is introduced to get satisfactory basin of attraction, and dynamic tunneling algorithm is employed to search for the optimal shape parameters. Our proposed methodology may also address other skeleton-based pattern recognition problems.

2. The description of radical ideas have been strengthened in Page 1.

3. Some important references related to nonlinear active shape models have been added in Page 3-4.

4. Semi-automatic landmark point labeling has been described in Page 5. We use 10 points to represent a stroke: a line is drawn from the start point to the termination point of a stroke, and automatically determines the other points.

5. The alignment issue has been discussed in Page 5-6. In the original active shape models, the $\mathbf{\Psi}$ is calculated after the training examples are aligned, and then each training example is rotated and scaled into the tangent space to $\mathbf{\Psi}$ so as to minimize the distance between them. This process is implemented iteratively until it converges ([6]; [10]). However, such an alignment strategy does not work well with handwritten Chinese radicals. One of the reasons is that the local rotation may lead to different radical classes. Hence, alignment in this case will be fulfilled in this way: 1. Characters are deskewed based on the document layout analysis; 2. Each segmented character is normalized to (64×64) dot matrix; 3. In recognition, the unknown normalized charter image is rotated to a few positions, and all the radical classes are matched with the rotated images.

6. The introduction of KPCA has been briefed, whereas the explanation to KPCA has been strenthened in Page 7.

7. Shape parameter searching by dynamic tunneling algorithm with gradient descent has been given in Page 9.

8. The future work of fully automatic landmark labeling has been stated.One feasible approach is stroke extraction method associated with labeled reference radicals. A reference radical is created for each class, in which all the strokes involved are labeled. An alternative is solve reparameterisation/alignment/clustering problem simultaneously ([19]). However, there has been no strong report to support this method applied to skeleton shape with no connectivity information. Further investigation is required to deal with touching/broken strokes in our case.