

# A Radical Approach to Handwritten Chinese Character Recognition Using Active Handwriting Models

D. Shi      S. R. Gunn      R. I. Damper

Image, Speech and Intelligent Systems (ISIS) Research Group,  
Department of Electronics and Computer Science,  
University of Southampton,  
Southampton SO17 1BJ, United Kingdom  
Email: [ds99r,srg,rid]@ecs.soton.ac.uk

## Abstract

*This paper applies active handwriting models (AHM) to handwritten Chinese character recognition. Exploiting active shape models (ASM), the AHM can capture the handwriting variation from character skeletons. The AHM has the following characteristics: principal component analysis is applied to capture variations caused by handwriting, an energy functional on the basis of chamfer distance transform is introduced as a criterion to fit the model to a target character skeleton, and the dynamic tunneling algorithm (DTA) is incorporated with gradient descent to search for shape parameters. The AHM is used within a radical approach to handwritten Chinese characters recognition, which converts the complex pattern recognition problem to recognizing a small set of primitive structures – radicals. Our initial experiments are conducted on 98 radicals covering 1400 loosely-constrained Chinese character categories written by 200 different writers. The correct matching rate is 94.2% on these  $2.8 \times 10^5$  characters. Comparison with existing radical approaches shows that our method achieves superior performance.*

## 1. Introduction

It is known that handwritten Chinese character recognition is one of the most challenging topics in pattern recognition, because it involves a large number of characters with complex structure, serious interconnection among the components, and considerable pattern variation [14]. There are more than 70,000 Chinese characters, of which about 4,000 are used daily, and this scale of usage challenges current recognition systems. However, most of Chinese characters can be decomposed into a few fundamental simple graphs [10], called *radicals* [2]. These radicals have different sizes and are placed in specific positions with respect to each other to make a legitimate character. Depending on the rad-

icals chosen, codes of different levels of complexity are required to generate a character using the pre-defined radicals.

Conventionally, there are two possible ways to extract radicals, from a character skeleton or on the basis of strokes. Skeleton-based methods, such as [5] and [4], treat a radical as a subimage of the character skeleton image. These methods aim to discover the relationship among the hierarchically represented graphs and capture their variations. Stroke-based methods, such as [15] and [8], decompose a radical further into its primitive structural parts, i.e., straight-line strokes, and then recognize the whole character by structural analysis. The advantage of the latter approach is that it requires far less computation than skeleton-based methods, but it suffers from a problem of ambiguity when strokes intersect.

Ip et al. [5] applied snake fitting [7] to Chinese radical extraction with energy functional minimization. Their experiments were conducted on 36 character classes written by 10 people, and the initial results were promising. However, snakes are forced to fit the target image by applying constraints of smoothness and some salient features. In their work, they did not mention how to deal with false salient features resulting from broken strokes and the inherent defects of thinning algorithms.

Fukushima et al. [4] proposed a skeleton-based radical approach to handwritten Japanese Chinese character recognition using the neocognitron—a neural network capable of recognizing distorted patterns as well as tolerating positional shift. Up to now, it is still difficult to bring neocognitron-based methods to practical use, because too much expert domain knowledge is required to design its training patterns. In previous work, GA-based learning of the neocognitron was used to search for parameters and training patterns automatically in handwritten numeral recognition [11]. However, it cannot work well on handwritten Chinese character recognition, as there is a huge search space in this case.

Wang and Fan [15] proposed a radical-based system for recognizing handwritten Chinese characters. In their approach, a recursive hierarchical scheme is developed to perform radical extraction first. Character features and radical features are then extracted for matching. Finally, a hierarchical radical-matching scheme is devised to identify the radicals embedded in the character, so enabling recognition. Using this approach, the complexity of off-line handwritten Chinese character recognition is greatly reduced.

The merit of Liao and Huang's work [8] is that radical extraction is not confused by spurious strokes and the inherent defects of thinning algorithms, as mentioned above. Their method consists of three parallel matching algorithms, which extract radicals parallelly at stroke, skeleton and pixel level. However, this method is quite time consuming. In addition, one practical issue is that a decision must be made as to which algorithm to select.

