I.J. Information Technology and Computer Science, **2009, 1, 41-48**
Published Online October 2009 in MECS (http://www.mecs-press.org/)

Hierarchical Matching for Chinese Calligraphic Retrieval Using Skeleton Similarity

Jie Chen
Computer School of Hubei University of Technology , Wuhan , China
Email: cjjjack@163.com

Fuxi Zhu
Computer School of Wuhan Universoty, Wuhan, China
Email: fxzhu@public.wh.hb.cn

***Abstract*—Individual Chinese characters are identified mainly by their skeleton structure instead of texture or color.In this paper, an approach based on skeleton similarity for Chinese calligraphic characters retrieval is proposed.By this approach,first,the skeleton of the binarized individual characters are acquired by an improved multi-level module analysis algorithm.Second,the first round of skeleton matching based on the invariant moment-descriptor guarantees the recall rate;the second round of skeleton matching based on the comprehensive characteristic difference in the polar coordinates system guarantees the retrieval precision.Finally,different styles of the same Chinese characters are ranked and displayed according to the two rounds of matching score.Besides,the efficiency of our approach is manifested by the preliminary experiment.**

***Index Terms*—hierarchical matching, calligraphic retrieval, skeleton similarity**

1. Introduction

With more than 2000 years' evolution, the calligraphy has many treasure works in different historical periods as a special traditional Chinese art form. And the problem of calligraphic retrieval in e-library and other on-line applications has not been resolved well yet, because the different calligraphic styles of the same Chinese character vary so much. The current retrieval method is based upon the artificial index with the information about the author, the caption, the period and the description of the works content. Since the calligraphic characters are different with the standard characters, the usual method like OCR can not handle the retrieval problem. A novel method based on the calligraphic content and character shape feature is needed badly. However the special structure features of the calligraphic characters (like abnormal stroke junction, end stroke subtraction and other art styles) make the retrieval problem even more difficult.

In the recent 20 years, various offline hand-written characters recognition methods has improved in some degrees, which mainly deal with the characters of the alphabetical systems [1][2][3].The center of gravity as the calligraphic retrieval feature is proposed to index the ancient Chinese works[4], but the computation cost of the algorithm is intolerable in the real-time retrieval query in the web application. Shape similarity is also proposed as a key feature in calligraphic retrieval[5], but the precision is not high enough because the contour of characters is the only considered shape feature.

Chinese character is shaped by the constructure-related stokes, and the hand-written Chinese characters retrieval is reduced to the skeletons recognition. Calligraphic characters as an aesthetic medium contain the sentiment of the author, which makes the recognition more difficult than the usual hand-written characters. The skeletons of calligraphic characters can be acquired in most cases as the rudiment of the first-round retrieval to reduce the globe-searching space and enhance the recall rate. And the second-round retrieval based upon the character regional shape features can guarantee the satisfying retrieval precision.

The remainder of this paper is organized as follows: Section 2 gives an effective method of skeleton extraction. Section 3 describes the two-round hierarchical matching for Chinese calligraphic retrieval. Section 4 presents the experiments and evaluation.

Manuscript received Febuary 13, 2009; revised July 25, 2009; accepted September 15, 2009.
No. 2006CB303000, No.2007CB316505, corresponding author:Peng Yali .

Conclusion and future work are given in the final section.

2. Precise skeleton structure of calligraphic characters

2.1 Basical Structure of Chinese Calligraphic characters

Various models are used in the constructional analysis of Chinese calligraphic characters, which shows good performance in the new artistic style generation and simulation [20-23]. For the calligraphic retrieval process, the hierarchical structure is useful to improve the accuracy ratio and recall ratio

There are usually four hierarchical levels for the basic Chinese characters shape features: basic strokes, complicated strokes, stroke identification symbol and character. The first level of basic strokes contains five members: vertical line, horizontal line, two kinds of oblique lines and dot. The three other upper levels are the combination of the lower level ones. That means the complicated strokes are the combination of several basic strokes, and the storks identification symbols are the combination of the basic strokes and the complicated strokes, and any Chinese character is composed by the basic strokes, complicated strokes and strokes identification symbols. Here, the relevant structural information of the various basic stokes is enough for the description of the shape features of any Chinese calligraphic character.

Five kinds of basic Chinese character strokes have their unique shape features, which is relevantly invarible. The hierarchical model is used as the normal description of the shape features. According to the human visional conception system, the skeleton is a good descriptor of the basic strokes direction and other shape features, which can also be used in the stroke order information extraction.

