Handwritten Chinese Character Recognition Using Kernel Active Handwriting Model*

Daming Shi, Yew Soon Ong and Eng Chong Tan
School of Computer Engineering
Nanyang Technological University
50 Nanyang Avenue, Singapore 639798
[asdmshi,asysong,asectan]@ntu.edu.sg

Abstract - This paper describes a kernel active handwriting model (K-AHM) and its application to handwritten Chinese character recognition. In the model, the kernel principal component analysis is applied to capture nonlinear variations caused by handwriting, and a fitness function on the basis of chamfer distance transform is introduced to search for the optimal shape parameters using genetic algorithms (GAs). The K-AHM is applied to handwritten Chinese character recognition, which converts the complex pattern recognition problem to recognizing a small set of primitive structures call radicals. Treating Chinese character composition as a discrete-time Markov process, the character composition is carried out with the Viterbi algorithm. The proposed methodology has been successfully implemented in an experimental recognition system.

Keywords: Active handwriting model, kernel principal component analysis, genetic algorithms, Viterbi algorithm.

1 Introduction

The fundamental goals in handwriting recognition are to capture the inter-writer and intra-writer shape variations. There has been a growing interest in deformable models or deformable templates. Such deformable models/templates are able to synthesize a very close approximation to any image of the target object, and can adjust themselves to the target shape. They are capable of capturing the natural variability within a class of shapes and can be used in image search.

Jain and Zonker [4] investigated the application of deformable templates to handwritten character recognition. This deformation system represents a binary image in terms of its intensity contours, and then iteratively computes parameters of a continuous displacement function to map the contour template accurately onto the edges of the target image. They reported a 99.25% recognition rate on a 2000 character subset of the NIST Special Database 1.

A milestone in the development of deformable templates is the work of active contour models (or snakes) [5]. A snake is described as an elastic line which moves under the influence of forces derived from an energy function. This energy minimization is modeled as having stiffness and elasticity and is attracted toward features such as lines and edges. Constraints can also be applied to ensure that the snake remains smooth and to limit the degree to which it can be bent. Snakes can be considered as parameterized models, the parameters being the spline control points. They are usually free to take almost any smooth boundary with few constraints on their overall shape. The idea of fitting by using image evidence to apply forces to the model and minimizing an energy function is effective.

Cootes et al. [3] concluded that a general drawback of the snakes is the lack of specificity, since their model parameters are limited on the basis of their distributions determined from a training set. As the model parameters are correlated over the training set, they do not effectively restrict the shapes which can be generated to ones similar to those found in the original training set. It is necessary that a deformable model should be able to accommodate the range of variation found in the objects it is used to represent. The principal role of a model is to facilitate robust automatic interpretation even in images which are noisy or cluttered or where parts of the objects of interest may be occluded. If the model is nonspecific, in the sense that it is able to deform so as to represent objects which are not valid examples of the class to be recognized, then this robustness is compromised. To this end, Cootes et al. developed active shape models (ASMs) to find a basis for shape representation in which the shape parameters are uncorrelated over the training set. In this case, simple limits on each parameter constrain the model to generate shapes similar to those in the training set.
The original ASMs are only suitable for representing linear variations within the point distribution models. However, nonlinear shape variations may be common in handwriting, such as different writing styles from person to person, and time-varying distortion. To generalize ASMs to the nonlinear case, Romdhani et al. [6] are the first to introduce the kernel principal component analysis (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.

Exploiting ASMs, Shi et al. [8] proposed active handwriting models (AHMs), and extended them by the kernel PCA [10]. This paper describes the improvement on our previous work by incorporating genetic algorithms and provides an effective solution to handwriting recognition. In Section 2, kernel active handwriting models are introduced. The roles of K-AHM for feature extraction in handwritten Chinese character recognition and the Viterbi algorithm for the optimal character composition, are described in Section 3. Finally, the experimental results are given in Section 4 followed by the conclusions in Section 5.

2 Kernel Active Handwriting Model

In the following discussion, any handwriting character is represented as a binary skeleton image. Active handwriting models extract the eigenvectors, $\mathbf{U}$, of the training examples using PCA, and then the model, $\mathbf{\Gamma}$, can be generated by adjusting the shape parameters, $\mathbf{b}$, corresponding to the principal modes: $\mathbf{\Gamma} = \mathbf{\Psi} + \mathbf{Ub}$, where $\mathbf{\Psi}$ is the mean vector of the training examples.

The use of kernel PCA for developing the nonlinear active handwriting models is discussed as two phases in this section. During training, the handwriting nonlinear variation is captured using kernel PCA, and at recognition, each radical model will be fitted to the target character by adjusting the shape parameters.