Since there is extensive interconnection among the strokes in handwritten Chinese characters, stroke extraction is extremely difficult and brings considerable ambiguity. Active shape modeling [3] can build up a deformable model by changing a small number of parameters to generate the principal variations of the models. Active shape models have similarities to snakes, in which a contour is fitted to the image evidence by minimizing an energy function. However, a snake has only generic prior knowledge, such as smoothness, whereas a much greater amount of prior information exists about radicals, which can be recovered from training data and encoded within an active shape model.

In our previous work, active shape modeling was applied to radical extraction for handwritten Chinese characters [12]. Its weakness was due to local minima because of the standard gradient descent technique employed. In Section 2, we improve this work by introducing a dynamic tunneling algorithm into the gradient descent, and we call our new method active handwriting models (AHM), which is expected to be suitable for any skeleton-based character recognition problem. In Section 3, the AHM is applied to Chinese radical recognition. Experiments and their results are given in Section 4, followed by conclusions and future work in Section 5.

## 2 Active Handwriting Models

This section will discuss active handwriting models (AHM), which can be expected to be suitable for any skeleton-based character recognition problem, although our current experiments are conducted on Chinese.

### 2.1 Capturing Shape Variation by Principal Component Analysis

Principal component analysis (PCA) 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 [6].

This section describes how to obtain the eigenvectors to capture the main directions of variance within the skeleton character. Given a set of examples for a character  $\{\Gamma_1, \Gamma_2, \dots, \Gamma_M\}$ , which are represented by  $N$  landmark points, i.e.,  $\Gamma_k = (x_{k0}, y_{k0}, \dots, x_{k(N-1)}, y_{k(N-1)})^T$ . The mean vector of the set is defined by  $\Psi = \frac{1}{M} \sum_{k=1}^M \Gamma_k$ . The centralized vector difference from the mean vector is given by  $S_k = \Gamma_k - \Psi$ . The centralized character matrix  $\mathbf{A}$ , is then formed as by  $[S_1 S_2 \dots S_M]$ , having  $2N$  rows and  $M$  columns. The eigenvectors,  $\mathbf{u}_i$ , and eigenvalues,  $\lambda_i$ , of the covariance matrix  $\mathbf{A}\mathbf{A}^T = \frac{1}{M} \sum_{k=1}^M S_k S_k^T$  are solutions to

$$(\mathbf{A}\mathbf{A}^T)\mathbf{u}_i = \lambda_i \mathbf{u}_i.$$

The variance explained by each eigenvector is equal to the corresponding eigenvalue. Most of the variation can usually be explained by a small number of modes,  $M'$ . 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:

$$\Gamma = \Psi + \mathbf{U} \cdot \mathbf{b} \quad (1)$$

Here  $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_{M'})$  is the matrix of the first  $M'$  eigenvectors, and the shape parameters,  $\mathbf{b} = (b_1, \dots, b_{M'})$  control the variation in shape around the mean radical.

### 2.2 Chamfer Distance Transform

Given a handwriting character model, we can employ gradient descent to adjust the shape parameters,  $\mathbf{b}$ , in equation 1 to fit the character image. A satisfactory basin of attraction is expected to aid finding the optimal shape parameter efficiently. The chamfer distance transform [1] can reach this goal. The most significant property of the chamfer distance transform is its ability to handle noisy and distorted data, as the edge points of one image are transformed by a set of parametric transformation equations, which describe how the images can be geometrically distorted in relation to one another. The transform approximates global distances by propagating local distances at image pixels. In the binary edge image, each edge pixel is first set to zero and each non-edge pixel is set to infinity. A  $3 \times 3$  window is

used to scan the image iteratively until no change has been brought to any pixel. For a pixel valued  $I(x, y)$ , its new value is calculated as follows:

$$I(x, y) = \min \begin{cases} I(x-1, y-1) + C_2 \\ I(x, y-1) + C_1 \\ I(x+1, y-1) + C_2 \\ I(x-1, y) + C_1 \\ I(x, y) \\ I(x+1, y) + C_1 \\ I(x-1, y+1) + C_2 \\ I(x, y+1) + C_1 \\ I(x+1, y+1) + C_2 \end{cases}$$

where, the pair of distance transform coefficients,  $C_1$  and  $C_2$ , are constants and satisfy:  $0 < C_1 < C_2 < 2C_1$ . In our experiments,  $C_1$  and  $C_2$  are set to 3 and 4 respectively, which makes the maximum difference from Euclidean distance 8% percent [1].

