# Statistical Structure Modeling and Optimal Combined Strategy based Chinese Components Recognition

Bowen Yu, Xiaohui Liang, Jiajia Hu, Linjia Sun
*State Key Laboratory of Virtual Reality Technology and Systems*
*Beihang University*
*Beijing, China*
*Email: {yubw, lxh, hjj, sunlinjia}@vrlab.buaa.edu.cn*

***Abstract*—Extracting perceptually meaningful components plays an essential role in Chinese character studying process. This paper proposes an improved statistical structure modeling method to pick up all meaningful components in one character. Each stroke is represented by the distribution of the feature points both in model component and input character. The stroke relations are effectively reflected by the statistical dependency. The mutual information among strokes can be calculated to measure the importance of relationships. Considering the local features of components' difference from the whole character recognition, this paper proposes a method based on local feature to select local components rather than the whole character. At last, we adopt optimal combined strategy to select the best component recognition result. By this method, all the components in one character can be achieved.**

***Keywords*-Chinese component recognition; statistical structure modeling; neighbor selection; optimal combined strategy**

## I. Introduction

Chinese writing, as the representative of ideographic systems, is not based on the irreducible elements used in speaking (syllabic or alphabetic) like phonetic system. In Chinese information processing, each character is treated as a separate symbol, which brings a major problem of mass encoding set. According to Structure Study of Chinese Ideography, the elements of Chinese writing are called components which brought in certain meanings when they were used to construct characters. For example, those characters like "桃", "李", which have a common component "木", are all used to record a name of "tree". Therefore, extracting perceptually meaningful components plays an essential role in the study of Chinese writing as well as the information processing of Chinese characters.

The Chinese recognition problem includes on-line recognition and off-line recognition. In on-line recognition, the trajectories of pen-tip movements and pen-up/down switching information are recorded for recognition. On the other hand, for off-line character recognition, the only available information is the input image. Comparing to the wide use of on-line recognition, off-line recognition is still a difficult problem.We have to extract necessary features from the input character image for recognition. In this paper, we only refer to the knowledge of off-line character recognition.

Generally, the recognizer represents and analyzes the Chinese character image by statistical or structural methods. The statistical methods usually process the image to form a feature vector so that their similarity can be calculated by the distance between two vectors. On the other hand, the structural methods extract the structure information from the image which the relationship between strokes are emphasized. The structural recognizer regards the relationship between strokes as the most stable feature of one character.

Various statistical methodologies have been proposed for character recognition, such as k-nearest-neighborhood classifer[1], K-Means clustering and Gaussian distribution selector[2], nonlinear active shape models[3], contextual vector quantization[4] and Mahanalobis distance[5].

In structural character recognition method, character are often decomposed into parts to analyze. For instance, the part of characters can be contours, strokes or hierarchical models. The structural matching method is approximately equal to stroke matching.

As the basic component of Chinese character, the description of stroke is very important in structural matching. However, the natural strokes are difficult to extract in off-line character recognition. Instead of the natural strokes, straight line segments are more frequently used as the structural primitive. Also the line between feature points including end points and junction points is another form of stroke[6] .

In order to extract more information in stroke, the strokes are categorized into several types, such as horizontal, vertical, slash, back-slash, dot and hook[7]. In these methodologies, all the strokes are classified into the predefined stroke types where each type has different tolerances of orientation, angles and length. However, this method depends heavily on the heuristic knowledge because the process of manually design.

Another essential element in stroke modeling is relationship between strokes. Many researchers regard the relationship between strokes as the most important structural information because stroke relationship is more stable than stroke itself. The relation information can also distinguish two characters which are very similar in shape.

There are both advantages and disadvantages in two meth-

978-0-7695-4911-8/12 $26.00 © 2012 IEEE
DOI 10.1109/SITIS.2012.43

ods. The statistical methods such as k-nearest-neighborhood classifier, K-Means clustering and Gaussian distribution selector and contextual vector quantization, they all perform better when process the image with noises. However, they reflect the character structure indirectly, thus have difficulties to distinguish characters with similar shapes such as "王" and "玉". However, the structure methods are more adaptable on the similar characters. In this method, characters are decomposed into parts to analyze. The component of characters can be contour, strokes, and hierarchical models. Most of structural matching methods adopt strokes as component, structural matching is approximately equal to stroke matching so far.

