A Novel Method for Off-line Handwriting-based Writer Identification

Zhenyu He[1], Bin Fang[1,2], Jianwei Du[3], Yuan Yan Tang[1] and Xinge You[1,4]
[1] Department of Computer Science
Hong Kong Baptist University, Hong Kong
Email: zyhe, bfang, yytang, xyou@comp.hkbu.edu.hk
[2] Center for Intelligent Computation of Information
Chongqing University, P.R.China
[3]Department of Mathematics and Physics
Beijing Institute of Petrochemical Technology, P.R.China
Email: dujianwei@bipt.edu.cn
[4] Faculty of Mathematics and Computer Science
Hubei University, P.R.China

Abstract

Handwriting-based writer identification is a hot research topic in the pattern recognition field. Nowadays, on-line handwriting-based writer identification is steadily growing toward its maturity. On the contrary, off-line handwriting-based writer identification still remains as a challenging problem because writing features only can be extracted from the handwriting image in this situation. As a result, plenty of dynamic writing information, which is very valuable for writer identification, is lost. At present, 2-D Gabor filter method is widely acknowledged as a good method for off-line handwriting identification, however it still suffers from some inherent disadvantages, such as the high computational cost. In this paper, we present a novel wavelet-based GGD method to replace the traditional 2-D Gabor filters. Shown in our experiments, this novel method not only achieves better experiment results but also greatly reduces the elapsed time on calculation.

1. Introduction

Even in such a highly developed, modern society, handwriting has continued to play as a main means of communication and recording in people's daily life. Usually, people regard the handwriting as an important personal sign of the writer. Given its ubiquity in human transactions, automatic writer identification based on handwriting has a great significance in many real-world applications, such as postal addresses on envelopes, bank checks, private letters, and etc.

We can classify handwriting-based writer identification in several ways. However, the most straightforward one is to classify it into on-line (also called dynamic) and off-line (also called static) writer identification based on different input methods. The former assumes that a transducer device is connected to the computer, which can convert writing movement into a sequence of signals and then transmit this signal sequence into computer. Since information on the writing order and dynamics of the writing process, which is captured by the transducer device contains many useful writing features of the writer, on-line handwriting-based writer identification, compared with off-line handwriting-based writer identification, is comparatively easier to cope and achieve a higher accuracy. In contrast, off-line handwriting-based writer identification usually deals with handwriting materials directly scanned into computer and thus much dynamic information of writing process is lost. As a result, despite continuous effort, off-line handwriting-based writer identification still remains a highly challenging research issue [1]. Unfortunately, on-line methods of writer identification are inapplicable in many cases where the transducer can not be used, so developing techniques on off-line writer identification is an urgent task.

Further, the off-line writer identification can also be divided into two types: text-dependent and text-independent [1]. Text-dependent methods only match the same character and so require the writers to write the same fixed texts. Contrastively, given that the writing styles of different people are often visually distinctive, text-independent methods do not match the same characters but extract writing style features from the global writing text. In other words, in

Proceedings of the 2005 Eight International Conference on Document Analysis and Recognition (ICDAR'05)
1520-5263/05 $20.00 © 2005 IEEE

the case of text-independent, the handwriting images are considered as different textures. Generally, text-dependent methods have better identification results, however as mentioned above, they are inapplicable in many practical applications because of their strict requirement on same characters. On the contrary, though owning a wider applicability, text-independent methods don't obtain the same high accuracy as text-dependent methods do. In sum, more research works should be done in text-independent writer identification. Our research work focuses on the off-line, text-independent writer identification based on handwriting.