Fig 1: Five kinds of basic Chinese character strokes

2.2 Calligraphic image pre-processing

Various calligraphic works are scanned, digitized and stored in the data-base with the image format, which has not any artificial index of the works content information. The deform strokes are so common in the calligraphic characters (like convolution, transformation and contortion) that the traditional horizontal and vertical projection [6] is not effective in the individual character segmentation process. Successive projection [7] and other necessary manual adjustments are adopted. The Gauss noise suppressor is used to eliminate most back-ground noise and inner white noise of the calligraphic image. The local optimal threshold algorithm based on the canny operator [8] manifests excellent performance in the calligraphic binary image conversion. All binary images in the data-base are normalized into 128*128 pixels. Figure 1 illustrates the pre-processing result of the calligraphic image of the character "Yong".

Fig 2: a Chinese character binarized image

2.3 Acquisition of the precise skeleton of calligraphic character

According to the knowledge of the human vision perception system, the skeleton illustrates the stroke direction and other shape feature of the Chinese character, which provides the fundamental stroke-order information. Skeleton has three geometric features: continuity ,one-pixel width and centralization[9].The traditional skeleton extraction algorithm based on morphology used in the hand-written character recognition appears noneffective to the calligraphic images because of the vast noise and the unique aesthetic styles. Thru following difficulties are obvious:

1) The distinct stroke deforming in the stroke-junction locate;

2) The traditional method can not guarantee the one-pixel width of the skeleton;

3) The skeleton deviates from the stroke contour center;

4) The plausible short branches of the skeleton disturb the retrieval and reduce the retrieval precision;

The hierarchical template analysis method is used to extract the skeleton of the calligraphic characters [10].In this approach, firstly the important structural junction pixels of the inner characters' contour are remained to maintain the continuity of the skeleton. Secondly, the short branches of the skeletons should be removed by the standard of the pixels' border distance which manifests the nearest distance between one skeleton pixel to the background pixels, because in so most cases the short branches would be inevitable. Lastly, some measures will adopt to guarantee the one-pixel width of the skeleton.

The process of skeleton extraction mainly contains five kind of data structure, which has three hierarchical templates parameters, one distance boundary distance, one skeleton crossing pixels counter. In the binary images of the calligraphic characters, the foreground pixels of the characters have the gray value 1, while the background pixels have the gray value 0.Following are the five kinds of data structure signification:

S4(p)-------The sum of gray value of four neighboring specific pixels of the target foreground pixel (p) of the calligraphic characters image, which contains four directions: up, down, left and right.

S8(p)------In the template of 3*3,whose center is the target foreground pixel (p) of the calligraphic characters image, the sum of the gray value of eight surrounding pixels.

S16(p)------ In the template of 5*5,whose center is the foreground pixel (p) of the calligraphic characters image, the sum of the gray value of 16 periphery surrounding pixels.

C(p)------ In the template of 3*3, whose center is the target foreground pixel (p) of the calligraphic characters image, the sum of pixels which accord with the following equation.

$$g(P)-g(Nk)=1 \qquad (1)$$

(g(p) is the target pixel gray value, g(Nk) is the kth pixel gray value of the periphery pixels in the template.)

D(p)------The nearest Euclidean distance between the target foreground pixel of character image to the background pixels.

Following are the main steps of the skeleton extraction process:

1. Keep the foreground pixels, important structural pixels and the extrusive pixels of the stokes, if they have one of the following features:

If S4(p)=4;
If C(p)>1;
If S4(p)=1,S8(p)=1 and S16(p)=1;
p(x,y) is the gray value of the target pixel with the coordinates (x,y),if p(x-1,y)=0,p(x+1,y)=0,and p(x+2,y)=0;
If p(x,y-1)=0,p(x,y+1)=1 and p(x,y+2)=0;

2. In order to remove the short branches in the skeleton, the skeleton pixels which have one of the following features should be picked out. If their D(p) is less than a certain threshold value, they should be removed.