For the special application such as component recognition and segmentation, they both can't solve the recognition problem perfectly[8] . The statistical methods tend to focus on the global features, while the component recognition always needs local features. And the structure methods can't deal with nonstandard characters because of the existence of unexpected stroke segment.

However, after the appearance of the methods combined with statistical and structural methods is proposed, the defects mentioned above could be fixed. It mainly uses various kinds of statistical methodologies to describe the structure of Chinese characters in the images. In the special application of component recognition, the combined methods could achieve better performance. Compared with pure statistical method, the structure of Chinese character is emphasized to achieve better recognition rate. In comparison with pure structure method , the combined method's statistical modeling of strokes have more strong resistance to noise which allows us to process not only the regular character images[9].

In combined method, the stroke is represented by the distributions of its length, position and angle[10]. The model stroke is represented by the joint distribution of its feature points, while the character is represented by the joint distribution of the component strokes. The position and the shape of strokes are statistically modeled and their distributions are estimated from training samples.This statistical modeling is adopted to represent both statistical and structural information of characters. Subsequently the MRF-based methods are proposed by Jia Zeng[11], which makes use of the energy of cliques to measure the similarity between the target character and the model character. It provides a much more enriched vocabulary to describe character structures. The difference between different characters can be emphasized by the special design of neighborhood system and clique potentials.

We mainly focus on transferring the combined methods from Chinese character recognition to Chinese component recognition. By combined method, the strokes in input character could be modeled to reflect both statistical and structural information, then the recombination of input strokes could realize the target of component recognition.

There are three main contributions in our work:

1) We improved the combined method, which separates the feature extraction into two parts. This improved method could emphasize the local feature for component recognition and be regardless of the same component in different position of different characters.

2) Sampling the feature points of the stroke to reduce the influence caused by different scale of input character.

3) Adopting the optimal combined strategy to select the best component recognition result.

## II. Chinese Component Modeling

### A. Statistical Structure Modeling

In statistical structure modeling method, the component and the whole character are both represented by strokes. The stroke of the model components is manually labeled for the component recognition in character images. The labeling process is to manually combine the sub-strokes extracted by the initial extraction algorithm. The strokes are assumed to obey the multivariate normal joint distribution $X \sim N(\mu, \Sigma)$, and their attributes could be represented by attributes of points belong to them, where we use Gabor filter to measure direction information of the these points[12]. Given the input character image S, the formula 1 is utilized to measure the similarity between S and the model component C. The similarity can be measured by the joint probability of all matching instances between $s_i$ and $r_i$, where $s_i$ represents the strokes in input character and $r_i$ represents the strokes in model component.

$$Pr(S = C) \equiv Pr(s_1 = r_1, s_2 = r_2, ..., s_n = r_n). \tag{1}$$

However, recognizing the character with joint distribution of all component strokes is too complicated. To solve this problem, the conditional probability is applied.

A multivariate normal joint distribution $X \sim N(\mu, \Sigma)$ is composed of sub-distributions $X_A \sim N(\mu_A, \Sigma_A)$ and $X_B \sim N(\mu_B, \Sigma_B)$ as follows:

$$X = \begin{bmatrix} X_A \\ X_B \end{bmatrix}, \Sigma = \begin{bmatrix} \Sigma_{AA} & \Sigma_{AB} \\ \Sigma_{BA} & \Sigma_{BB} \end{bmatrix}. \tag{2}$$

The parameters of conditional distribution $Pr(X_B|X_A = x_A)$ are computed by the following formula:

$$\mu_{B|A} = \mu_B + \Sigma_{BA}\Sigma_{AA}^{-1}(x_A - \mu_A). \tag{3}$$

$$\Sigma_{B|A} = \Sigma_{BB} - \Sigma_{BA}\Sigma_{AA}^{-1}\Sigma_{AB}. \tag{4}$$

Then the matching probability $Pr(s_1 = r_1, s_2 = r_2)$ can be calculated by conditional probability.

### B. Neighbor Selection and Local Feature Extraction

The joint distribution contains all information about the strokes and their relationships without the consideration of their importance. In fact, reflecting all relationships is not necessary both in time complexity and the proper accuracy.

Therefore, we tend to select the necessary relationship to simplify the relationship model.