Let  $I'$  be the chamfer distance transformed image of  $I$ , which is a functional:  $I'(x, y) = D_{\text{chamfer}}(I(x, y))$ . Hence, the energy of a radical model  $\Gamma$  is given by:

$$E(\mathbf{b}) = \sum_{j=1}^N D_{\text{chamfer}}(I(\Gamma_j(\mathbf{b}))) \quad (2)$$

where  $\Gamma_j$  is the  $j$ th point of  $\Gamma$ , which is located in a 2-dimensional position  $(x_j, y_j)$ .

We want to find the minimum,

$$\frac{\partial E}{\partial \mathbf{b}} = \sum_{j=1}^N \frac{\partial}{\partial \mathbf{b}} D_{\text{chamfer}}(I(\Gamma_j(\mathbf{b}))) = 0. \quad (3)$$

One solution to this problem is to use a gradient descent based approach. Given an initial position of a model  $\Gamma$ , the solution is obtained by adjusting the model coefficients,  $\mathbf{b}$ , using the gradient information, equation 3.

## 2.3 Gradient descent with dynamic tunneling algorithm

Standard gradient descent leads to local minima in the basin of attraction nearest to the starting point. In other words, the efficiency of the algorithm depends heavily on the initial point and topology of the surface associated with the objective function. Some methodologies find global minima but may fail to do so within a reasonable time, such as simulated annealing and the multiple random start method. They are unsuitable for our problem, in which the computational time is an important indicator of the system performance.

Yao [16] proposed the dynamic tunneling algorithm (DTA), which is based on a physical analogy to the quantum-mechanical tunneling of a particle through a potential barrier. Also, the degree of tunneling allowed is a function of time, increasing during iterative search. The DTA concept is governed by the fact that any particle placed at a small perturbation from the equilibrium point will move away from the current point to another within a finite amount of time.

RoyChodhury et al. [9] hybridized gradient descent algorithms with DTA, in which they introduced a dynamic system

$$\frac{du}{dt} = u^{1/3}.$$

Any initial condition which is infinitesimally close to the repelling point  $u = 0$  will escape the repeller, to reach point  $u_0 \neq 0$  in a finite time given by

$$t_1 = \int_0^{u_0} u^{-1/3} du = \frac{3}{2} u_0^{2/3}.$$

Hence, 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.

Since all the eigenvectors are orthogonal, the shape parameters can be specified one by one. The algorithm for searching the  $k$ th shape parameter  $b_k$  by gradient descent with DTA is given as follows [13]:

- Step 1**  $b_k^* = 0, b_k^+ = 0$ .
- Step 2** Searching the local optimum starting with the point  $b_k^*$ , and the current global minimum  $b_k^*$  is obtained.
- Step 3** Dynamic tunneling procedure begins, setting the tunneling time  $t = 0$ .
- Step 4**  $t = t + 1$ .
- Step 5** In the case of positive direction, select the point  $b_k = b_k^* + \frac{2}{3\rho} \sqrt{t^3}$ ;  
In the case of negative direction, select the point  $b_k = b_k^* - b_k^+ - \frac{2}{3\rho} \sqrt{t^3}$ .  
Here  $\rho$  represents the strength of the repeller, and  $\rho = 100$  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 2**;  
In the case of negative direction, if  $b_k < -3\sqrt{\lambda_k}$ , go to **Step 10** to end the algorithm.
- Step 7** Calculate the energy function in  $b_k$ .
- Step 8** If  $E(b_k) > E(b_k^*)$  go to **Step 3** to continue dynamic tunneling procedure.
- Step 9** Otherwise, go to **Step 2** to start a gradient descent procedure.
- Step 10** End.