If S4(p)=3,S8(p)>3 and S16(p)>3;
If C(p)=3,S8(p)>3 and S16(p)>3;
If S8(p)=3 and S16(p)=3;

3. In order to maintain the skeleton one-pixel in width, the target foreground pixel p(x,y) in the skeleton which has the following feature should be removed:

If p(x,y-1)=1 and p(x-1,y)=1;
If p(x,y-1)=1 and p(x+1,y)=1;
If p(x,y-1)=1 and p(x-1,y)=1;
If p(x,y+1)=1 and p(x-1,y)=1;
If p(x,y+1)=1 and p(x+1,y)=1；
If p(x,y-1)=1,p(x-1,y)=1 and p(x+1,y)=1;
If p(x,y-1)=1,p(x-1,y)=1 and p(x,y+1)=1;
If p(x,y+1)=1,p(x-1,y)=1 and p(x+1,y)=1;
If p(x,y+1)=1,p(x-1,y)=1 and p(x-1,y)=1;

The Figure2 shows the comparison of two skeletons of the same character after removal of the short branches.

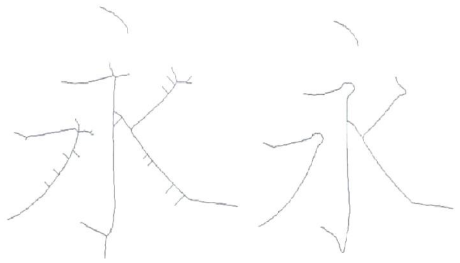

Fig 3: comparison of raw skeleton and precise skeleton with short branchs removed

3. Comparison of the calligraphic character skeleton similarity

The shape features of the Chinese characters are more indicative than the color and texture features, and the stroke structure is the most proper descriptor of the shape features. Five fundamental strokes all have directional feature. The stroke "Heng" is horizontal, the stroke "Shu" is vertical, the stroke "Pie" and "Na" are inclined in 45 degrees and 135 degrees. The computation of the stroke direction is relatively easy because of the single-pixel width of the skeleton. The following figure3 demonstrates a character and its skeleton in the polar coordinates system, and the origin of the coordinates is the center of the character and its skeleton.

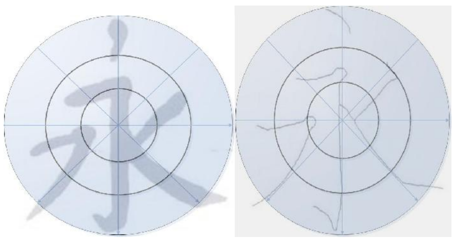

Fig 4: structure of a Chinese character and its skeleton in polarized coordinate system

3.1 Skeleton similarity based on the invariable moment-descriptor

Considering the shape feature invariability of the deforming skeleton in revolving, zooming, transformation, and extrusion, the seven invariable moment-descriptors are used as the feature vector in the first round of coarse retrieval. For the binary calligraphic image f(x,y),the (p+q) matrix descriptor is defined as:

$$m_{pq} = \sum \sum x^p y^q f(x,y)$$

The central (p+q) matrix descriptor is defined as :

$$u_{pq}=\Sigma\ \Sigma\ (x-x_0)^p(y-y_0)^q f(x,y)$$

The point (x_0,y_0) is the gravity center of the binary calligraphic image.

The unified central matrix descriptor is defined as:

$$t_{pq}= u_{pq}\ /\ u^{(p+q+1)}_{00}$$

The seven invarible moment-descriptor are following as:

$$C1= t_{20}+ t_{02}$$

$$C2=(t_{20}-t_{02})^2+4\ t^2_{11}$$

$$C3=(t_{30}-3t_{12})^2+(3t_{21}-t_{03})^2$$

$$C4=(t_{30}+t_{12})^2+(t_{21}+t_{03})^2$$

$$C5=(t_{30}-3t_{12})(t_{30}+t_{12})[(t_{30}+t_{12})^2-3(t_{21}+t_{03})^2]+(3t_{21}-t_{03})(t_{21}+t_{03})[3(t_{30}+t_{12})^2-(t_{21}+t_{03})^2]$$

$$C6=(t_{20}-t_{02})[(t_{30}+t_{12})^2-(t_{21}+t_{03})^2]+4\ t_{11}(t_{30}+t_{12})(t_{21}+t_{03})$$

$$C7=(3t_{21}-t_{03})(t_{30}+t_{12})[(t_{30}+t_{12})^2-3(t_{21}+t_{03})^2]+(3t_{12}-t_{30})(t_{21}+t_{03})[3(t_{30}+t_{12})^2-(t_{21}+t_{03})^2]$$

Due to the tremendous value difference of the seven moment-descriptors, some scale normalization measure is necessary to acquire the weight concord in the feature vector, which enable the quick and proper calculation of the different skeletons similarity distance. The Gauss scale-normalization method is a good option in this application, because the effect of the noise in the collected data is stemmed due to the logical speculation of the noise probability distribution.