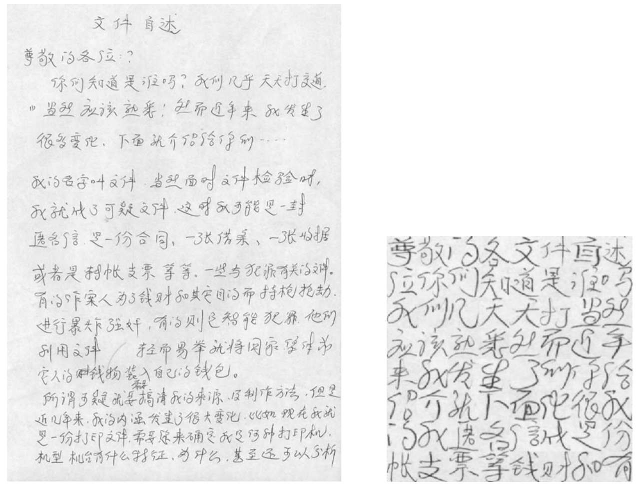
2. Relative work

Commonly, writer identification is regarded as a typical problem of pattern recognition and contains 3 basic steps: pre-processing, feature extraction, feature matching. Nowadays, it is an active research field in academia because it is a combination of traditional pattern recognition with jumped-up biometric authentication. In the last several decades, some researchers have touched this field [2]. But most previous research works are on text-dependent writer identification, especially on signature verification. Even so, there still exist a few papers on text-independent writer identification, in the following we introduce these relative works briefly.

In 1975, Duverony has reported that the most important variation of the writers handwriting is reflected in the low-frequency band of Fourier spectrum of the handwriting images [3]. Similarly, Kuckuck has used Fourier transform techniques to process handwritten text as texture [4]. Later, inspired by the idea of multichannel spatial filtering technique, Said et al have proposed a texture analysis approach for writer identification [1]. In their method, they regard the handwriting as an image containing some special textures and then use a well-established 2-D Gabor filter to extract features of such textures. Except those methods in frequency domain, Schrihari and Cha extracte twelve shape features from the handwriting text lines to represent personal handwriting style. The features mainly contain visible characteristics of the handwriting, such as width, slant and height of the main writing zones [5]. Besides, to improve the accuracy and enhance the robustness, some other papers adopte multiple features integration for writer identification [5][6].

3. Pre-processing

The input image contains characters of different sizes, spaces between text lines and even noises. So before feature extraction, origin image should be processed firstly. Common steps adopted for pre-processing are as follows: firstly, removing the noises in the handwriting image; secondly, locating the text line and separating the single character using projection; thirdly, normalizing each character into a same size; finally, creating the texture image by text padding. In our application, we design a pre-processing method which produces texture image for text-independent feature extraction. Since some papers have discussed pre-processing [1][2], and this problem is not our focus in this paper, here we don't introduce our methods on pre-processing in details. An example of image pre-processing is shown in fig 1.



(a) Original image (b) Texture image

Figure 1. An example of pre-processing

4. Gabor algorithm

In reference [1], a well-designed 2-D Gabor filter is proposed for text-independent writer identification. Besides this paper, reference [2] also applies the same Gabor filter on Chinese text-independent writer identification. Both of the two papers say good results are achieved in their experiments. And the academia also widely acknowledges the Gabor method is an effective method on text-independent writer identification. To display the advantage of our new algorithm, we will compare it with the Gabor method. While firstly, we introduce the Gabor method briefly.

The computational model of the 2-D Gabor filter proposed in reference [1][2] is as follows:

$$h_e(x, y) = g(x, y) \cos[2\pi f(x \cos\theta + y \sin\theta)] \quad (1)$$

$$h_o(x, y) = g(x, y) \sin[2\pi f(x \cos\theta + y \sin\theta)] \quad (2)$$

where h_e and h_o denote the so-called even- and odd- symetric Gabor filter, and $g(x, y)$ is an isotropic Gaussian function.

The spatial frequency responses of the Gabor function are

$$H_e(u, v) = \frac{[H_1(u, v) + H_2(u, v)]}{2} \tag{3}$$

$$H_o(u, v) = \frac{[H_1(u, v) - H_2(u, v)]}{2j} \tag{4}$$

where $j = \sqrt{-1}$ and

$$H_1(u, v) = \exp\{-2\pi^2\sigma^2[(u - f\cos\theta)^2 + (v - f\sin\theta)^2]\}$$

$$H_2(u, v) = \exp\{-2\pi^2\sigma^2[(u + f\cos\theta)^2 + (v + f\cos\theta)^2]\}$$

Here, f, θ, σ are the spatial frequency, orientation, and space constant of the Gabor envelope, separately. For a given input image, $h_e(x, y)$ and $h_o(x, y)$ will combine to provide different channel outputs of the input image with different f, θ and σ.

