21st International Conference on Pattern Recognition (ICPR 2012)
November 11-15, 2012. Tsukuba, Japan

Discriminative Normalization Method for Handwritten Chinese Character Recognition

Yuanping Zhu
Department of Computer Science,
Tianjin Normal University
zhuyuanping@mail.tjnu.edu.cn

Jun Sun, Satoshi Naoi
Fujitsu R&D Center Co. Ltd.
{sunjun, naoi}@cn.fujitsu.com

Abstract

Comparing with conventional character normalization methods not taking the discriminative information into account, this paper proposes a novel normalization method -- Discriminative Normalization. Saliency regions contain most of discriminative information among similar characters. According to different types, they are enlarged in character normalization to increase their influence in recognition. As a result, discrimination power among similar characters is enhanced which is benefit to separating similar characters. The experiment on CASIA dataset shows that error rate is reduced by 9.97%. Comparing with similar character recognition without discriminative normalization, 46.0% more errors are reduced. That verifies its effectiveness.

1. Introduction

Handwritten Chinese Character Recognition (HCCR) is important but difficult in OCR. The essential reason is that handwritten character samples have greatly different shape distortion which causes smaller compactness in shape distribution. To compensate the shape distortion, character normalization tries to convert character samples to approximate a uniform shape mode. Therefore, character normalization is indispensable for current HCCR system [1]. Many normalization methods have been proposed and shown good performance, such as [2-4].

As we know, discriminative information is crucial for recognition. Making full use of it is necessary. Conventional character normalization methods work in global or character-independent mode. The same criterion is applied for all character samples and discriminative information among characters is never considered. In this paper, combining with discrimination enhancement, a novel normalization method -- Discriminative Normalization (DN) is proposed.

Discriminative information in character recognition is from dissimilarity among characters. Similar characters usually share similar or even same parts. Their discriminative information is not obvious and easy to be disturbed. This causes the difficulty in similar character recognition and brings about most of recognition errors. To utilize discriminative information among similar characters, Suzuki et al. [5] and Gao et al. [6] used Mahalanobis and LDA classifier to recognize similar characters respectively. Namely, they used discriminative information from sub-space. Xu et al. [7] and Leung et al. [8] proposed similar character recognition based on critical region analysis. Two different methods for locating critical region are given by them. Actually, they use critical region to select features with better discrimination.

In Discriminative Normalization, a character is divided into two parts according to similar character, similar and dissimilar part. As Fig. 1 shows, the dissimilar part is named saliency region which contains most of information to separate similar characters. Usually, saliency region occupy smaller area. That causes weak description and insufficient influence of saliency region sometimes. Meanwhile, it brings recognition instability from noises and shape distortion in similar part. So, increasing saliency region's influence in recognition will be benefit to separating similar characters. Observing Fig. 1, precisely locating critical region is difficult, especially in the border region. That is, direct recognition on critical region is not reliable. If enlarges saliency regions and compresses similar regions, it will make two effects: 1) strengthen description of dissimilar part; 2) weaken the influence of similar part. As a result, discrimination power will be enhanced and similar character errors will be reduced. By this means, discrimination enhancement is embedded in normalization. How to locate saliency region and implement character shape transformation are key points. The method details are given in section 2. Section 3 gives experimental results. Finally, section 4 will discuss conclusions.

Fig. 1. Saliency region of similar character pair

2. Discriminative normalization Method

2.1. System Framework

As Fig. 2 shows, the system has two stages: Training and Recognition. There are two key points, saliency region detection and character image transformation.

Training:

- Create similar character table
- Saliency region detection
- Create normalization model dictionary

Each similar character pair has one discriminative normalization model storing transformation parameters.

- Similar character classifier training

Recognition:

- Initial character recognition

Use a recognizer to yield initial recognition result.

- Similar character search.

The first two recognition candidates regarded as a possible similar character pair and search the corresponding pair in similar character table.

- Character image transformation

Perform shape transformation if similar character pair is found. Transformation parameters are stored in normalization models.

- Similar character recognition

A pair-wise classifier for similar character pair is designed and fisher classifier is adopted in this paper.