The joint probability can be transformed to every conditional probability multiplied together, which affords us to calculate the joint probability by the conditional probability. The equation is as followed:

$$\begin{aligned} Pr(S) = Pr(s_1, s_2, ..., s_n) &= Pr(s_1)Pr(s_2, ..., s_n|s_1) \\ &= ... = \prod_{i=1}^{n} Pr(s_i|s_1, s_2, ..., s_{i-1}). \end{aligned} \tag{5}$$

Using the equation above, we can calculate the joint probability by conditional probability of each stroke. However, we have to calculate the dependencies from all preceding strokes in order to get one conditional probability. The neighborhood system can simplify this problem by selecting some important relationship to calculate the conditional probability. This approximation is acceptable both on accuracy and time cost. The approximated equation is shown below:

$$Pr(S) \approx \prod_{i=1}^{n} Pr(s_i|nei(s_i)). \tag{6}$$

The joint distribution contains all information of the strokes and their relationships. Owing to the complexity of reflecting all mutual relationships, we tend to select the most important relationship by minimizing the Kullback-Leibler measure to simplify the relationship model. The details of derivation process can be found in reference [10].

Only one stroke is selected as the neighbor of each stroke. According to the specialty of the data, the $M(s_i; s_j)$ tends to be positive if two strokes are tending to be parallel. To the contrary, the $M(s_i; s_j)$ tends to be negative if two strokes are tending to be vertical. Also, the longer one stroke is, the more stable feature it has, and we finally choose the strokes which have maximum absolute value as the neighbor (see Fig. 1).

Here, we put more concentration on the extraction of the local feature. We decided to calculate the bounding box of two strokes which are neighbors to each other. Then the position relation, the relative length ratio are all extracted, as shown in Fig. 2. These features could effectively reflect the local characters of component, which allows us to better extract the Chinese component in follow-up work.

At last, the individual and neighbor information every stroke has in one Chinese character component are stored in database, where the database could be normal ”txt” or binary file and so on.

## III. Chinese Components Recognition System

### A. Candidate Stroke Extraction

First, we use skeletonizing algorithm to thin the images. Second, the cross points and the end points are both extracted for searching the basic strokes. When the component images in database are processed, the manual interaction can be adopted to combine some strokes in order to reduce the number of model strokes. On the process of character image which is input by users to recognize the components, the basic strokes should be combined into candidate strokes when we try to find the proper component. The algorithm is shown in algorithm 1. The array called record is used to write down the details about which strokes the current combination of stroke is consisted of. After the algorithm 1 is carried out, all candidate strokes to the proper component are produced. For example, when we input the Chinese character “吧”, comparing with the component “巴”, the candidate stroke extraction results are shown in Fig. 3.

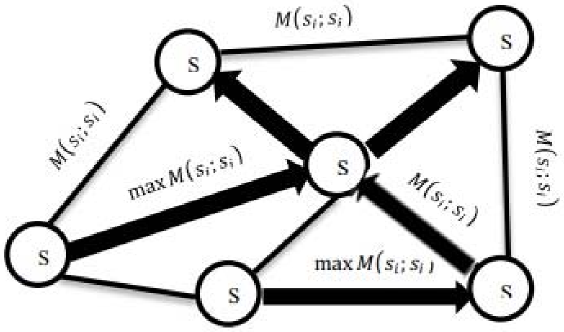


Figure 1. Finding optimal set of the neighbor relationships.

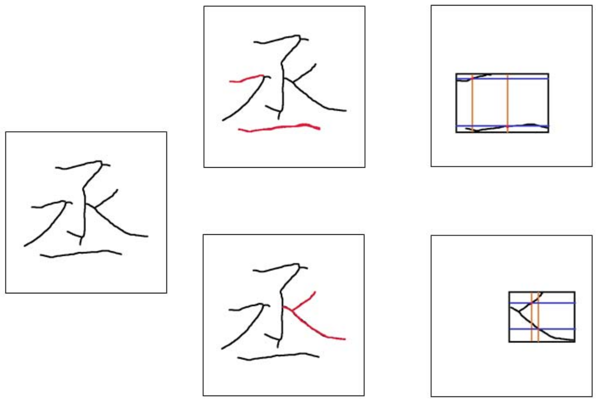

Figure 2. Local features extraction between neighbor strokes.