The mean values(M) and the standard deviation(S) of the channel outputs are used to represent writer global feature for writer identification. If J orientations and L frequencies for each orientation are selected for Gabor filter, a total of J×L features will be obtained from a given handwriting image, as form a feature vector with J×L elements.

After extracting the writing features, Weighted Eucliden Distance(WED) is applied for feature matching.

$$WED(k) = \sum_{i=1}^{N} \frac{(M_i - M_i^k)^2}{\delta_i^k} \tag{5}$$

where M_i denotes the ith feature of the handwriting image whose writer is unknown, M_i^k and δ_i^k denote the ith feature and its standard deviation of the handwriting written by writer K separately, and N denotes the total number of mean values.

5. Wavelet-based GGD algorithm

Though 2-D Gabor filter is effective in handwriting-based writer identification, it still suffers from some inherent disadvantages. One of the most serious disadvantages is the intensive computational cost, because the 2-D Gabor filter has to convolute the whole image for each orientation and each frequency.

Compared with the Gabor filter, 2-D wavelet can decompose the image into sub-bands with different frequencies and orientations. So, we only need to deal with the specified wavelet sub-bands according to the selected values of frequency and orientation.

Wavelet transform is a tool that cuts up data or functions or operators into different frequency components, and then studies each component with a resolution matched to its scale. The basic introduction of wavelet theory has been published in a series of classic papers and books. Here, we only introduce basic concepts of wavelet.

A function $\psi \in L^2(R)$ is called an admissible or basic wavelet if it satisfies the following "admissibility" condition.

$$C_\psi := \int_R \frac{|\widehat{\psi}(w)|^2}{|w|} dw < \infty \tag{6}$$

With translation and dilation of the basic wavelet $\psi(x)$, a 1-D continuous or integrable wavelet transform is defined by

$$(W_\psi f)(a, b) = \int_{-\infty}^{\infty} f(t)\overline{\psi_{a,b}(t)}dt \tag{7}$$

Where $\psi_{a,b}(t) = a^{-\frac{1}{2}}\psi(\frac{x-b}{a})$. If only discrete value of a and b are used, the eq 7 represents the 1-D discrete wavelet transform.

The extension of the 1-D wavelet transform to 2-D wavelet transform is usually achieved by using separable wavelet. An image can be transformed into four sub-images by 2-D separable wavelet transform, namely: (1) LL sub-image: both horizontal and vertical directions have low frequencies. (2) LH sub-image: the horizontal direction has a low frequency, and the vertical one has a high frequency. (3) HL sub-image: the horizontal direction has a high frequency, and the vertical one has a low-frequency. (4) HH sub-image: both horizontal and vertical directions have high frequencies. That is, a handwriting image can be decomposed into sub-bands with different frequencies and orientations via 2-D wavelet transform, and therefore we only need to extract features from specified sub-bands but not from the whole handwriting image. In this way, calculational cost is greatly reduced.

The most sample and direct features of wavelet coefficients are the energies of wavelet sub-bands. Mostly, L^1-norm and L^2-norm are selected as measures. Suppose $\{WI_j(x, n)\}_{x\in R}$ is the collection of wavelet coefficients in the jth wavelet sub-band, then L^1 and L^2 of this sub-band are given as:

$$L_i^1 = \frac{1}{R}\sum_{x\in R}|WL_j(x, n)| \tag{8}$$

$$L_i^2 = (\frac{1}{R}\sum_{x\in R}WL_j(x, n)^2)^{1/2} \tag{9}$$

where, R refers to domain of the jth wavelet sub-band.

But energy is too simple to completely represent the essential features of wavelet coefficients, for example, energy can not reflect statistical distribution of wavelet coefficients. Previous experimental work revealed a particular property of the wavelet coefficients $\{WI_j(x,n)\}_{x\in R}$ of almost every natural textured image at a given scale n and the sub-band number j, that the histogram of these coefficients corresponds to a Generalized Gaussian Density model.