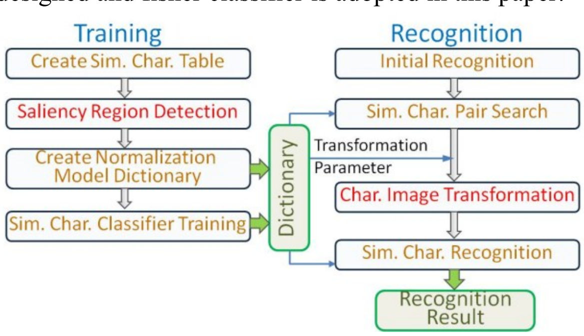

Fig. 2. System framework

2.2. Saliency region detection

Registering similar part is difficult, the dissimilar part as well. In character feature extraction, information is extracted zone by zone from image. 2D geometric information is maintained in feature map. Analyzing feature map can locate approximate dissimilar part. A generalized concept of saliency region based on dissimilar part is constructed. We divide a character in zones and dissimilar part contains some of them. Saliency Region (SR) is a closed region covering main dissimilar zones. There are three main steps.

- Initial character normalization

Global character normalization is used to increase compactness of all characters' data distribution. The line density equalization [3] is adopted.

- Feature extraction

A 392-dimensional feature, WDCH (weighted direction code histogram) [9], is used. In this feature, 8-directional information in 7x7 zones is extracted. For the convenience of saliency region analysis, feature is represented as a tensor $x(i,j,k)$. Where, i and j are the zone labels of horizontal and vertical axes respectively, k is the direction label.

- Saliency region analysis

Saliency region is constructed based on dissimilar zones. We compare the feature tensor templates to find dissimilar zones. A feature difference map between two characters is utilized to measure distribution of dissimilarity in a character image.

Let $T^m(i,j,k)$, $T^n(i,j,k)$ denote tensor templates of character C_m and C_n. In fact, they are character class centers. The difference map is:

$$D^{m,n}(i,j)=\sum_{k=1}^{8}\left\|T^m(i,j,k)-T^n(i,j,k)\right\| \quad (1)$$

Its variance is δ^2. Then, the zones with large difference value are labeled as dissimilar zones.

$$D^{m,n}(i,j)>Th \quad (2)$$

and $Th=\overline{D^{m,n}(i,j)}+\delta\cdot t, t=0.1$ (3)

Fig. 3 gives an example of detecting saliency region of similar pair "阿" and "何". Dissimilar zones are marked as 1. Actually, dissimilar zones are scattered in image and not always continuous. But we don't use dissimilar zones directly and the rough location of dissimilar part is enough. The saliency region is a rectangle covering main dissimilar zones. The most important parameters of SR are position and range. Let m_{10}, m_{00} and m_{01} be moments and μ_{20}, μ_{02} be central moments. The saliency region center is:

$$(cx,cy)=(m_{01}/m_{00},m_{10}/m_{00}) \quad (4)$$

Two axes' lengths are:

$$ax=2\mu_{02}/m_{00},ay=2\mu_{20}/m_{00} \quad (5)$$

According to center position, saliency regions are assigned to 9 types shown in Fig. 4. The arrow of each type denotes the deformation direction. Considering margin effect, type 9's region range is set by axis length of x and y. Let w and h be character width and height. Saliency region ranges are:

$$Type\,1:(rx,ry)=(2\cdot cx,2\cdot cy); \quad (6)$$

$$Type\,2:(rx,ry)=(2\cdot cx,h); \quad (7)$$

$$Type\,3:(rx,ry)=(2\cdot cx,2\cdot(h-cy)); \quad (8)$$

$$Type\,4:(rx,ry)=(w,2\cdot(h-cy)); \quad (9)$$

$$Type\,5:(rx,ry)=(2\cdot(w-cx),2\cdot(h-cy)); \quad (10)$$

$$Type\,6:(rx,ry)=(2\cdot(w-cx),h); \quad (11)$$

$$Type\,7:(rx,ry)=(2\cdot(w-cx),2\cdot cy); \quad (12)$$

$$Type\,8:(rx,ry)=(w,2\cdot cy); \quad (13)$$

$$Type\,9:(rx,ry)=(ax,ay); \quad (14)$$

2.3. Character image transformation

In recognition stage, if similar pair is found, shape transformation will be performed based on coordinate mapping function. Keeping the image size, the mapping function should extend saliency region and compress similar region at the same time. Here, sine function is chosen whose curve is shown in Fig. 5. As the same as saliency region detection, a global normalization should be performed in advance. Let *I(x,y)* and *I'(x',y')* are images before and after transformation. The coordinate mapping functions are:

$$Type\,1:\begin{cases} x'=\sin(\frac{\pi}{2}\alpha_1\cdot x)/\sin(\frac{\pi}{2}\alpha_1), \\ y'=\sin(\frac{\pi}{2}\alpha_2\cdot y)/\sin(\frac{\pi}{2}\alpha_2), \end{cases} \quad (15)$$

$$Type\,2:x'=\sin(\frac{\pi}{2}\alpha_1\cdot x)/\sin(\frac{\pi}{2}\alpha_1) \quad (16)$$

$$Type\,8:y'=\sin(\frac{\pi}{2}\alpha_2\cdot y)/\sin(\frac{\pi}{2}\alpha_2) \quad (17)$$