After gradient descent with DTA, the optimal shape parameters for each radical class specify the corresponding energy minimum according to equation (2). The decision rule used here is then simply to select the radical class with the overall minimum, although better performance might be expected taking the context of the radical into account (e.g. using a precompiled lexicon).

### 3 Chinese Radical Extraction Using Active Handwriting Models

Although there are many thousands of Chinese characters, it is fortunate that only a small number of radicals can compose many different Chinese characters [2, 10]. Radical approaches decompose Chinese characters into a small set of radicals, so the complex character recognition problem is converted to a simpler problem of radical extraction and optimization of a combination of the radical sequences.

In this section, active handwriting models will be applied to radical extraction for handwritten Chinese characters. Since there is less interconnection within the peripheral structures than within inner structures in a typical Chinese character, the peripheral radicals are usually preferred for classification. In our research, all the peripheral radicals can be sorted into 9 types: left-hand side, right-hand side, upper side, lower side, surrounding and 4 diagonal corner radicals. So, in our active radical models, any mean radical has its own initial location. This treatment simplifies the matching phase in which we do not need to search over the whole character image.

#### 3.1 Training Phase

Training phase includes landmark point labeling and principal component analysis. The first step is to extract character skeletons by an image thinning algorithm. Then, landmark points are labeled manually to represent a radical, and principal component analysis is applied to obtain the main shape parameters which capture the radical variations. Figure 1 illustrates the process for the simple radical *mu* (meaning *tree*).

Figure 2 shows eigenvectors corresponding to the top three eigenvalues calculated from 60 training examples of the category in Figure 1, where  $\lambda_T = \sum_{i=1}^M \lambda_i$ .

#### 3.2 Matching Phase

Recognition phase includes chamfer distance transform on the target image and shape parameter searching via gradient descent with dynamic tunneling algorithm. From Figure 3, shows examples of some character skeletons and their chamfer distance transformed images.

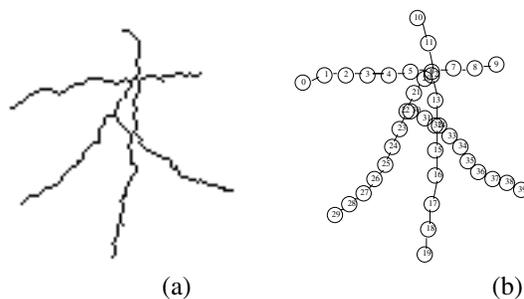


Figure 1: Example of (a) the skeleton for the radical (*mu* or *tree*). (b) the landmark points for the skeleton in (a).

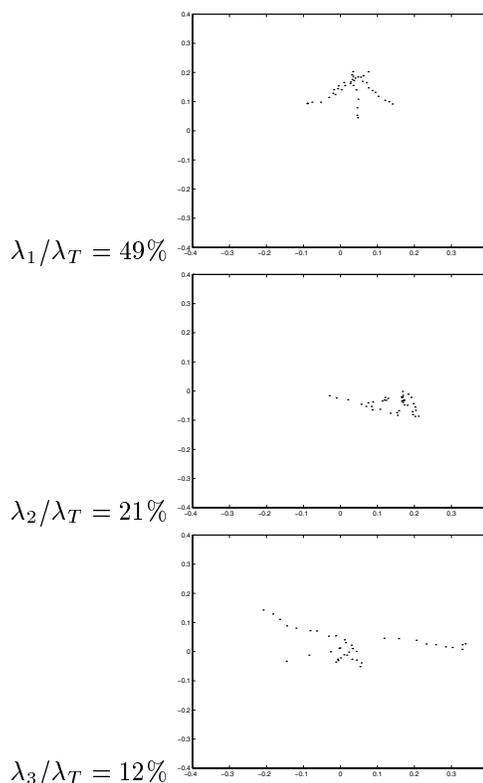


Figure 2: Eigenvectors corresponding to top three eigenvalues,  $\lambda_i$  for the radical in Fig. 1.