After the Gauss scale-normalization, any moment-descriptor value is mapped to the standard Gauss probability distribution, most feature value will in range [-1,1] with probability more than 99%.The feature value out of the range will be set to -1 or 1,so as to maintain the normative value. The skeleton images with characteristic distance to the sample skeleton larger than a certain threshold value g should be removed before the second-round of precise retrieval.

The first round of coarse retrieval based on the invariable moment-descriptor can guarantee the recall rate, which means the retrieval result contains correct samples as more as possible, if the standard difference threshold value g is set relatively larger.

3.2 Skeleton similarity comparison in the polar coordinates system

According to the algorithm proposed in the [12],some improvements are made here to increase the second-round retrieval precision,which means incorrect samples are included in the final retrieval result as little as possible. As illustrated in the figure3, the calligraphic skeleton was put in the polar coordinates system, which is divided into 12 even sectors. At the same time, the polar coordinates system is segmented by 4 concentric circles. Thus, the whole coordinates system is divided into 48 parts, which is similar to the radar scanner. The distribution difference of every skeleton's pixels in the segmented region is an indicator of the possibility whether two skeletons belong to the same character. Here skeleton A has (Ca)i pixels in the ith radar region, and skeleton B has (Cb)i pixels in the ith radar region. Some characteristic values are proposed to assess the pixels distribution difference between two skeletons:

1) (Xai,Yai) and (Xbi,Ybi) are the coordinates of two skeletons barycenter in the radar region qi, and the definition of barycenter difference Gi is the following formula :

$$Gi=[(Xai-Xbi)^2+(Yai-Ybi)^2]^{1/2}/Max[(X^2ai+Y^2ai)^{1/2},(X^2bi+Y^2bi)^{1/2}] \quad (2)$$

Gi demonstrates the Euclidean distance of the two skeletons barycenter.

2) Skeleton A and skeleton B have Nai and Nbi pixels in the radar region qi separately, and the definition of the least mean distance between skeleton A and skeleton B is Dab in the following formula:

$$Dab=\sum da(xk,yk)/Nai \quad (3)$$

da(xk,yk) is the distance between pixel (xk,yk) of the skeleton A and the nearest pixel of skeleton B.The Dba is the least mean distance between skeleton B and skeleton A in the following formula:

$$Dba=\sum db(xk,yk)/Nbi \quad (4)$$

According to formula 3 and 4,the Di is the mutual least mean distance of skeleton A and skeleton B defined in the following formula:

$$Di=(Dab*Nai+Dba*Nbi)/(Nai+Nbi) \quad (0<i<49) \quad (5)$$

Di is the characteristic of the extent of two skeleton superposition.

3) Diff_near is the variance of the Di defined in the following formula:

$$Diff\_near=(Nai*\sum(da(xk,yk)-Dab)^2+Nbi*\sum(db(xk,yk)-Dba)^2)^{1/2}/(Nai+Nbi) \quad (6)$$

4) Vab is the statistical mean distance between skeleton A and skeleton B in the radar region qi defined in the following formula:

$$Vab=\sum dis(xk,yk)/Nbi \quad (7)$$

$\sum dis(xk,yk)$ is the sum of distance between the pixel (xk,yk) of the skeleton A to every pixel of the skeleton B in the radar region qi. Accordingly, the definition of Vba is in the following formula:

$$Vba=\sum dis(xk,yk)/Nai \quad (8)$$

According to the formula 7 and 8,the mutual statistical mean distance between skeleton A and skeleton B is Vi defined in the following formula:

$$Vi=(Vab*Nbi+Vba*Nai)/(Nai+Nbi) \quad (9)$$

Vi is also an indicator of the extent of two skeletons superposition.

5) Diff_stat is the mean square deviation of Vab and Vba in the following formula:

Diff_stat=(Nai*∑(dis(xk,yk)-Vab)^2+Nbi*∑(dis(xk,yk)-Vba)^2)$^{1/2}$/(Nai+Nbi) (10)

6) The definition of the directional parameter of skeleton A and B in the radar region qi is in the following formula:

Fa=∑fx(xk,yk)/(Nai*Nai)

Fb=∑fy(xk,yk)/(Nbi*Nbi) (11)

fx(xk,yk) is the angular degree between certain pixel of skeleton A to its barycenter, and fy(xk,yk) is defined to the skeleton B.And the angular degree is measured toward the positive direction of the horizontal coordinate axis. The value of angular degree is in the range （0°,180°）.

