Chinese Cursive Script Character Image Retrieval Based On An Integrated Probability Function *

Irwin King[1], Zhong Jin[2], and David Yuk-Ming Chan[1]

[1] Department of Computer Science and Engineering,
The Chinese University of Hong Kong,
Shatin, N.T., Hong Kong
king@cse.cuhk.edu.hk
[2] Department of Computer Science,
Nanjing University of Science and Technology,
Nanjing, People's Republic of China
jinzhong@mail.njust.edu.cn

Abstract. Often in content-based image retrieval, a single image attribute may not have enough discriminative information for retrieval. On the other hand, when multiple features are used, it is hard to determine the suitable weighting factors for various features for optimal retrieval. In this paper, we present an idea of integrated probability function and use it to combine features for Chinese cursive script character image retrieval. A database of 1400 monochromatic images is used. Experimental results show that the proposed system based on Legendre moment feature, Zernike moment feature, and pseudo Zernike moment feature is robust to retrieval deformed images. Using our integrated probability function, ninety-nine percent of the targets are ranked at the top 2 positions.

1 Introduction

The last few years have seen an upsurge of interest in contend-based image retrieval (CBIR) — the selection of images from a collection via features automatically extracted from images themselves [1]-[6]. Chinese calligraphy is invaluable in the history of Chinese civilization. There are five styles of Chinese calligraphy. Among them, the cursive script style usually has a low level of legibility and most closely approaches abstract art. Figure 1 shows a page of Chinese cursive style calligraphy [7]. Therefore, it is challenging to develop a Chinese cursive script character image retrieval system. The goal is not to classify but to rank based on similarity.

The ultimate goal of designing information retrieval systems is to achieve the best possible retrieval performance for the task at hand. This objective traditionally led to the development of different retrieval schemes for any information

* This paper is supported in part by an Earmarked RGC Grant from the Hong Kong Research Grant Council # CUHK4176/97E and # CUHK 4407/99E.

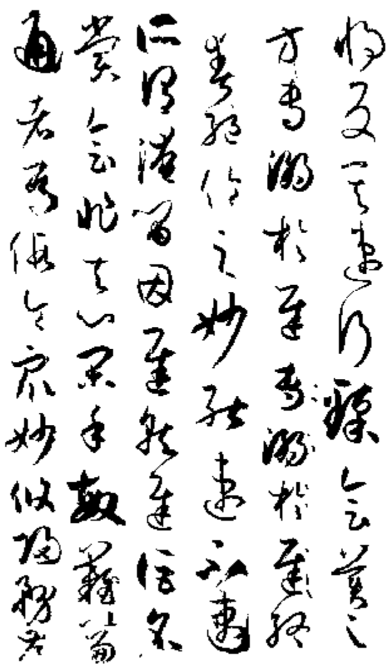

Fig. 1. A page of Chinese cursive style calligraphy

retrieval problem to be solved. The results of an experimental assessment of the different designs would then be the basis for choosing one as a final solution to the problem. It has been observed in such design studies, that different retrieval scheme design potentially offers complementary information which can be harnessed to improve the performance of the selected design.

One idea is not to rely on a single decision making scheme. Instead, all designs, or their subset, are used for decision making by combining their individual opinions to derive a consensus decision. This motivates the relatively recent interest in combining techniques. Cao et al. [8] presented to recognize handwritten numerals with multiple features and multistage classifiers. Kittler et al. [9] developed a common theoretical framework for combining classifiers and demonstrated that the combination rule — the sum rule — outperforms other classifier combination schemes. Recently, Chan and King [6, 10] proposed a weight assignment method in dissimilarity function using genetic algorithm.

This paper presents a combination technique of multi-feature based on the posterior probability estimators. Experiments with 6 kinds of features have been conducted on a database of 1400 Chinese cursive script character images. It is organized as follows. In section 2, a new combination technique is proposed. Section 3 performance experiments are conducted, and experimental results are discussed. Section 4 gives a conclusion.

2 Combination Technique

2.1 Problem Definition

Suppose an image database DB is composed of c distinct images $\{I_1, I_2, \cdots, I_c\}$ and there are K images $\{I_{q1}, I_{q2}, \cdots, I_{qK}\}$ for queries. For a query I_q, image retrieval decision can be made according to the dissimilarity between I_q and any image $I \in DB$. This dissimilarity can be called as a kind of decision function.

