Recognition of Characters by Adaptive Combination of Classifiers

WANG Fei, LI Zai-ming
(School of Communication and Information Engineering, UESTC Chengdu 610054 China)

Abstract In this paper, the visual feature space based on the long Horizontals, the long Verticals, and the radicals are given. An adaptive combination of classifiers, whose coefficients vary with the input pattern, is also proposed. Experiments show that the approach is promising for character recognition in video sequences.

Key words character recognition; adaptive combination; multiple classifiers

Recognition of Chinese characters is difficult. However, most of the Chinese characters can be decomposed into a couple of fundamental parts, such as left-hand radical, right-hand radical, upper radical, and so on. In other words, a large number of different characters can be composed by a simple combination of a small number of radicals. If each of these radicals is recognized as a single pattern, the number of categories, whose patterns have to be classified, can be greatly reduced. These radicals are taught to be the model as individual patterns during the training phase. And a lot of Chinese characters have Horizontals (called “HENG” in Chinese) and Verticals (called “SHU” in Chinese). Considering the robust of the Horizontal and the Vertical, only the long Horizontal and the long Vertical are extracted as visual feature in this paper.

Recently, there is a popular belief that features and classifiers of different types could complement each other to some extent, therefore fine combination of multiple classifiers would result in improvement in performance [1]. The combination method is usually linear combination model and coefficients are usually constant, which restricts its adaptability and performance. In fact it is more appropriate and flexible to premise that the coefficients should vary with the input pattern. Based on this assumption, an adaptive combination model is proposed in this paper.

1 Visual Feature Extractor

1.1 The Long Horizontal Extractor
Suppose that the image is described as

$$[I(x,y)]_{M \times N}$$

where $M$ is the width of the image, $N$ is the height of the image. Scanning the text images in the horizontal direction, the number of the row object could be obtained. Suppose that $OR(i)$ represents the number of the $i$th row object, and $WR(i)$ represent the pixel numbers of the $i$th row. The long Horizontal model is described as following [2]

$$\sum_{k= j}^{j+h} OR(k) = h$$

where $h$ represents the density of the long Horizontal, $WR_i$ represents the width threshold of the long Horizontal, $h_s$ represents the height threshold of the long Horizontal.

The extraction of the long Vertical is similar to the Horizontal.

1.2 The Radical Extractor
Scanning the text images $[I(x,y)]_{M \times N}$ in the Horizontal direction, the number of the row object
could be obtained. And then the rows that have the same number of the row object in the vertical direction are combined. If the numbers of the row object change, then the new layer starts, thus the character images could be layered in certain layers. According to the shape of the layer, the layer can be classified into three kinds: the Punctuate layer, the Vertical layer, and the Horizontal layer.

The radical can be extracted based on the relation of the conjointly layer and the object number of the layers. For example, the radical “++” have three layers, the first and third layer both have two objects whose style are Punctuate, the second layer has one object whose style is Horizontal.

1.3 Location of the Visual Feature
In some Chinese characters, the same pattern can be placed in several different positions. In other words, the information only on the shape of radical is not enough to decide whether it is the left-hand, the right-hand, the upper radicals or the bottom radicals. In order to get the information on the position of the radical in a character, the approach also detect the center of gravity of the pattern segmented in layer.

2 Multiple Classifiers

The schematic diagram of the multiple classifiers is shown in Fig. 1

2.1 Individual Classifiers
Consider a pattern recognition problem where an input pattern \( X \) is assigned to be one of \( m \) possible classes. Let \( S \) represents the \( m \) possible classes, and \( C \) represents the \( n \) available classifiers.

\[
S = \{ S_i \} \quad i = 1, 2, \ldots, m \tag{4}
\]

\[
C = \{ C_j \} \quad i = 1, 2, \ldots, n \tag{5}
\]

Each classifier extracts different feature vector and represents the given pattern by a distinct measurement vector. Denote the feature vectors extractors extracted by the classifiers by \( F \) and the measurement vectors produced by the classifier by \( Y \). Then, the feature vector and measurement vector used in the classifier are defined respectively as

\[
F_j = (F_{j1}, \ldots, F_{jk}) \quad j = 1, 2, \ldots, n \tag{6}
\]

where \( k \) is the dimension of \( F_j \)

\[
Y_j = (Y_{j1}, \ldots, Y_{jm}) \quad Y_j \in R^m \tag{7}
\]

Thus, each classifier of \( C \) can be considered as a mapping process from feature vector to measurement vector. For classifier

\[
C_j : F_j \rightarrow Y_j \tag{8}
\]

the measurement level output \( Y \) may be estimates of posteriori probabilities, similarities or distances to the \( m \) prototypes or templates. Then, each individual classifier of \( C \) can make a decision as following

For classifier \( C_j \), if

\[
Y_{j\rho} = \min_{k} (y_{jk})
\]

\[
X \rightarrow S_j \tag{9}
\]

2.2 Adaptive Combination Scheme
For classifiers mentioned above, the commonly used combination method is linear combination model[3]. The combination function can be described as follows
Then, the decision rule can be described as
\[ C_i = \min_k (C_k) \quad k = 1, 2, \cdots, n \] (11)
\[ X \rightarrow S_i \]

In this model, the coefficient \( \beta \) is usually constant, which restricts its adaptability and performance.

Suppose that \( \Phi = \{ \phi_1, \phi_2, \cdots, \phi_P \} \) is the set of all Chinese characters that may exist in video sequences in certain application, \( i = 1, 2, \cdots, P \), where \( p \) is the total number of the Chinese characters in video sequences.

For classifier \( C_i \) and in \( \Phi = \{ \phi_1, \phi_2, \cdots, \phi_P \} \)
Define
\[ \eta_j = \max_j (\| p - \phi_j \|) \quad j = 1, 2, \cdots, p \] (12)

where \( \| \| \) is the measurement of \( \phi_i \) and \( \phi_j \), the measurement may be estimates of posteriori probabilities, similarities or distances, \( \eta_j \) means the maximum distance or the minimum similarity in \( \Phi \), then
\[ \beta_i = \| p - \phi_i \| / \eta_i \] (13)

The coefficient \( \beta_i \) varies with the input pattern.

3 Experiments

To justify the approach, this approach has been applied in the system of the automatic recognition of the licenses of automobiles. In the system, there are 59 Chinese character classes with 786 training and 1578 test cases. Each samples is a 48×96 binary image which has been zoomed.

Fig.2 shows some examples of Chinese characters. Due to the bad illumination, heavy fog, rain, or another reason, these Chinese characters are incomplete in shape and have some parts missing after they have been segmented from the video sequences, but these characters are still recognized correctly. We can see that this approach is promising for Chinese character recognition.

![Fig.2](image)

Fig.2 The recognition result of some characters of automobile license

References


Brief Introductions to Authors

WANG Fei (王飞) was born in 1972. He is now a doctor candidate at UESTC. His research interests include: multimedia processing and communications, image recognition and coding technologies.

LI Zai-ming (李在铭) was born in 1939. He is now a Professor and Doctor Supervisor at UESTC. His research interests include: communication theory, signal detection and estimation, multimedia information processing and communications.