The Generalized Gaussian Density(GGD) model is given as

$$p(x;\alpha,\beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} \exp^{-1(|x|/\alpha)^{\beta}} \tag{10}$$

where $\Gamma(\cdot)$ is the Gamma function, i.e.,

$$\Gamma(\cdot) = \int_0^\infty \exp^{-t} t^{Z-1} dt, Z > 0.$$

The parameter $\alpha > 0$, called scale parameter, describes the standard deviation and $\beta > 0$, called shape parameter, is inversely proportional to the decreasing rate of the peak. The basic idea of GGD model is to use the GGD model to approximate the statistical distribution of the wavelet coefficients in one wavelet sub-band and then take the parameter couple $\{\alpha, \beta\}$ of GGD model as the features to represent wavelet sub-band. There are varied methods to estimate α, β, here we adopt the maximum-likelihood estimator(MLE). The following is how to use MLE for GGD.

The likelihood function of the data vector $x = (x_1, ..., x_L)$ (here we should convert the sub-band image s into a multi-dimensional vector x) having independent component can be defined as

$$L(x;\alpha,\beta) = \log \prod_{i=1}^{L} p(x_i;\alpha,\beta) \tag{11}$$

And using MLE, α, β can be deduced as the roots of following likelyhood equations [9]:

$$\frac{\partial L(x;\alpha,\beta)}{\partial \alpha} = -\frac{L}{\alpha} + \sum_{1}^{L} \frac{\beta |x_i|^{\beta} a^{-\beta}}{\alpha} \tag{12}$$

$$\frac{\partial L(x;\alpha,\beta)}{\partial \beta} = -\frac{L}{\beta} + \frac{L\Psi(1/\beta)}{\beta^2} - \sum_{i=1}^{L} (\frac{|x_i|}{\alpha}) \log(\frac{|x_i|}{\alpha}) \tag{13}$$

where $\Psi(.)$ is the digamma function, i.e. $\Psi(z) = \frac{\Gamma'^{z}}{\Gamma(z)}$. We ignore the deduction process to solve the equations above. For more details, please refer to reference [9].

To replace the typical norm-based distance (e.g. Euclidean distance), we use Kullback-Leibler Distance (KLD) for feature matching. The Kullback-Leibler Distance (KLD) between two sub-bands is as

$$D(p(\cdot;\alpha_1,\beta_1)\|p(\cdot;\alpha_2,\beta_2)) = \log(\frac{\beta_1\alpha\Gamma(1/\beta_2)}{\beta_2\alpha_1\Gamma(1/\beta_1)}) + (\frac{\alpha_1}{\alpha_2})^{\beta^2} \frac{\Gamma((\beta_2+1)/\beta_1)}{\Gamma(1/\beta_1)} - \frac{1}{\beta_1} \tag{14}$$

and the KLD between two handwriting image is the sum of all the distances across all selected wavelet sub-bands.

6. Experiment result

In our experiments, as mentioned above, all handwriting are scanned into computer with a resolution of 300 dpi. Then via certain pre-processing algorithm, we produce the handwriting texture images from the original scanned images, as shown in fig 1. Experiments display the size of handwriting texture image should be suitable, since large size image leads to a high computational cost and small size image reduces the identification accuracy. In our experiment, we select size 512×512 pixels.