According to the formula (11), the mutual mean directional parameter of skeleton A and skeleton B is defined in the following formula:

Fi=(Fa*Nai+Fb*Nbi)/(Nai+Nbi) (12)

Every radar region has a characteristic vector (Gi,Di.Diff_near_i,Vi,Diff_stat_i,Fi),(0<i<49).The Gauss normalization is used to calculate the characteristic distance between skeleton A and B. Four characteristics have the same weight, but the every region has different characteristic distance weight inverse proportional to its distance to the coordinate origin. The computation cost of the proposed algorithm is O(n^2),which has some obvious improvement on the base of algorithm in [5],because the skeleton image has much less pixels than the binary calligraphy image.

3.2 Fussy Support Vector Machine used in the calligraphic images retrieval

Support vector machine(SVM)[13-19] has been used in the characters images retrieval for at least 10 years, which shows good performance in printed characters processing. For the hand-written ones SVM has poor performance, especially for the Chinese calligraphic images. Since the Chinese characters vary tremendously in shape, texture and other features, the normal SVM can not decide which group the calligraphic character belongs to. And the result of such retrieval is wrong at some extent. Sometimes the denial of retrieval is a trouble due to the inherent fault of normal SVM. The fundamental ideal of SVM is to transit the n groups classification problem to the n two-group classification problem, which leads to the sample can be classified into two or more group. But such cases can not be tolerated in SVM processing. Here the Fussy Support Vector machine (FSVM) is used to handle the cases. The fundamental ideal of FSVM is the relevant decision function. The relevant decision function of i group and j group is defined as :

$$D_{ij}(x) = w^t{}_{ij} + b_{ij}$$

Here $D_{ij}(x)= -D_{ji}(x)$. For the new input sample, the decision function of the i group is defined as:

$$D_i(x) = \sum_{j=i,\, j\neq i}^{n} D_{ij}(x)$$

The condition for the i group classification is following formula:

$$i = \arg\max D_i(x) \quad \text{(i=1…n)}$$

But there are still the samples which can not be classified into any group. In order to handle the problem, the fussy classification function is introduced, which can have the negative value. For the classification panel $D_{ij}(x)=0$, the fussy classification function defined in the direction vertical to the super panel:

$$m_{ij}(x) = \begin{cases} 1 & \text{if } D_{ij}(x) > 1 \\ D_{ij}(x) & \text{otherwise} \end{cases}$$

The ith fussy classification function is defined as :

$$m_i(x) = \min D_{ij}(x) \quad \text{(j=1…n and } j\neq i)$$

The condition for the sample x classified in the ith group is:

$$i = \arg\max m_i(x) \quad \text{(i=1…n)}$$

The introduction of fussy classification of FSVM can handle the denial of classification in the calligraphic retrieval during the skeleton features comparison.

The following figures are several cases which shows the improvement of FSVM and limit of SVM

