

Active Radical Modeling for Handwritten Chinese Characters

Daming Shi, Steve R. Gunn and Robert I. Damper

Image, Speech and Intelligent Systems Research Group,
Department of Electronics and Computer Science,
University of Southampton,
Southampton SO17 1BJ, UK.

Abstract

Handwritten Chinese character recognition is one of the most difficult problems of pattern recognition. Since the majority of Chinese characters are made up from just a small set of primitive structures – radicals – this paper describes an approach to active radical modeling for such handwritten characters. The most significant characteristic of our method is that radicals can be found robustly without stroke extraction, and the principal variations of the radical can be encoded in a small number of parameters. In the training phase, the example radicals are represented by manually-labeled ‘landmark’ points. Then a small number of principal components of the eigenvectors are calculated to capture the main variation of the training examples from the mean radical. In the matching phase, each radical model is fitted to the image evidence by adjusting the shape parameters in terms of chamfer distance minimization. Initial experiments are conducted on 1100 loosely-constrained Chinese character categories written by 200 different writers. The correct matching rate is 95.8%, showing that our radical modeling is effective and capable of forming a sound basis for handwritten Chinese character recognition.

1 Introduction

As a logographic writing system, Chinese script represents a precious cultural heritage (Sampson 1985). Nowadays, one third of the world’s population reads and writes in Chinese or Chinese-like scripts. So, the advancement of Chinese information processing will undoubtedly bring a great commercial and social benefit. Further, Chinese characters are known to be highly complex, and form a large character set (> 4000). Hence, handwritten optical Chinese character recognition is one of the most challenging topics in pattern recognition (Tang, Tu, Liu, Lee, Lin, and Shyu 1998).

Although there are many different Chinese characters, the majority can be decomposed into a few fundamental parts, or *radicals* – what Sampson (1985) calls “simple graphs”. These occur in distinct regions of the character, such as left-hand radical, upper radical and so on. Hence, decomposition into radicals offers an attractive approach to recognition. Wang and Fan (2001) proposed a radical-based optical character recognition (OCR) system for handwritten Chinese, in which a recursive hierarchical scheme is used to extract radicals from an input character represented by strokes.

Since there is extensive interconnection among the strokes in handwritten Chinese characters, stroke extraction is extremely difficult and brings considerable ambiguity. Our research is aimed at avoiding stroke extraction by building up a deformable model for each radical with active shape modeling (Cootes, Taylor, Cooper, and Garaham 1995; Blake and Isard 1998), in which the principal variations of the radical are encoded in a small number of parameters. Active shape models have similarities to snakes (Kass, Witkin, and Terzopoulos 1988), in which a contour is fitted to the image evidence by minimizing an energy function. But a snake has only generic prior knowledge, such as smoothness. A much greater amount of prior information exists about radicals, which can be recovered from training sets and encoded within an active shape model.

The remainder of this paper is organized as follows. In Section 2, active radical

modeling is described, including landmark point labeling and eigenvector computation. In Section 3, we describe how radical matching within a character is achieved by adjusting the shape parameters in terms of chamfer distance minimization. Experiments and their results are given in Section 4, followed by conclusions in Section 5.

2 Active Radical Modeling

Although there are many thousands of Chinese characters, 90% of these can be classified with only about 200 radicals. If each radical is recognized as a single pattern, the difficult problem of Chinese character recognition can be converted to a simpler character composition from the radicals. In our research, a radical is defined by a position (or perhaps, more strictly, a region) and its shape in the characters. This treatment simplifies the matching phase such that we do not need to search over the whole character image.

2.1 Landmark point labeling

Active shape modeling requires some *landmark points* to capture shape constraints in training and construct plausible new examples of the shape for use in matching. As Chinese characters consist of strokes, the terminating points together with some interpolated points are labeled as landmark points to represent strokes, and 10 points are chosen to represent a stroke in our research. In training, these are labeled manually. Landmark point labeling is labor intensive: For instance, 5000 points must be labeled for a 50-point radical class with 100 examples. In matching, the landmark points are found automatically and used as a basis to measure the difference between the image object and a radical model.



Figure 1: Some examples of (a) a radical (*mu* or *tree*) and (b) the landmark points for the first example in (a).

2.2 Computation of eigenvectors

This section describes how to obtain the eigenvectors to capture the principal directions of variance within the radical, which is a crucial point in active radical modeling. Given a set of examples for a radical $\{\Gamma_1, \Gamma_2, \dots, \Gamma_M\}$, which are represented by N landmark points, i.e., $\Gamma_i = (x_{i0}, y_{i0}, \dots, x_{i(N-1)}, y_{i(N-1)})^T$. The mean radical of the set is defined by $\Psi = \frac{1}{M} \sum_{k=1}^M \Gamma_k$. The centralized radical difference from the mean radical is given by $\Phi_i = \Gamma_i - \Psi$. The centralized radical matrix \mathbf{A} , is then formed as by $[\Phi_1 \Phi_2 \dots \Phi_M]$, having M columns and $2N$ rows. The eigenvectors \mathbf{u}_i and eigenvalues λ_i of the covariance matrix $C = \frac{1}{M} \sum_{k=1}^M \Phi_k \Phi_k^T = \mathbf{A} \mathbf{A}^T$ 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' . In practice, a small M' is sufficient for pattern recognition, since accurate reconstruction of the image is not a requirement. Figure 2 shows eigenvectors corresponding to the top three eigenvalues calculated from 50 training examples of the category in Figure 1, where $\lambda_T = \sum_{i=1}^M \lambda_i$.