Type 9:

$$\begin{cases} x'=cx+\frac{x-cx}{|x-cx|}\cdot\frac{1}{\sin(\frac{\pi}{2}\alpha_1)}\cdot\sin(\frac{\pi}{2}\alpha_1\cdot\frac{|x-cx|}{\max\limits_{p_{ij}\in SR}|x_{ij}-cx|}) \\ y'=cy+\frac{y-cy}{|y-cy|}\cdot\frac{1}{\sin(\frac{\pi}{2}\alpha_2)}\cdot\sin(\frac{\pi}{2}\alpha_2\cdot\frac{|y-cy|}{\max\limits_{p_{ij}\in SR}|y_{ij}-cy|}) \end{cases} \quad (18)$$

where p_{ij} is a pixel, α_1, α_2 control deformation extent.

Type 3, 5 and 7 have similar processing with type 1. Type 6 and 4 are mirror types of type 2 and 8.

Observing sine function curve, 1-order derivative of right part changes too fast. This tends to yield serious deformation. To control deformation extent, α is used

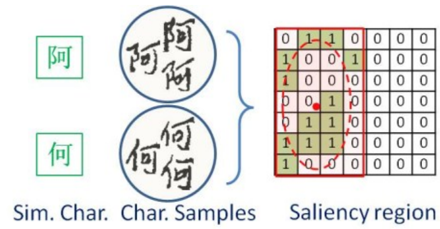

Fig. 3. Saliency region detection (dissimilar zones are marked as 1).

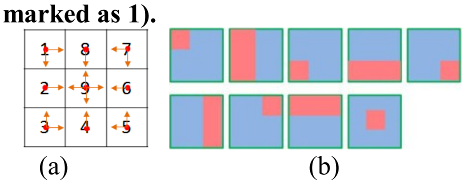

Fig. 4. Saliency region types. (a) 9 types of saliency region assigned by centers; (b) saliency region type 1 to type 9 (from top to bottom and left to right);

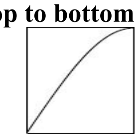

Fig. 5. Sine function, range is regularized to [0, 1].

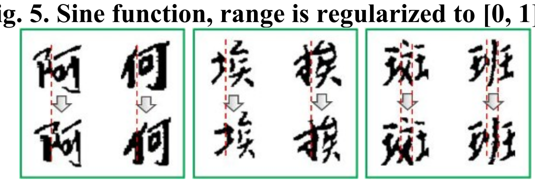

Fig. 6. Several examples of character image shape transformation

to limit the range of sine function used. In shape transformation, the extended room of saliency region is from similar part. Let β indicate the compression extent of similar part. Then, α_1 is:

$$\alpha_1=[w_1+w_2\cdot(1-\beta)]/w \quad (19)$$

where w is character width, w_1 is saliency region width, w_2 is similar region width. β is set to 0.2. α_2 is similar.

Discriminative normalization performs character shape transformation following above strategies. Fig. 6 gives several examples. It is able to see that proportion of dissimilar parts is increased. Dissimilar parts will have more influence in recognition.

3. Experimental results

Evaluation is based on CASIA dataset which was collected by Institute of Automation, Chinese Academy of Sciences and contains 3755 Chinese characters of the GB level-1charset, 300 samples per character. We choose 290 samples per class for training and 10 samples per class for testing.

The HCCR engine used in experiment has similar architecture with Gao's [6]. Feature is 392-dimesional WDCH mentioned in section 2. LDA reduces original feature to 120 dimensions. Dimension reduced feature is fed into MQDF[10] with 10 principle components. Before classification on MQDF, a prototype classifier on 120-dimensional feature outputs 100 recognition candidates. MQDF only runs on 100 coarse candidates to save time. Thus, recognition of MQDF is the initial recognition of DN. In similar character recognition, a pair-wise fisher classifier is adopted. To measure the effectiveness of discriminative normalization, we just compare the result of pair-wise fisher classifier with or without DN in the experiment. The similar character pair is decided by first two recognition candidates (d_0 & d_1) as follows. e is 0.1.

$$|(d_0 - d_1)/d_0| \le e \quad (20)$$

Table 1 and Fig. 7 give comparison data. DN denotes discriminative normalization. The 1st column is the number of similar character pair utilized in experiment. Zero means no similar character recognition and only using conventional normalization, which is baseline. From 2nd to 5th column are results of similar character recognition without or with DN. ERR is error reduction rate. When 5k similar pairs adopted, DN method reduced error rate about 9.97%. The best achievement of without DN is 6.83% when 7k similar pairs used. That is to say, similar character recognition with DN eliminates 46.0% more errors than without DN while fewer similar pairs are used.