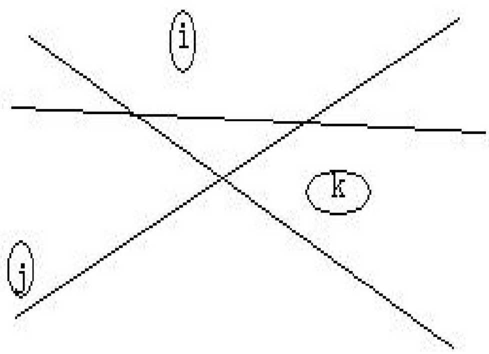

Fig 5: Central region can not be classified in the any group by SVM

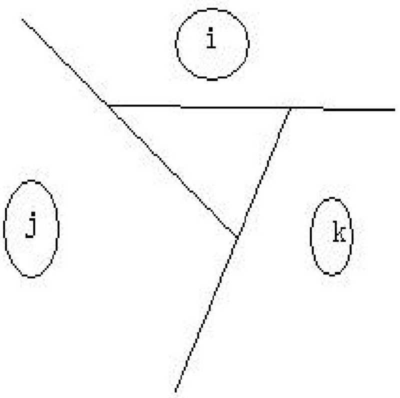

Fig 6: Central region can not be classified in the any group by SVM

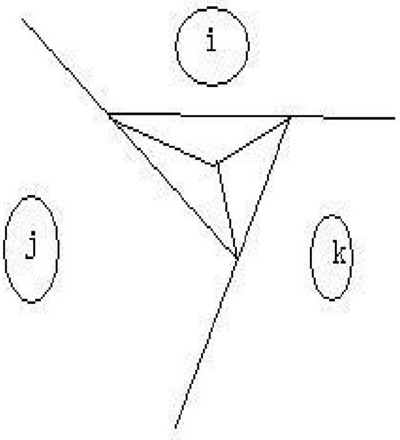

Fig 7: Central region can be classified in the any group by FSVM if fussy classification function is used

4. Experiment and Analysis

After the collection and scanning of famous calligraphic works in various dynasties, there are 2100 calligraphic images with the format of 128*128 pixels in the experimental data-base, which is binarized and segmented. Figure 4 illustrates the two-round of retrieval result of the Chinese character "yong" by the algorithm proposed in this paper, and the previous 15 retrieval images are correct ones, according to the criterion of recall rate ratio and precision ratio:

Recall ratio=|correct samples in the retrieval result sum of retrieval result|/|sum of correct samples in the data-base|

Precision ratio=|correct samples in the retrieval result sum of retrieval result|/|sum of retrieval result

Fig 8: a hand-written Chinese character retrival result in the second round of matching

Fig 9: a printed style Chinese character retrival result in the second round of matching

Further experiment is made in the data-base which includes more than 100 different Chinese characters in more than 5 styles. The retrieval result by our approach is compared with the result by algorithm in [5] and the method based on the traditional multi-direction projection in [7], which shows that our approach maintains higher recall rate and precision in the voluminous retrieval context. Figure 5 illustrates the comparison of the three algorithms

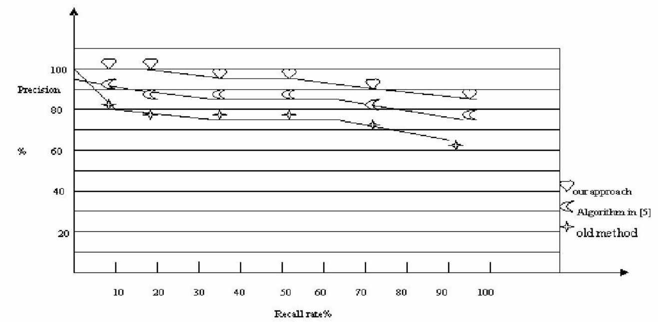

Fig 10: Comparison of recall and precision in three retrieval methods

Table 1 illustrates the comparison of the mean time cost between the three algorithms. Because in our approach the pixels in the skeletons characteristics matching process are about 1/10 of those in the other two algorithms, the mean time cost of our approach is far less than the other two algorithms.

Table 1：Average Time cost of three retrieval method

Algorthim	Projection	Algorthim in [5]	Our approach
Time cost (min)	5.5	2.05	1.42

5. Conclusion

In this paper a novel hierarchical matching algorithm for Chinese calligraphy retrieval based on the skeleton similarity is proposed and manifests good performance in the preliminary experiment. It doesn't need any complexed characters recognition process. The skeletons of various characters in different styles are acquired by an improved method based on the structural analysis, then the invariant moment-descriptors are used in the first round of coarse retrieval to enhance the recall rate, lastly the comprehensive characteristic comparison in the polar coordinates system is used in the second round of precise retrieval to enhance the precision. Since the skeleton contains the precise and concise structural information of the calligraphy content and aesthetic style, and the proposed algorithm doesn't need any training samples, it can guarantee the real-time retrieval application in the online e-library. At the same time, our approach can provide the technique of online calligraphy simulation and appreciation. Further work will be focused on the enlargement of calligraphic data-base, style extraction and stroke order retrieval of different characters. Stroke order information can improve the retrieval efficiency because it contains the kernel information of the characters' structure. The calligraphic retrieval of comprehensive factor including texture, skeleton, boundary and stroke order may be possible in the near future.

References

[1] Palmondon R, Srihari S N. On-line and off-line handwriting recognition:A comprehensive survey[J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2000,22(1):63-84.