Because of the difference both in shape and scale between the single component and the component in Chinese character, using the traditional calculating method can not achieve preferable accuracy. In order to deal with this problem, we should sample the stroke data. After the Gabor filter is adopted, we firstly decide to select key points in the stroke like end points, corner points and cross points. Then we further sample the points between the key points. The equal number of points are sampled between key points compared to the component in database. Referring to the stroke of the

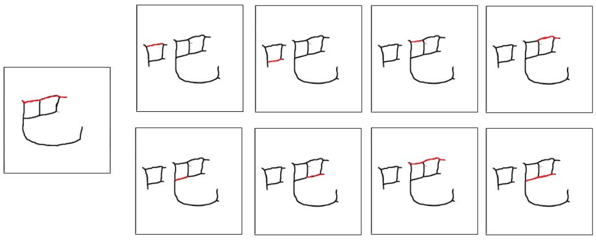

Figure 3. Candidate stroke extraction.

**Input**: All basic strokes detected by the stroke extraction algorithm.
**Output**: The candidate strokes in character image to one model stroke in Chinese component.

Initialization:
  We keep input strokes which are either similar in direction or rather small in open list.
  Build up one vector to record the information of combination.
Extension:
**for** $i = 0$ to the size of similar strokes **do**
  record[i] = MAX; j = i+1;
  **while** record[i]!=-1 **do**
    **if** stroke i,j are connected and $j < size$ **then**
      Combine them and push the new stroke into candidate strokes;
      Update the stroke i, then record j in record[i+1];
      i = i + 1;
    **else if** $j < size$ **then**
      j = j + 1;
      Continue;
    **end if**
    i = i - 1;
    j = record[i] + 1;
    record[the number of combination/2 + i-1]=-1;
  **end while**
**end for**

**Algorithm 1:** Candidate Stroke Extraction

model component, we could calculate the similarity between the sample data and the model stroke. By sampling the data, we reduce the differences caused by different scales as much as possible. In our experiment, we find that this method could produce better results than original one.

### *B. Matching Algorithm*

The whole stage is divided into two sub-stages, single component matching and all components matching. The first problem can be solved by heuristic search algorithm. The aim of single component matching is to find out the best set of strokes in input character to every model component in

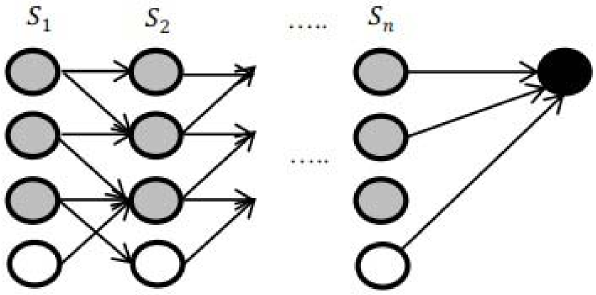


Figure 4. Finding optimal matching.

database. The second problem is categorized as the variant knapsack that demands us to find the best combination of the component found by the first stage, which could fulfill the knapsack as many as possible without stroke usage collision. This is the optimal combined strategy we mentioned above.

*1) Single Component Matching:* The heuristic search algorithm is applied to match the strokes of character to every stroke of the model component. The matching graph should be firstly built up by the model strokes in component. Since the longer strokes represent the more stable structure information, they are sequenced in front. The strokes which are the neighbor of several strokes in component is also put in front.

In this stage, a matching result is defined by a set of correspondences between the model strokes and their matching instances. The matching stage is illustrated in Fig. 4. The $S_i$ represents the model strokes in the model component. The gray nodes represent the candidate stroke which is calculated by the algorithm 1. The white ones show that the optimal candidate stroke $s_i$ for model stroke $r_i$ is not found. The last black node records the final result of this matching process, which the component will be selected if the result is greater than current best record, otherwise, discarded. At last, for one model component, we could find the best set of strokes in input character which indicates the combination of these strokes are most similar to the certain component in database. The algorithm is shown in algorithm 2, which the record[i] here illustrates the selection in candidate strokes generated by algorithm 1 for the certain component.