2.1 Training Phase

Kernel PCA provides a way to extend linear PCA to nonlinear subspaces of the data [7]. Here, linear PCA is performed in some high-dimensional feature space, which is related to the input space by a nonlinear map. The number of nonlinear components obtained by kernel PCA can be greater than the original input dimension. The dimension size of feature space is specified by the number of training examples. However, the method will confer no advantage if the data lies in a linear subspace. The main challenge with this method is how to choose an appropriate nonlinear transformation.

The details of variation capture by kernel PCA can be seen in [10]. The projections of a data point $\mathbf{e}$ onto the eigenvectors $\mathbf{V}^k$ in the feature space can be defined as

$$\beta_k(\mathbf{e}) = \mathbf{V}^k \mathbf{\Phi}(\mathbf{e}) = \sum_{i=1}^{M'} \alpha_i^k \tilde{K}(\mathbf{e}_i, \mathbf{e})$$  \hspace{1cm} (1)

where $M'$ is the number of principal components, $\mathbf{\Phi}(\mathbf{e}_i)$ is the centralized point in the feature space corresponding to the mean vector $\mathbf{\Psi}$, and $\tilde{K}(\mathbf{e}_i, \mathbf{e})$ is the centralized kernel matrix.

Active handwriting models can be generated on the basis of kernel PCA with the following two steps:

(i) Generate active handwriting models in feature space with the mean vectors $\mathbf{\Psi}$. An operator $\mathbf{P}^1_{M',b}$ is defined by:

$$\mathbf{P}^1_{M',b} = \sum_{k=1}^{M'} \beta_k(\mathbf{\Psi}) b_k \mathbf{V}^k$$  \hspace{1cm} (2)

(ii) Find the active model $\mathbf{\Gamma}$ which is a pre-image in the feature space so as to minimize

$$\rho(\mathbf{\Gamma}) = \left\| \mathbf{P}^1_{M',b} - \mathbf{\Phi}(\mathbf{\Gamma}) \right\|^2$$  \hspace{1cm} (3)

2.2 Recognition Phase

The recognition phase consists of performing a chamfer distance transform [1] on the target image and shape parameter searching via gradient descent with dynamic tunneling algorithm. Optimal shape parameters will be obtained by minimizing the mean-square distance between the model and the input character in the input space. The most significant property of this 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.

Given an unknown character, all the model classes will adjust their shape parameter to fit to it. Such procedures are terminated once the optimal shape parameters are found. The optimal active shape model can be found by genetic algorithms as follows: 1) Let the population size be $100 \times M'$. 2) The searching range for
any shape parameter $b_k$ is $-3\sqrt{\lambda_k} < b_k < 3\sqrt{\lambda_k}$ at step 0.001. 3) Generate active shape models, whose fitness are their chamfer distance to the target image. 4) Let the probability of crossover be 0.8, the probability of mutation be 0.05. 5) The searching procedure is terminated when the number of different individuals is $1000 \times M'$.

3 Handwritten Chinese Character Recognition with K-AHM

Handwritten 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. Chinese characters follow a hierarchical representation. A graph of the Chinese writing system stands not for a unit of pronunciation but for a morpheme, a minimal meaningful unit of the Chinese language. These simple graphs which are known as radicals, can compose many different Chinese characters. Radical approaches decompose Chinese characters into a small set of categories, so the complex character recognition problem is converted to a simpler problem of radical extraction and optimization of combination with the radical sequences.

3.1 Radical extraction with kernel active handwriting modeling

The Chinese radical extraction with kernel active handwriting models is illustrated in figure 1. An input skeleton is converted to its chamfer distance transformed image. Then all the radical classes generate active models and find their optimal shapes according to chamfer distance minimization. The output of radical extraction level is a set of radicals ranked by their chamfer distance to the given character.

3.2 Radical composition with Viterbi algorithm

This section describes the character composition based on these radical sequences. Treating Chinese characters as discrete-time Markov processes, the optimal radical combination is equivalent to the best path in the graph made up by all radical classes. Markov methods invoke the assumption that the language is a Markov source and uses transition probabilities.

In this research, Chinese character recognition is associated with a graph where the nodes contain radical recognition scores. A one-to-one correspondence exists whereby every path through the graph branches corresponds to a particular legal segmentation of the input character into radicals, and conversely, every possible legal segmentation of the input character corresponds to a particular path through the graph. In this graph representation, the Viterbi algorithm provides a convenient method for rapid determination of the best-scoring path (corresponding to an interpretation of a character). The radicals are extracted from 9 different peripheral positions. The outputs of the radical extraction are then considered as symbol probabilities.