[2] Rath T M, Kane S,Lehman A.Indexing for a digital library of George Washington's manuscripts:A study of word matching techniques[R]. Massachusetts: Center for Intelligent Information Retrieval, Computer Science Department, University of Massachusetts,2004

[3] Itay Bar Yosef, Klara Kedem, et al. Classification of Hebrew calligraphic handwriting style[A]. In: Proceedings of the 1st International Workshop on Document Image Analysis for Libraries, Palo Alto, California,2006.299-305

[4] Shi Baile, Zhang Liang, Wang Yong , et al. Content-based Chinese script retrieval through visual similarity criteria[J]. Journal of Software,2001,12(9):1336-1342(in Chinese)

[5] Zhang Xiafen, Zhuang Yueting , et al.Chinese calligraphic chracter retrieval based on shape similarity[J]. Journal of Computer-Aid Design & Computer Graphics,2005,17(11):2565-2569)

[6]Wu Youshou, Ding Xiaoqing. Chinese Chracter Recognition: Theory, Approach and Implementation[M].Beijing: High Education Press,1992(in Chinese)

[7]Zhang KuangZhong . Chinese Character Recognition Technology[M].Beijing: Tsinghua University Press,1992(in Chinese)

[8] John Canny. A computational approach to edge detection[J]. In: IEEE transactions Pattern Analysis and Machine Intelligence,1986,11(2),678-698.

[9] Blum H. A transformation for extracting new description of shape[A]. In Wathen-Dunn W ed. Model for the Perception of Speech and Visual[C].Cambiride, Massachusetts: MIT Press,1967.362-380

[10] Zhao chunjinag, Shi wenkang. A robust algorithm for distilling the skeletons of images[J]. Computaer Appliacation,2005,6(1),1305-1306(in Chinese)

[11] Wan Hualin, Morshed U, Hu Hong, etal. Texture feature and its application in CBIR[J]. Journal of Computer-Aid Design & Computer Graphics, 2003,15(2):195-199(in Chinese) products of Bessel functions," *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955.

[12] Haili Chui, Anad, Rangrarajan. A new point matching algorithm for non-rigid registration[J]. Computer Vision and Image Understanding archive,2003,89(2):114-14

[13] L u Y, Zhang H J , Yin L W , et al. Joint semantic and feature based image retrieval using relevance feedback [J]. IEEE Transaction on Multimedia, 2003, 5 (3) : 3392347.

[14] Inoue T, A be S. Fuzzy support vector machines for pattern classification[C]. In: Proceedings of International Joint Conference on N eural Networks (IJCNN ' 01) , July 2001, 2: 1449-1454.

[15] A be S, Inoue T. Fuzzy support vector machines for multiclass problems [C]. In: Proceedings of the Tenth European Symposium on Artificial Neural Networks (ESANN " 2002) , Bruges,Belgium, April 2002, 113-118.

[16] Wang Shang fei, Xue J ia, W ang Xifa. Contented-based emotion image retrievalmodel[J]. Computer Science, 2004, 31 (9) : 186-190.

[17] Hsia T C. A note on invariantmoments in image processing[J].IEEE Trans. On SMC, 1981, 11 (12) : 831-834.

[18] Vapnik V N. Statistical learning theory [M]. John W iley &Sons, NewYork, N Y, 1998.

[19] Abe S. Analysis of multiclass support vector machines[C]. In:Proceedings of International Conference on Computational Intelligence fo rModelling Control and Automation (C IMCA ' 2003) ,Vienna, Austria, February 2003, 385-396.

[20] Xu Yang.Chinese Calligraphy Production Method Based on HMM Genetics Analogy[J].Journal of Wuhan University,2008,54(1):85-89.
[21] Xu Song Hua,Lau Francis,Cheung William K,Pan Yun.He.Auto matic generation of artistic Chinese calligraphy.IEEE Intelligent Systems,2005,20(3):32-39
[22] Zhang Jun-song,Yu Jin-hui,Mao Guo-hong,Ye Xiu-zhi.Generating Brush Texture for Cursive Style Calligraphy with Auto regressive and Stratified Sampling[J].Journal of Computer-Aided Design & Computer Graphics, 2007, 19(11): 1399-1403.
[23] Dong Jun,Xu Miao,Pan Yun-he,Ye. Statistic Model-Based Simulation on Calligraphy Creation [J].Chinese Journal of Computers, 2008, 31(7): 7720-7725

Jie. Chen: Post-graduated student for doctor degree for computer science in Wuhan University, major in multi-media and images processing

Fuxi Zhu: Professor of school of computer science in Wuhan University,interested in multi-media and web mining.