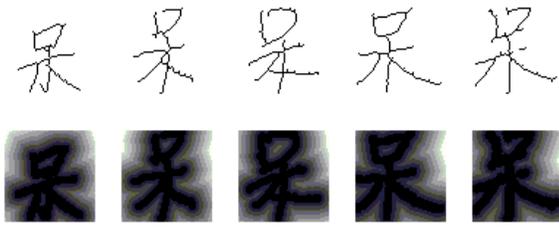


Figure 3: Character skeletons and their chamfer distance transformed images.

From Figure 3, we see that the effect of the transform is to blur the original image so as to tolerate noise and distorted data. Hence, a basin of attraction is created. The distance between a model and target image will be calculated by superimposing the model onto the chamfer distance transformed image.

The above-mentioned dynamic tunneling algorithm is incorporated with gradient descent technique to search for the optimal shape parameters for each radical class. The shape parameters corresponding to the eigenvectors can be set respectively, as all the eigenvectors are orthogonal.

## 4 Experiments and Results

Our database was collected by Harbin Institute of Technology and Hong Kong Polytechnic University, and comprises a collection of 751,000 loosely-constrained handwritten Chinese characters, consisting of 3755 categories written by 200 different writers [12, 13]. In our current experiments, the radical training sets include 98 radical classes, and each class has 60 different examples (each from a different writer) which are the relevant parts from the character examples.

The experiments were conducted on a PC (Pentium III 450 MHz, 128 M RAM) using C++. The 98 radicals in our experiments can cover 1400 commonly-used Chinese characters. A lexicon of these 1400 characters has been built up, in which each character is a 9-dimensional feature corresponding to 9 types of radical, namely, left, right, upper, down, surrounding and 4 corner radicals [13].

If the first candidate radical extracted from the character image in a particular position is the same as the relevant one in the lexicon, the matching is treated as correct. On this basis, the matching rate is 94.2% radicals correct, on the  $2.8 \times 10^5$  characters, where each character is composed of 1 to 4 radicals. The average number of radicals per character is approximately 2.6.

We are now in a position to compare our radical recognition results with those of other researchers.

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

	Test set size (characters)	Number of radicals trained	% radicals correct
Method 1 (AHM)	$2.8 \times 10^5$	98	94.2
Method 2 (ARM)	$2.8 \times 10^5$	66	75.4
Method 3 (stroke-based)	1856	32	92.5
Method 4 (snake fitting)	360	not reported	85.8
Method 5 (neocognitron)	N/A	8	N/A

**Method 1:** Active handwriting modeling, with gradient descent and dynamic tunneling, as described immediately above. The experiments are conducted on 98 radicals covering 1400 loosely-constrained Chinese character categories written by 200 different writers (i.e.,  $2.8 \times 10^5$  characters).

**Method 2:** Active radical modeling without dynamic tunneling algorithm [12]. The experiments are conducted on the same database as above.

**Method 3:** Stroke-based approach [15]. Their experiments for radical extraction were conducted on just 1856 test characters.

**Method 4:** Snake-fitting approach [5]. Their character image database consists of 36 character categories written by 10 people (i.e., 360 test examples only).

**Method 5:** Neocognitron-based approach [4]. In their preliminary experiment, designed to show the potential of the technique, only 8 illustrative radical classes were considered. No explicit correct rate was given.

From Table 1, we can see that use of dynamic tunneling in conjunction with gradient descent (Method 1) improves results considerably over use of gradient descent alone (Method 2). Our method of active handwriting modeling is easily the best among the existing radical approaches. It deals with the largest number of radicals on a test set much larger than other workers have used, and still achieves the best correct matching rate.

The advantage of our method is not only the avoidance of straight-line stroke extraction, but also the ability to capture the variations with only a small number of shape parameters. As stated earlier, it is difficult to implement stroke extraction within a handwritten Chinese character as there will

be considerable interconnection among the strokes as well as many broken strokes. It is also a crucial point to handle the individual writer variation in handwriting recognition.

The disadvantage of our method is its long matching time caused by working at the pixel level and shape-parameter searching. Fortunately, radical extraction in our method can be run in parallel, so that problems of computation time can be addressed.