The survivor is defined the shortest path corresponding to a node. Let us define the following symbols:

\begin{align*}
    f(j) & : \text{the symbol probability of the } j\text{th radical}. \\
    P((i,j) \mid (a,b)) & : \text{the transition probability from the node } (a,b) \text{ to } (i,j). \\
    \pi(j) & : \text{the initial probability of the } j\text{th radical}. \\
    \chi(i,j) & : \text{the node of } i\text{th row}, j\text{th column}. \\
    \hat{x}(i,j) & : \text{the survivor path ending at } \chi(i,j). \\
    L(i,j) & : \text{the survivor path value}. \\
    \nu (=4) & : \text{the total number of rows in the graph}. \\
    J (=200) & : \text{the total number of columns in the graph}. \\
\end{align*}

The optimal radical combination can be found by the following viterbi algorithm:

**STEP 1** Initialization.

\[ L(i,j) = 0 \; , \; \forall i, j \neq 0 \; ; \; j=1. \]

**STEP 2** \[ L(1,j) = \pi(j) \times f(j) \; ; \; \hat{x}(1,j) = (1,j). \]
STEP 3 $i = 2$.
STEP 4 Calculate:
\[
L(i, j) = \max_{1 \leq m \leq J} \left[ \hat{x}(i - 1, j) \times f(j) \right];
\]
\[
\hat{x}(i, j) = (i, m), \quad \text{s.t.}
\]
\[
\max_{1 \leq m \leq J} \left[ \hat{x}(i - 1, j) \times f(j) \right].
\]
STEP 5 $i \leftarrow i + 1$; Repeat Step 3 while $i \leq \nu$.
STEP 6 $j \leftarrow j + 1$; Go to Step 2 while $j \leq J$.
STEP 7 Termination and backtracking for best path. The best path is:
\[
\hat{x}(1, \nu), \hat{x}(2, \nu), \ldots, \hat{x}(\nu, \nu)
\]
\[
\text{s.t. } L(\nu, \nu) = \max_{1 \leq i, j \leq J} L(i, j)
\]

4 Experiments and Results

Our experiments are conducted on HITPU, a database collected by Harbin Institute of Technology and Hong Kong Polytechnic University, which comprises a collection of 751,000–loosely-constrained handwritten Chinese characters, consisting of 3755 categories written by 200 different writers (The HITPU database is freely available in http://www.ntu.edu.sg/home/asdmshi/hitpu.html).

Our proposed nonlinear active handwriting modeling with Viterbi algorithm achieves 94.8% character recognition correct rate on a subset of HITPU, which includes 200 radicals covering 2154 loosely-constrained Chinese character categories written by 200 different writers (430,800 characters). In contrast, the recognition rate obtained in our previous work is 93.5% characters correct [9]. The result is based on the dynamic tunneling algorithm, which only works well with effective gradient descent algorithm. However, there is a lack of image gradient information for use in the nonlinear case. Hence, the genetic algorithm which needs no sensitivity information, provides the perfect means to obtain the optimal shape parameters.

Other existing representative radical approaches can be found in [2] and [13]. Chung and Ip [2] applied snake fitting to Chinese radical extraction with energy functional minimization. The external energy in their work consists of two different functions, i.e., displacement and intersection functions. Their experiments were conducted on 100 character categories written by 10 people, and the initial results reported were 79.1% character correct. In fact, snakes are forced on to the image by smoothness and some salient features. In their work, they did not give further discussion on how to deal with false salient features because of broken strokes and thinning algorithms. Wang and Fan [13] proposed a radical based optical character recognition system for recognizing handwritten Chinese characters. Their recursive hierarchical radical extraction consists of three layers. Layer 1 is character pattern detection which classifies a given character into a shape pattern, such as left-right, up-down, etc. Layer 2 is straight cut-line detection which detects gaps among radicals. A stroke clustering technique is devised in layer 3 to decompose Chinese characters that are left-right or up-down patterns into radicals. Their hierarchical radical-matching scheme also consists of three matching phases. The recognition rate on their test set was 80.9% character correct. The lower correct rate of the method resulted 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.

5 Conclusions

Kernel active handwriting modeling is proposed and applied to handwritten Chinese characters recognition in this paper. In training, nonlinear active handwriting models capture the handwriting variations by kernel PCA. In matching, the genetic algorithm is employed to search for the optimal shape parameters by minimizing the chamfer distance between generated models and the input skeleton image. The proposed kernel active handwriting model is capable of handling the individual writer variation with only a small number of shape parameters, as well as searching for the shape parameters in the feature space in terms of the image information in the input space. Furthermore, the proposed method can provide a novel radical approach to handwritten Chinese character recognition to avoid stroke extraction, which is difficult in handwriting recognition as there will be considerable interconnection among the strokes. The Viterbi algorithm helps finding the maximum a posteriori probability estimation of the radical combination. Experiments conducted on 430,800 characters show that our methods outperform our previous work and other representative approaches.

References