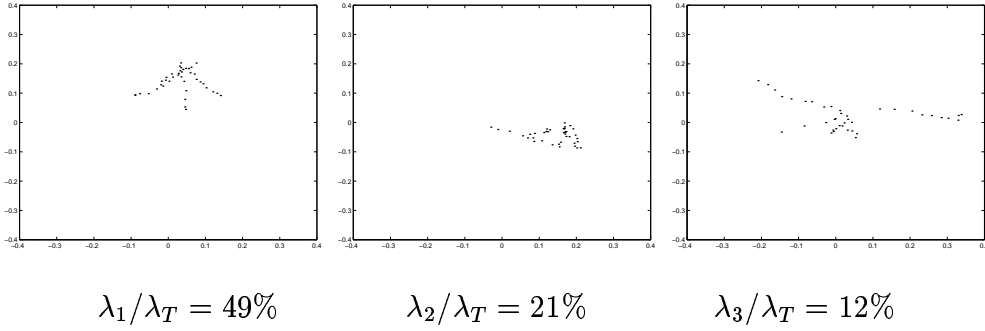


Figure 2: Eigenvectors corresponding to top three eigenvalues, λ_i for the radical in Fig. 1.

Any radical in the training set can be approximated using the mean radical and a weighted sum of these deviations obtained from the first M' modes:

$$\Gamma = \Psi + \mathbf{U} \cdot \mathbf{b}$$

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. The variance of b_k is limited depending on the application: In our case, $-3\sqrt{\lambda_k} \leq b_k \leq 3\sqrt{\lambda_k}$.

Figure 3 shows the shapes generated by varying the shape parameters separately. Obviously, the first shape parameter captures the variation of image scale. The matching phase will be carried out by varying b_k actively to fit the radical model to the image object.

3 Radical Matching Within a Character

In this section, we describe how the shape parameters \mathbf{b} are adjusted to fit the character image in terms of chamfer distance minimization.

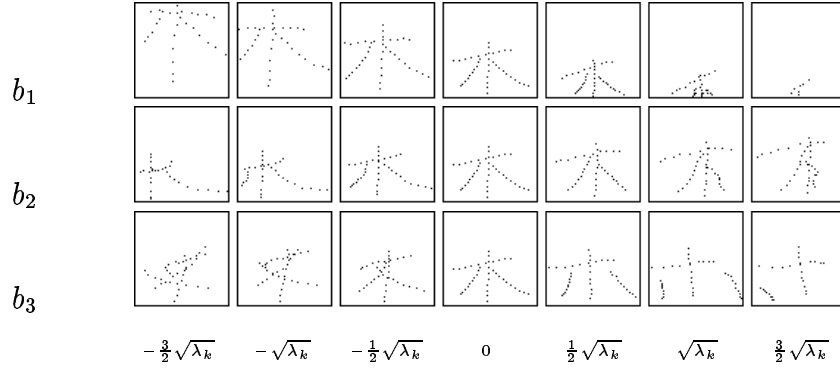


Figure 3: The effects of varying the first three shape parameters.

3.1 Chamfer distance transform

The chamfer distance transform was first proposed by Barrow, Tenenbaum, Bolles, and Wolf (1977), and then systematically developed by Borgefors (1988). Its most significant characteristic is its ability to handle noisy and distorted data, as the edge points of one image are transformed by a set of parametric 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 chamfer distance is a sequential distance transform. A (3×3) window is used to scan the image forwards (from top-left to bottom-right) and backwards (from bottom-right to top-left). For a pixel valued $f(x, y)$, its new value will be calculated as follows:

$$\text{Forwards: } f(x, y) = \min \begin{cases} f(x-1, y-1) + C_2 \\ f(x, y-1) + C_1 \\ f(x+1, y-1) + C_2 \\ f(x-1, y) + C_1 \\ f(x, y) \end{cases} \quad \text{Backwards: } f(x, y) = \min \begin{cases} f(x, y) \\ f(x, y+1) + C_1 \\ f(x-1, y+1) + C_2 \\ f(x, y+1) + C_1 \\ f(x+1, y+1) + C_2 \end{cases}$$

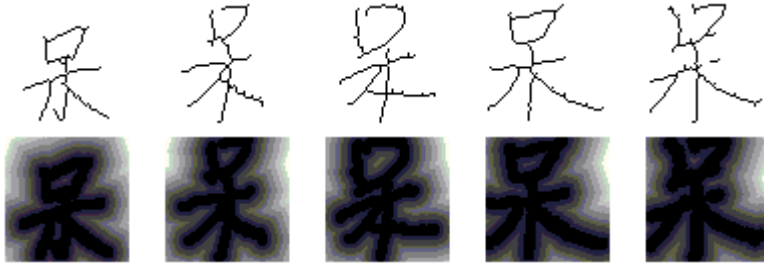


Figure 4: Character skeletons and their chamfer distance transformed images.

where C_1 and C_2 are constants equal to 3 and 4 respectively in the work of Borgefors (1988).