During the stroke matching process, for example, in stage $S_i$, if stroke i's neighbor is matched before, then the conditional probability is calculated instead of the common probability. Otherwise, the common probability is used if the neighbor of $S_i$ is not matched. Two input strokes $s_i$ and $s_j$, whose matching goal are $r_i$ and $r_j$, respectively, are calculated by conditional probability if they satisfy the following conditions:

(1) $s_i$ is already matched with $r_i$.
(2) $r_i$ is the neighbor to $r_j$.

```
Input: Candidate strokes to all strokes of the model
       component.
Output: The best set of strokes that match the model
        component.
Procedure:
while the path are not all found do
    if candidate stroke j is selected as the model
    stroke i and i < size of model stroke then
        i=i+1;
        j=0;
    else if i=0 then
        break;
    else
        i=i-1;
        j=record[i]+1;
    end if
end while
select best path which has the maximal similarity.
```

**Algorithm 2:** Searching Algorithm

```
Input: Every match result of model component.
Output: The components which consist of the
        character in input image.
Procedure:
for j = 1 to size of input strokes do
    if j > c[n] then
        f(1, j) = v[j];
    else
        f(1, j) = 0;
    end if
end for
for i = 1 to n do
    for j = 1 to size of input strokes do
        if j >= c[i] then
            f(i, j) = max
            [f(i − 1, j), f(i − 1, j − c[i]) + v[i]];
        else
            f(i, j) = f(i − 1, j);
        end if
    end for
end for
```

**Algorithm 3:** Component Recognition Algorithm

Otherwise, the probability in stage j is calculated without the conditional probability.

Here, we prefer to consider not only the conditional probability but also the local feature extracted in Section 2. The local feature should be emphasized when the graph is searched for the component. To emphasize the local feature, we should utilize the ratio of relative position and length calculated in the training stage. Both the local features and the conditional probability are applied to promote the performance of component recognition.

*2) All Components Matching:* Here, we modify the matching result in last section, in order to deal with the Chinese character like “磊”, “炎”, which has reduplicate component. If there is another result which has absolutely different strokes usage that is also similar to the model component in one searching process, for example, the “木” exists in “森”, not only one searching result produced in stroke matching is saved.

In fact, the single component matching is prepared for this subsection. After the single component matching stage, we could obtain possible components composed of many kinds of combination of the strokes in input characters to different component. The different combination results we all record have different input strokes usage. We define the best recognition result as a set of possible components, which the possible components are different in input strokes usage, as well, minimize the unselected strokes of strokes in input characters. We regard this problem as variant knapsack problem. The knapsack's size is equal to the size of strokes in input character. Every matching result has the attribute to describe the usage of input strokes. Then the problem can be classified into knapsack problem, which we aim to find the best combination in those result produced by the single matching problem to fulfill the knapsack without stroke usage collision as many as possible. The algorithm is shown in algorithm 3, where $f(i, j)$ indicates the best value under the knapsack condition j, when component i to n can be selected. And c,v indicates stroke usage and the value of stroke usage respectively. The ">" and "<" in algorithm 3 means the knapsack could afford matching result i without collision to other existed strokes in knapsack or not. Through the solution of variant knapsack problem, we could finally obtain the best component combination in input character and the aim of component recognition is able to be completed.

## IV. Experimental Results

In this section, we will introduce our experiment in five sub-parts. The first subsection will introduce the process of sampling data, in order to reduce the difference caused by different scales. The content in second subsection includes the neighbor selection and conditional probability. The third subsection shows the result of component recognition. The forth and fifth parts will give out the contrast experiment and statistical data respectively.

### A. Sample the Data

As mentioned above, we define the stroke as the distribution of feature points. Here, the feature points are extracted from all points in one stroke. In order to reduce the influence caused by the different scales, we decide to sample the points both in training stage and recognizing stage. The end points,

Figure 5. The feature points in sampling process.

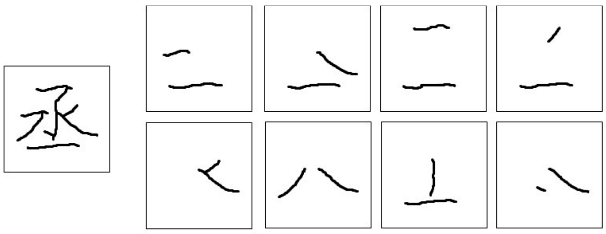

Figure 6. The result of neighbor selection.

crossing points and corner points are extracted firstly, then sample the data between these points according to the ratio of length. The whole sample stage complies with the rule that ensure the equal number of point data. The details are shown in Fig. 5 .