4. Conclusions

This paper proposed a novel normalization method -- discriminative normalization. To strengthen the description of dissimilar part and weaken the influence of similar part, saliency region of similar characters is detected and enlarged. By this means, discrimination power is enhanced among similar characters and results in obvious error reduction. In experiment, more errors are reduced than similar character recognition without discriminative normalization. It verifies that DN is effective in enhancing discrimination among similar characters. The future work will focus on more effective shape transformation strategies, especially for dissimilar regions with complicate structure.

Table 1. Error rates on CASIA dataset

#Sim. Char. Pair	Without DN		With DN	
	ER(%)	ERR(%)	ER(%)	ERR(%)
0	2.884	-	2.884	-
1k	2.818	2.31	2.794	3.14
2k	2.751	4.62	2.719	5.72
3k	2.719	5.72	2.668	7.48
4k	2.703	6.28	2.626	8.96
5k	2.695	6.56	**2.597**	**9.97**
6k	2.690	6.74	2.602	9.79
7k	2.687	6.83	2.605	9.70

*DN: discriminative normalization,
*ER: error rate, ERR: error reduction rate

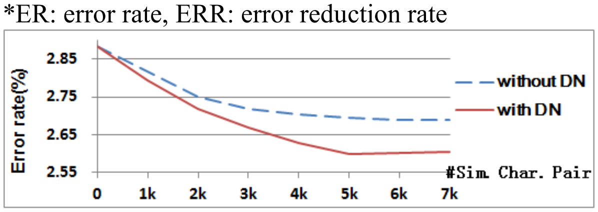

Fig. 7. Error rates curves on CASIA dataset

Acknowledgement: This work was supported by the Foundation of Hunted Talents of Tianjin Normal University (Grant No. 5RL112).

References

[1] C.-L. Liu, "Handwritten Chinese character recognition: Effects of shape normalization and feature extraction," In: Arabic and Chinese Handwriting Recognition, LNCS Vol.4768, Springer, pp.104-128, 2008.

[2] J. Tsukumo and H. Tanaka, "Classification of Handprinted Chinese character Using Non-Linear Normalization and Correlation Methods," In Proc. 9th Int'l Conf. Pattern Recognition, pp.168-171, 1988.

[3] H. Yamada, K. Yamamoto, and T. Saito, "A Nonlinear Normalization Method for Handprinted Kanji Character Recognition - Line Density Equalization," Pattern Recognition, 23(9), pp.1023-1029, 1990.

[4] C.-L. Liu, K. Marukawa, "Pseudo two-dimensional shape normalization method for handwritten Chinese character recognition," Pattern recognition, 38(12): 2242-2255, 2005.

[5] M. Suzuki, S. Ohmachi, N. Kato, H. Aso and Y. Nemoto, "A discrimination method of similar characters using compound Mahalanobis function," Trans IEICE Japan, J80-DII(10):2752–2760, 1997.

[6] T.-F. Gao,C.-L. Liu,"High accuracy handwritten Chinese character recognition using LDA-based compound distances," Pattern Recognition,41(11):3442-3451,2008.

[7] B. Xu, K. Huang, C.-L. Liu, "Similar Handwritten Chinese Characters Recognition by Critical Region Selection Based on Average Symmetric Uncertainty," In Proc. of the 12th ICFHR, pp.527-532, 2010.

[8] K. Leung and C. Leung, "Recognition of handwritten Chinese characters by critical region analysis," Pattern Recognition, 43(3):949–961, 2010.

[9] F. Kimura, T. Wakabayashi, S. Tsuruoka, and Y. Miyake, "Improvement of handwritten Japanese character recognition using weighted direction code histogram," Pattern Recognition, 30(9):1329-1337, 1997.

[10] F. Kimura, K. Takashina, S. Tsuruoka, Y. Miyake, "Modified Quadratic Discriminant Functions and Its Application to Chinese Character Recognition," IEEE Trans. Pattern Anal. Mach. Intell. 9 (1):149-153, 1987.