Definition 1 (Decision Function) A decision function between two images I and I_q is defined as

$$D(I, I_q) : \Re^{|I|} \times \Re^{|I_q|} \to \Re^1, \tag{1}$$

where $|\cdot|$ indicates the number of elements of a matrix.

Definition 2 (Training Pair) For any query I_{qj}, a training pair TP_j is defined as

$$TP_j = (I_{i(j)}, I_{qj}),\ j = 1, 2, \ldots, K, \tag{2}$$

where $i(j) \in \{1, 2, \cdots, c\}$, and $I_{i(j)} \in DB$ is the best matched image for I_{qj} defined by the user.

Definition 3 (Retrieval Position) Given a training pair $TP_j = (I_{i(j)}, I_{qj})$. The decision function values $\{D(I_i, I_{qj})\}_{i=1}^{i=c}$ can be computed for the given j with equation (1) and ranked according to the minimum rule. The target position of $D(I_{i(j)}, I_{qj})$ in the ranking, which is denoted by N_j, can be called the retrieval position of the training pair $TP_j = (I_{i(j)}, I_{qj})$.

Definition 4 (Average Position) An average position for a training pair set $TP = \{TP_j\}_{j=1}^{j=K}$ can be defined as follows:

$$\overline{N} = \frac{1}{k} \sum_{j=1}^{K} N_j. \tag{3}$$

2.2 Integrated Dissimilarity Function

In general, the dissimilarity of two images I_i and I_{qj} can be determined with their features.

Definition 5 (Feature Extraction) For any image I, a feature extraction function F is defined as

$$F(I) : \Re^{|I|} \rightarrow \Re^d, \tag{4}$$

which extracts a real-valued d-dimensional feature vector.

Definition 6 (Integrated Dissimilarity Function) Assume that there are M feature extraction functions $\{F_i\}_{i=1}^{i=M}$. The decision function $D(I, I_q)$ can be defined as the following integrated dissimilarity function:

$$D(I, I_q) = \frac{\sum_{i=1}^{M} w_i D(F_i(I), F_i(I_q))}{\sum_{i=1}^{M} w_i}, \tag{5}$$

where

$$D(F_i(I), F_i(I_q)) = ||F_i(I) - F_i(I_q)|| \tag{6}$$

is the Euclidean distance between the ith feature vector $F_i(I), F_i(I_q) \in \Re^{d_i}$, and the w_i is the weight assigned to the feature extraction function F_i.

2.3 Integrated Probability Function

Combination techniques in pattern recognition problems have been extensively investigated in recent years. Consider a pattern recognition problem where pattern Z is to be assigned to one of the c possible classes $(\omega_1, \omega_2, \cdots, \omega_c)$. Assume there are M distinct measurement vectors $\{x_i\}_{i=1}^{M}$. Kittler et al. [9] demonstrated that under some assumptions, the following combination rule — the sum rule — outperforms other classifier combination schemes:

$$Z \longrightarrow \omega_j \quad \text{with} \quad \arg\max_j \frac{1}{M} \sum_{i=1}^{M} P(\omega_j | x_i). \tag{7}$$

The most commonly used classifier is the minimum distance classifier.

Minimum Distance Rule Given a measurement vector x for the pattern Z and the representative measurement vectors x^j for the class ω_j $(j = 1, 2, \cdots, c)$, the dissimilarity between two vectors x and x^j can be measured by the Euclidean distance. The minimum decision rule is

$$Z \longrightarrow \omega_j \quad \text{with} \quad \arg\min_j ||x - x^k||, \tag{8}$$

where $|| \cdot ||$ indicates the Euclidean distance.