In the example “七” shown in Fig. 5. We firstly extract the end points, crossing points and corner points. At the base of these points, we further sample the points between these special points. The same number of sample points should be extracted in the same character with different size. The sampling process could obviously promote the accuracy rate in single stroke matching.

### *B. Neighbor system and local feature*

The example of neighbor selection are shown in Fig. 6. The picture shown all presents the most possible neighbor by the calculation illustrated in section 3. This process is applied in the training stage, where the component are manually combined and the neighbors are needed to be decided. The relationship is either tending to be parallel or vertical; otherwise the neighbor stroke’s length is rather long which indicates it is more stable. Particularly, it successfully detects high-level relationships, which makes one ideal foundation for the recognition in next part.

In Fig. 7, the conditional probability of model stroke is visualized to test whether the statistical dependency could reflect the stroke relationships well. First, we select one stroke with two neighbors by the methods illustrated above. Then the user draws one stroke according to the model, which makes some changes at the location, length or angle,

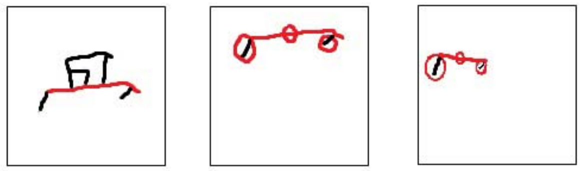

Figure 7. Conditional probabilities given the certain neighbors.

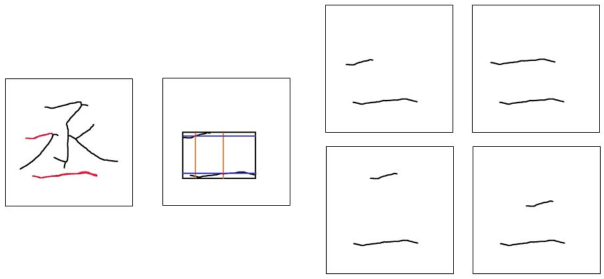

Figure 8. Stroke matching using local features.

to test whether the stroke can be matched to the right model stroke.

The local features are extracted between the strokes and their neighbors. We record the ratio of length and relative position for the component recognition.When the matching stage is processing, we calculate the relative ratio of length and position to measure the similarity .The local features are utilized in Fig. 8. Compared with the first picture in last four pictures, the relative length ratio, the center position and the center distance are not matched so well, respectively. The import of this kind of local feature could emphasize more on the relation between strokes in one component. Finally, we see that the probability we choose the first one as the matching result is much more than others.

### *C. Component Recognition*

The optimal combined strategy based recognition process are shown in Fig. 9. Here, the character “人”, “土” and “田” are collided in input strokes usage, so they could not be selected together. As the result, the optimal combined strategy makes the best result of “田” and “木”.

The model-driven component recognition are shown in Fig. 10.The first column shows the original character image. The second column shows the result after image thinning process, where the strokes are all extracted for the candidate stroke selection. The third and fourth column show recombination of the candidate stroke driven by the model component to realize the purpose of component recognition and segmentation. Then the target components are extracted by this method. Even the structure information can be further estimated by the distribution of the components which are found by the algorithm.

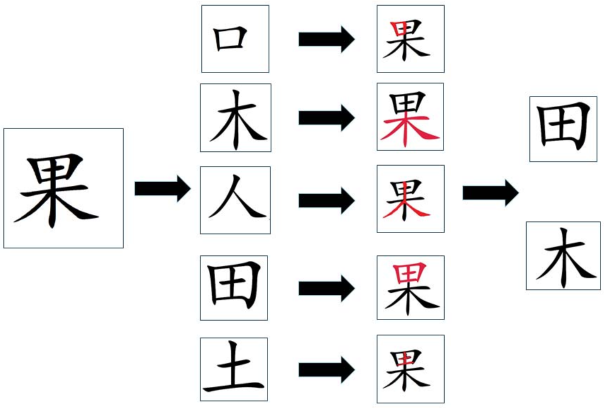

Figure 9. Optimal combined strategy based component recognition.