3.2 Chamfer distance minimization

Assume the current estimate of the position is Γ . To fit to an image object I , we need to find a set of adjustments which will move each point towards a better position.

First, the computation of chamfer distance is carried out by superimposing the model points on the distance transformed image I' . The chamfer distance can be measured by the root mean square average: $D(\Gamma, I') = \frac{1}{3} \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} v_i^2}$ where the v_i are the distance values and N is the number of landmark points. A perfect fit between the model and the distance image will result in a value of zero.

Then, for the sake of simplicity at this stage, gradient descent is used to adjust the shape parameters \mathbf{b} . On the basis of the current estimate Γ , adjustment is carried out iteratively to determine a displacement which moves each model point to a better location. Hence, the models attempt to deform to fit the character image evidence better. This is repeated until no more significant changes occur.

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 (Shi, Damper, and Shu 2000). Radical training sets are built up by taking the relevant parts from the character examples.

As 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. All peripheral radicals can be sorted into nine types: left-hand side, right-hand side, upper side, lower side, surrounding and four diagonal corner radicals. So, in our active radical models, any mean radical has its own initial location. We then need to select shape parameters to fit to the image evidence so as to minimize the chamfer distance.

The experiments were conducted on a PC (Pentium III 450 MHz, 128 Mbyte RAM). Initially, we trained 66 radicals which can compose 1100 commonly-used Chinese characters. A lexicon has been built up, in which each character is a 9-dimensional feature corresponding to 9 types of radical. 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. The correct matching rate is 95.8% and the average matching time is 0.95 seconds per character.

In previous work (Shi, Gunn, Damper, and Shu 2000), there were 18 types of outline primitive features used in the structural analysis. Since the outline primitive features are only a rather coarse description of Chinese characters, many particular feature descriptions (actually 34%) mapped to *different* characters. Hence, some other types of features (such as stroke density features) had to be used to specify the characters uniquely. By contrast, using radicals we can describe characters accurately. With the

new approach, the percentage of ambiguous feature codes falls to 2.5%.

Unlike the recent radical approach proposed by Wang and Fan (2001), we are able to capture the variation without stroke extraction. As a matter of fact, it is difficult to implement stroke extraction within a handwritten Chinese character as there will be much interconnection among the strokes as well as many broken strokes. We are currently trying to obtain the database of Wang and Fan so as to allow a performance comparison of the two methods.

The disadvantage of our method is its long matching time, which is expected to be solved in future work. Importantly, radical extraction in our approach can be run in parallel, so overall computation time can be reduced.

5 Conclusions

An approach to active radical modeling for handwritten Chinese characters has been proposed. Only a small number of principal components can capture the main variations of radicals, and chamfer distance minimization is used to match radicals within a character. Experiments for radical extraction were conducted on 1100 loosely-constrained characters and the correct matching rate of radical extraction is 95.8%. As radicals are extracted without stroke extraction (which is very difficult in handwriting recognition), our proposed method is competitive with previously-described radical modeling techniques. It is being further developed to form the basis of a handwritten Chinese character recognition system.

References

- Barrow, H. G., J. M. Tenenbaum, R. C. Bolles, and H. C. Wolf (1977). Parametric correspondence and chamfer matching: Two new techniques for image matching.

In *Proceedings of 5th International Joint Conference on Artificial Intelligence*, Cambridge, MA, pp. 659–663.

- Blake, A. and M. Isard (1998). *Active Contours*. Berlin, Germany: Springer.
- Borgefors, G. (1988). Hierarchical chamfer matching: A parametric edge matching algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 10(6), 849–865.
- Cootes, T. F., C. J. Taylor, D. H. Cooper, and J. Garaham (1995). Active shape models – their training and application. *Computer Vision and Image Understanding* 61(1), 38–59.
- Kass, M., A. Witkin, and D. Terzopoulos (1988). Snakes: Active contour models. *International Journal of Computer Vision* 1, 321–331.
- Sampson, G. (1985). *Writing Systems*. London, UK: Hutchinson.
- Shi, D., R. I. Damper, and W. Shu (2000). Chinese OCR using GA and extension matrix algorithm. *Communications of COLIPS, International Journal of Chinese and Oriental Language Information Processing Society* 9(1), 137–154.
- Shi, D., S. R. Gunn, R. I. Damper, and W. Shu (2000). Recognition rule acquisition by an advanced extension matrix algorithm. *International Journal of Engineering Intelligent Systems* 8(2), 97–101.
- Tang, Y. Y., L. T. Tu, J. Liu, S. W. Lee, W. W. Lin, and I. S. Shyu (1998). Offline recognition of Chinese handwriting by multifeature and multilevel classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(5), 556–561.
- Wang, A. B. and K. C. Fan (2001). Optical recognition of handwritten Chinese characters by hierarchical radical matching method. *Pattern Recognition* 34, 15–35.