Generally speaking, the smaller the distance value $||x - x^k||$, the larger the posterior probability $\hat{p}(\omega_k | x)$. Thus, we can present an estimator of the posterior probability $P(\omega_k | x)$ as follows:

$$\hat{p}(\omega_k | x) = \frac{1}{c-1}(1 - \frac{||x - x^k||}{\sum_{j=1}^{c} ||x - x^j||}) \stackrel{\text{def}}{=} \hat{P}(x^k, x). \tag{9}$$

Therefore, it is obvious that the minimum decision rule (8) is equivalent to the following maximum decision rule:

$$Z \longrightarrow \omega_j \quad \text{with} \quad \arg\max_j \hat{P}(x^k, x). \tag{10}$$

Definition 7 (Integrated Probability Function) Assume that there are M feature extraction functions $\{F_i\}_{i=1}^{i=M}$. The decision function $D(I, I_q)$ can be defined as the following integrated probability function:

$$D(I, I_q) = \frac{\sum_{i=1}^{M} w_i \hat{P}(F_i(I), F_i(I_q))}{\sum_{i=1}^{M} w_i}, \tag{11}$$

where the w_i is the weight and $\hat{P}(F_i(I), F_i(I_q))$ is the estimator of the posterior matching probability between image I and image I_q on the feature extraction function F_i, which is determined according to Equation (9).

3 Retrieval Experiments

Our aim is to develop a Chinese cursive script image retrieval system that is insensitive to variations on image deformations. In experiments, we evaluate the performances of six shape features, which are the invariant moments [15], eccentricity [16], edge direction histogram [17, 18], Legendre moments, Zernike moments and pseudo-Zernike moments [19–21]. We also evaluate the performances of the various combination schemes of features according to the frequency of the retrieval positions and the average retrieval position.

3.1 Database

Our Database DB has 1,400 binarized Chinese cursive script character images. Each image is normalized to the size of 200 by 200 pixels. These images were scanned from the book [22] with the resolution of 150dpi. In experiments, we test the behavior of our image retrieval system in the presence of the deformation transformations as shown in Figure 2.

10 images in DB are used to generate a set of 100 deformed images, as shown in Figure 3.

The deformed images were submitted as query images to the retrieval system to examine whether the deformed images can retrieval their original images or not.

3.2 Feature Extracted

For an image, six shape features are extracted. They are listed in Table 1. The first three kinds of features are from [10], and the last three kinds of features are computed on 50×50 low resolution images.

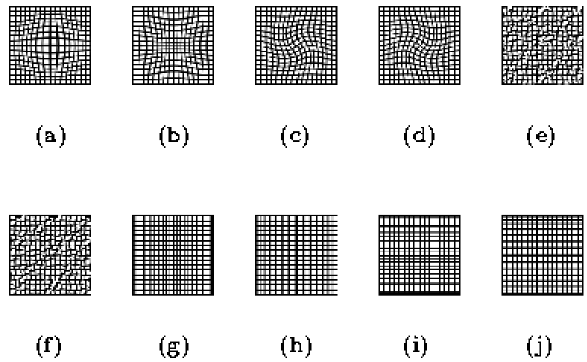

Fig. 2. Distortions include Pinch:(a) and (b); Twirl: (c) and (d); Ripple: (e) and (f); Horizontal Extension: (g) and (h); and Vertical Extension (i) and (j).

Table 1. Six Features for an Image

No.	Feature Description
x1	1-dimensional eccentricity [10]
x2	7-dimensional invariant moment vector
x3	30-dimensional vector of edge direction of the histogram
x4	36-dimensional pseudo Zernike moment vector $\{Z_{nk}(0 \le k \le n \le 7)\}$
x5	30-dimensional Zernike moment vector $\{Z_{nk}(0 \le k \le n \le 8,$ and $n-k=$ even$)\}$
x6	25-dimensional Legendre moment vector $\{\lambda_{mn}(0 \le m, n \le 4)\}$

3.3 Evaluation of Features

In order to evaluate the effectiveness of 6 features for Chinese cursive script character images, for any fixed j $(1 \le j \le 6)$, we can use the single feature x_j to conduct 100 queries with 100 training pairs. After 100 retrieval positions are obtained, the position frequency and the average position can be computed. These experimental results are listed in the Table 2.

From the Table 2, we have the following facts and discussions:

The 1-dimensional eccentricity and the 7-dimensional invariant moments are not effective features for Chinese cursive script character image retrieval since no less than 38% of the targets are ranked behind the top 70 positions. The dimensions of these two kinds of features are too small for our database. The low dimensional features do not have sufficient discriminative information for image retrieval on a large scale database.