## 5 Conclusions and Future Work

Radical approaches decompose Chinese characters into a small set of radicals, so 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, an approach to radical recognition for handwritten Chinese characters is proposed, based on active handwriting modeling (AHM). Only a small number of principal components can capture the main variations of radicals. Chamfer distance minimization is used to match radicals within a character. Use of a dynamic tunneling algorithm (DTA) in conjunction with gradient descent improves performance considerably over gradient descent alone. Experiments for radical extraction are conducted on 98 radicals covering 1400 loosely-constrained characters from 200 writers, and the matching rate obtained with AHM is 94.2% radicals correct. We also compare our method to existing radical approaches. The conclusion is that our method gives superior performance. It benefits from avoidance of (straight line) stroke extraction, as well as the ability to capture the variations by only a small number of shape parameters.

Future work includes investigating kernel PCA to determine whether nonlinear variations caused by handwriting can improve performance. The AHM will also be applied to other skeleton-based object recognition problems.

## References

- [1] G. Borgefors, "Hierarchical chamfer matching: A parametric edge matching algorithm", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 6, pp. 849-865, 1988.
- [2] S. K. Chang, "An interactive system for Chinese character generation and retrieval", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 3, No. 3, pp. 257-265, 1973.
- [3] T. F. Cootes, C. J. Taylor, D. H. Cooper and J. Garaham, "Active shape models—their training and application", *Computer Vision and Image Understanding*, Vol. 61, No. 1, pp. 38-59, 1995.
- [4] K. Fukushima, T. Imagawa and E. Ashida, "Character recognition with selective attention", In: *Proceedings of International Joint Conference on Neural Networks, IJCNN'91*, Seattle, WA, pp. A593-A598, 1991.
- [5] W. W. S. Ip, K. F. L. Chung and D. S. Yeung, "Offline handwritten Chinese character recognition via radical extraction and recognition", In: *Proceedings of Third International Conference on Document Analysis and Recognition*, Ulm, Germany, pp. 185-189, 1995.
- [6] I. T. Jolliffe, *Principal Component Analysis*, Springer-Verlag, New York, NY, 1986.
- [7] M. Kass, A. Witkin and D. Terzopoulos, "Snakes: Active contour models", *International Journal of Computer Vision*, Vol. 1, No. 4, pp. 321-331, 1988.
- [8] C. W. Liao and J. S. Huang, "A Transformation invariant matching algorithm for handwritten Chinese character recognition", *Pattern Recognition*, Vol. 23, Np. 11, pp. 1167-1188, 1990.
- [9] P. RoyChodhury, Y. P. Singh and R. A. Chansarkar, "Hybridization of gradient descent algorithms with dynamic tunneling methods for global optimization", *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, Vol. 30, No. 3, pp. 384-390, 2000.
- [10] G. Sampson, *Writing Systems*, Hutchinson, London, UK, 1985.
- [11] D. Shi, C. Dong and D. S. Yeung, "Neocognitron's parameter tuning by genetic algorithms", *International Journal of Neural Systems*, Vol. 9, No. 6, pp. 497-509, 1999.
- [12] D. Shi, S. R. Gunn and R. I. Damper, "Active radical modeling for handwritten Chinese characters", In: *Proceedings of Sixth International Conference on Document Analysis and Recognition*, Seattle, WA, 2001.
- [13] D. Shi, S. R. Gunn and R. I. Damper, "Radical recognition of handwritten Chinese characters based on active shape modeling", submitted to *Pattern Recognition*, 2001.
- [14] Y. Y. Tang, L. T. Tu, J. Liu, S. W. Lee, W. W. Lin and I. S. Shyu, "Offline recognition of Chinese handwriting by multi-feature and multilevel classification", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 20, pp.556-561, 1998.
- [15] A. B. Wang and K. C. Fan, "Optical recognition of handwritten Chinese characters by hierarchical radical matching method", *Pattern Recognition*, Vol. 34, No. 1, pp. 15-35, 2001.
- [16] X. Yao and Y. Liu, "Towards designing artificial neural networks by evolution", *Applied Mathematics and Computation*, Vol. 91, No. 1, pp. 83-90, 1998.