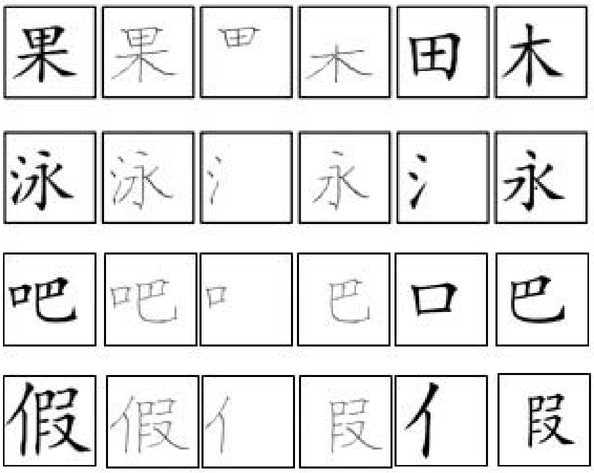

Figure 10. The results of component recognition.

## D. Contrast Experiment

We firstly compare our method with some other efforts we have tried. By comparison, the method based on the connected area appears to have over-segment problem, which always divides the character into fragmentary pieces. The recombination of these pieces is both time-consuming and difficult to find the correct result. The more single areas there are, the harder the recombination is. The Viterbi algorithm segmentation using HMM search for path from one side to the other, could segment the character into different parts. After the initial segmentation, the whole character recognition method could be applied to recognize the component. But the threshold is rather difficult to determine. Different threshold causes different results, as shown in Fig. 11. The comparison results show that this kind of method does heavily depend on the segmentation results, where our method does not. And the segmentation is built up on the

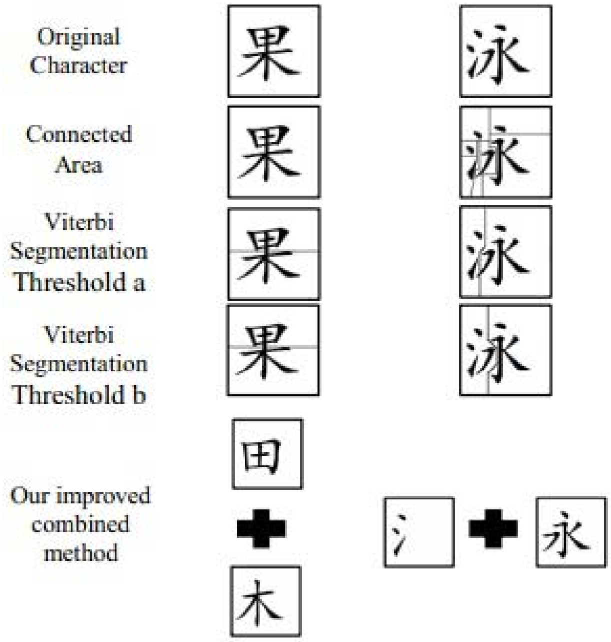


Figure 11. Comparison with the methods which recognize components after segmentation.

prior knowledge about the character's structure (left and right, up and down, etc.), which decides the segmentation is vertical, horizontal or others. On the one hand, the statistical structure method combined with dynamic window is not affordable on the complexity. On the other hand, it could not deal with the special structure called sub-surround structure and surround structure such as “病” and “国”.

From the experiment shows above, the method which firstly segment and then recognize can not produce satisfactory result all the time. Our method deals with the segmentation and recognition at the same time, which could produce rather preferable result.

Furthermore, we decide to compare our proposed method with some representative other works on radical recognition. The comparison result are shown in table I.

*1) Method 1:* Using the statistical structure modeling and optimal combined strategy based method described above, experiments were conducted on 100, 200 radicals covering 500 and 1000 Chinese characters, named experiment 1 and 2 respectively. Here, we only has single copy for each component regardless of their different position in different characters in database.

*2) Method 2:* Using the nonlinear PCA approach. The experiments were conducted on the same test set as for Method 1. Here we stress this method should train large number of components, so that it could recognize the same component in different position of different characters. For instance, the same component “木” has different position in “杉”, “沐”, “呆” and “杏”. Our method could avoid

Table I
THE CONTRAST EXPERIMENT OF COMPONENT RECOGNITION

| | Experiment 1 | Experiment 2 |
|---|---|---|
| Method 1(Our Method) | 96.8 | 92.6 |
| Method 2(Nonlinear ASM) | 96.4 | 89.3 |
| Method 3(Stroke-Based) | 92.6 | 86.4 |

this by the statistical structure modeling, which we only have one copy for one component without the consideration of its different position in different characters, the whole recognition process allows us to recognize the component in different position of different characters.