Fig. 3. 100 queries deformed from 10 character images for performance evaluation.

The 30-dimensional edge direction of histogram is not an efficient feature for Chinese cursive script character image retrieval as 21% of the targets are ranked behind the top 70 positions. As the histogram of an image is treated as a 1-D discrete signal, it can not have enough discriminative information for image retrieval on a large scale database.

The Legendre moments, Zernike moments, and pseudo Zernike moments are effective features for the Chinese cursive script character images, about 90% of the targets are ranked at the top 3 positions. But they are not very satisfactory as no less than 1% target was ranked after the top 70 positions. Generally speaking, moments can describe the global characteristic of an image. Therefore, high-order moments are required for image retrieval on a large scale database.

3.4 Combination Results

In order to evaluate the effectiveness of the integrated probability function, the following combination schemes are considered and listed in Table 3.

With each scheme, we can conduct 100 queries with 100 training pairs, and get 100 positions, with which the position frequency and the average position can be computed. The experimental results are listed in the Table 4.

From the Table 4, we have the following facts and discussions:

Table 2. Retrieval Results on Single Feature

Frequency of Positions	Features					
	x_1	x_2	x_3	x_4	x_5	x_6
1	2	33	27	89	81	94
2	4	3	7	5	7	2
3	1	5	5	1	3	0
4	0	0	4	0	0	0
5	2	1	1	0	1	1
(5,10]	3	7	7	2	2	1
(10,20]	10	6	7	0	1	0
(20,70]	16	7	21	1	4	1
>70	62	38	21	2	1	1
Average Position	200.65	114.09	62.76	5.64	4.58	4.31

Table 3. Seven Combination Schemes

No.	Scheme Description
C1	combining x_4 with x_5
C2	combining x_4 with x_6
C3	combining x_5 with x_6
C4	combining x_4, x_5 with x_6
C5	combining x_3, x_4, x_5 with x_6
C6	combining x_2, x_3, x_4, x_5 with x_6
C7	combining x_1, x_2, x_3, x_4, x_5 with x_6

With the proposed integrated probability function, the combination of features performs in general better than single feature does in decreasing the average position for retrieval.

Among seven combination schemes, the combination of Legendre moment, Zernike moment, and pseudo Zernike moment outperforms other combination schemes. Ninety-nine percent of the targets were ranked at the top 2 positions, and only one percent target was ranked at the fifth position. The combination of more than these three features cannot increase the effectiveness of image retrieval test. Therefore we proposed a retrieval system based on the combination of Legendre moments, Zernike moments, and pseudo Zernike moments.

The last column of Table 4 lists the retrieval results of [10]. It is obvious that our proposed system performs much better than the existing system.

4 Conclusion and Future Work

This paper presents an idea of using integrated probability function for retrieval of Chinese cursive script character images. It is based on the combination of Legendre moment, Zernike moment, and pseudo Zernike moment. Experiments

Table 4. Retrieval Results on Combination Schemes

<table>
<tr><th rowspan="2">Frequency of Positions</th><th colspan="7">Schemes</th><th rowspan="2">Chan [10]</th></tr>
<tr><th>C1</th><th>C2</th><th>C3</th><th>C4</th><th>C5</th><th>C6</th><th>C7</th></tr>
<tr><td>1</td><td>89</td><td>98</td><td>96</td><td>97</td><td>97</td><td>97</td><td>95</td><td>43</td></tr>
<tr><td>2</td><td>3</td><td>0</td><td>1</td><td>2</td><td>2</td><td>2</td><td>2</td><td rowspan="4">31</td></tr>
<tr><td>3</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>2</td></tr>
<tr><td>4</td><td>2</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td>5</td><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>0</td></tr>
<tr><td>(5,10]</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td rowspan="3">20</td></tr>
<tr><td>(10,20]</td><td>1</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td></tr>
<tr><td>(20,70]</td><td>2</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td>>70</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>6</td></tr>
<tr><td>Average Position</td><td>3.01</td><td>1.18</td><td>1.23</td><td>1.06</td><td>1.06</td><td>1.06</td><td>1.21</td><td></td></tr>
</table>

show that the present system is superior to the existing system [10]. Ninety-nine percent of the targets are ranked at the top 2 positions in a database containing 1,400 images. Further work can be done with the feedback of the weight coefficients.