*3) Method 3:* The stroke-based approach of Wang and Fan[13]. Their method's performance depends heavily on the segmentation result in first stage, which made their method hard do with the characters like "果" and "栽" which have the components attached between two components. However, our method skips the segmentation process, we utilize the recombination of strokes to form the most similar set to form certain component. Sequently, we use the optimal combined strategy to select the best combination result, which at the same time, the components in character are segmented.

## V. CONCLUSION

In this paper, a systematic statistical structure modeling method is improved. Combined with utilization of optimal combined strategy, we could deal with the problem of component recognition in Chinese characters. Both the characters and the components in image are represented by a set of strokes. Through matching strokes to instances using structure information, we could find the best matching result for every model component. Then solving the variant knapsack problem by optimal combined strategy, we could find the best combination of components for the input character. Then the components recognition results are extracted. Furthermore, the structural information could be achieved by structure model matching.

The experiment shows the method is acceptable in Chinese component extraction. But there are also problems. Some special characters like "森" and "品" demand us to search for more than one matching solution in the single matching stage, which makes the problem more complex. Also, the irregular character images should be put into experiment to show the robustness of our method. Optimizing the process of graph searching and improving experiment including more irregular character images will be our future work.

## ACKNOWLEDGMENT

This work was supported by the Beijing Natural Science Foundation(Researches on Human Body Segmentation Methods in Natural Environment based on Computer Vision, Grant No.4112032), the Natural Science Foundation of China (Grant No.61170186).

## REFERENCES

[1] Cheng Lin Liu and M.Nakagawa, *Prototype Learning Algorithms for Nearest Neighbor Classifier with Application to Handwritten Character Recognition*. Pattern Recgnition, 2001.

[2] Y.Y. Tang, L. T. Tu, J.Liu, S. W. Lee, and W. W. Lin, *Off-line recognition of Chinese Handwriting by Multifeature and Multi-level Classification*. IEEE Trans.Pattern Analysis and Machine Intelligence, 1998.

[3] D.Shi, S.R. Gunn, and R.I. Damper, *Handwritten Chinese Radical Recognition Using Nonlinear Active Shape Models*. IEEE Trans.Pattern Analysis and Machine Intelligence, 2003.

[4] P. K. Wong and C. Chan, *Off-Line Handwritten Chinese Character Recognition as a Compound Bayes Decision Problem*. IEEE Trans.Pattern Analysis and Machine Intelligence, 1998.

[5] N.Kato, M.Suzuki, S.Omachi, H.Aso and Y.Nemoto, *A Handwritten Chinese Character Recognition System Using Directional Element Feature and Asymmetric Mahalanobis Distance*. IEEE Trans. Pattern Analysis and Machine Intelligence, 1999.

[6] H.J. Lee, B. Chen, *Recognition of handwritten Chinese characters via short segments*. Pattern Recognition, 1992.

[7] Cheng Lin Liu, In Jung Kim and Jin H. Kim, *Model-Based Stroke Extraction and Matching for Handwritten Chinese Character Recognition*. Pattern Recognition, 2001.

[8] D. Shi, S.R. Gunn and R.I. Damper, *Offline Handwritten Chinese Character Recognition by Radical Decomposition*. ACM Transactions on Asian Language Information Processing, 2003.

[9] Yanwei Wang, Xiaoqing Ding and Changsong Liu, *MQDF Discriminative Learning Based Offline Handwritten Chinese Character Recognition*. International Conference on Document Analysis and Recognition, 2011.

[10] In Jung Kim and Jin Hyung Kim, *Statistical Character Structure Modeling and Its Application to Handwritten Chinese Character Recognition*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003.

[11] Jia Zeng, Wei Feng,Lei Xie and Zhi Qiang Liu, *Cascade Markov Random Fields for Stroke Extraction of Chinese Character*. Information Sciences, 2010.

[12] Yih Ming Su, Jhing Fa Wang, *A Novel Stroke Extraction Method for Chinese Characters using Gabor Filters*. Pattern Recognition, 2003.

[13] A.B. Wang and K.C. Fan, *Optical Recognition of Handwritten Chinese Characters by Hierarchical Radical Matching Method*. Pattern Recognition, 2001.