5 Acknowledgement

This paper is supported in part by a Earmarked RGC Grant from the Hong Kong Research Grant Council #CUHK4407/99E.

References

1. C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, and W. Equitz. Efficient and effective querying by image content. *Journal of Intelligent Information Systems*, 3(3/4):231-262, July 1994.
2. C. P. Lam, J. K. Wu, and B. Mehtre. STAR–A System for trademark archival and retrieval. In *2nd Asian Conf. on Computer Vision*, volume 3, pages 214-217, 1995.
3. Y. S. Kim and W. Y. Kim. Content-based trademark retrieval system using visually salient feature. In *IEEE Computer Society Cnf. on Computer Vision and Pattern Recognition*, pages 307-312, 1997.
4. A. K. Jain and A. Vailaya. Shape-based retrieval: A case study with trademark image databases. *Pattern Recognition*, 31(9):1369-1390, 1998.
5. J. P. Eakins, J. M. Boardman, and M. E. Graham. Similarity retrieval of trademark images. *IEEE Multimedia*, 5(2):53-63, 1998.
6. D. Y. M. Chan and I. King. Genetic algorithm for weights assignment in dissimilarity function for trademark retrieval. In *Third International Conf. on Visual Information and Information Systems(VISUAL'99)*, volume 1614 of *Lecture Notes in Computer Science*, pages 557-565, The Netherlands, 1999. Springer.
7. K.C. Ma. *Shu Pu I Chu / Ma Kuo-Chuan I Chu*. Hong Kong: Shao-Hua Wen Hua Fu Wu She, 1977.

8. J. Cao, M. Ahmadi, and M. Shridhar. Recognition of handwritten numerals with multiple feature and multistage classifier, *Pattern Recognition*, 28(2): 153-160, 1995.
9. J. Kittler, R. Hatef, R. P. W. Dubin, and J. Matas. On combining classifiers. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 20(3):226-239, 1997.
10. D. Y. M. Chan and I. King. Weight assignment in dissimilarity function for Chinese script character image retrieval using genetic algorithm.
11. J. Ashley, R. Barber, M. Flickner, J. Hafner, D. Lee, W. Niblack, and D. Petkovic. Automatic and semiautomatic methods for image annotation and retrieval in QBIC. In *Proceedings of Storage and Retrieval for Image and Video Database III*, volume 2420, pages 24-35, February 1995.
12. J. K. Wu, B. M. Mehtre, Y. J. Gao, Chian-Prong Lam, and A. Desai Narasimhalu. STAR–A multimedia database system for trademark registration. In Witold Litwin and Tore Risch, editors, *Applications of Databases, First International Conf.*, volume 819 of *Lecture Notes in Computer Science*, pages 109-122, Vadstena, Sweden, 21-23 June 1994. Springer.
13. A. Soffer and H. Samet. Using negative shape features for logo similarity matching. In *14th International Conf. on Pattern Recognition*, volume 1, pages 571-573, 1998.
14. F. Mokhtarian, S. Abbasi, and J. Kittler. Efficient and robust retrieval by shape content through curvature scale space. In *First International Workshop on Image Databases and Multi-Media Search*, pages 35-42, Amesterdam, The Ntherlands, August 1996.
15. M. K. Hu. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 8, 1962.
16. B. Jahne. *Digital Image Processing: Concepts, Algorithms and Scientific Applications*. Springer-Verlag, Berlin; New York, 4 edition, 1997.
17. J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679-698, 1986.
18. A. K. Jain and A. Vailaya. Image retrieval using color and shape. *Pattern Recognition*, 29(8): 1233-1244, 1996.
19. S. X. Liao and M. Pawlak. On image analysis by moments. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 18(3): 254-266, 1996.
20. Cho-Huak Teh and R. T Chin. On image analysis by the methods of moments. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 10(4): 496-513, 1988.
21. R. R. Bailey and M. Srinath. Orthogonal moment features for use with parametric and non-parametric classifiers. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 18(4): 389-398, 1996.
22. H. H. Lu. *Han Tai Chien Tu Tsao Tzu Pien*. Shang-hai: Shang-hai Shu Hua Chu Pan She, 1989, 1 edition.