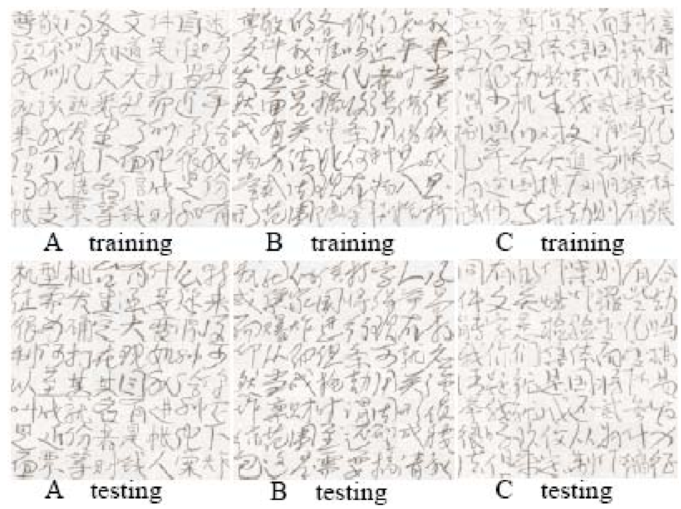


Figure 2. Some samples of texture image used in our experiments. "A Training" refers to training sample of writer A, and "A Testing" refers to testing sample of writer A.

20 Chinese handwritings written by 10 persons have been carried out in the experiments, with one training handwriting and one testing handwriting for each person. We produce one handwriting texture image from each handwriting, and thus a total of 20 handwriting texture image are obtained. The training and testing texture image, whose size is 512×512 pixels, consists of 64 Chinese characters with size 64×64 pixels, as is shown in fig 2. In Gabor method, 4 spatial frequencies are used: 32, 54, 128, 256, and for each spatial frequency, we select 0, 45, 90 and 135 degree as orientations, because both reference [1] and [2] say that the highest accuracy is obtained in this case. In GGD method, we firstly decompose the handwriting image via db4 wavelet transform at 3 levels, and then apply GGD model on the wavelet decomposition sub-bands except the

HH sub-band at the finest scale.

A testing handwriting texture image is matched with all training handwriting texture images. Then we sort the matching results in an ascending order to produce a list. And the position of writer of the testing handwriting in the list is regarded as the experiment result to evaluate algorithm accuracy. (For example, if the matching result between the training handwriting and testing handwriting, both of which are written by the same writer, is the minimum in the list and consequently occupies the position 1, we say the position of real writer is top 1; in other words, the topper the position of one writer is, the more possibility of being the real writer of the testing handwriting the writer has.) The experiment result is in the table 1.

Table 1. IDENTIFICATION ACCURACY OF EXPERIMENT

Method name	Top 1	Top 2
GGD	80%	20%
Gabor	70%	30%

Not only contrast the identification accuracy, we also contrast the average elapsed time of the two methods. We run and record the average elapsed time of GGD and Gabor methods in our computer. The software environment of our computer is: Window XP, Matlab 7.0; and the hardware environment is: Intel Pentium IV 2.66GHZ CPU, 512MB RAM. The experiment result is shown in the table 2.

Table 2. ELAPSED TIME OF THE GGD AND GABOR METHOD

Method name	Average elapsed time (second)
GGD	19.338
Gabor	124.812

7. Conclusion

In this paper, we have presented a novel text-independent method for Chinese handwriting-based writer identification. It shows, based on our experiments, the 2-D Gabor filter is still a good method for handwriting-based writer identification while our new method achieves better results and greatly reduces the elapsed time. Because the text-independent methods do not consider the writing content, the methods involved in this paper are also available for English, Korean, Japanese and Latin Languages, etc.

References

[1] H.E.S.Said, T.N.Tan. Personal identification based on handwriting. *Pattern Recognition*, vol.33, no.1(7):149–160, January 2000.

[2] Y.Zhu and T.Tan. Biometric personal identification based on handwriting. *Proc. 15th International Conference on Pattern Recognition*, pages 801–804, Barcelona, Spain, Sep3-7 2000.

[3] J.Duvernoy. Handwriting synthesis and classification by means of space-variant transform and karhunen-loeve analysis. *J.Opt.Soc.Am.*, 65:1331–1336, 1975.

[4] W.Kuckuck. Writer identifcation by spectra analysis. *Processdings of the international Conference On Security Through Science Engineering*, pages 1–3, West Berling, Gerany, 1980.

[5] S.Cha, S.N.Srihari. Multiple feature integration for writer verification. *the Precddings of 7th IWFHR2000*, pages 333–342, amstredam, Netherland, September 2000.

[6] E.N.Zois, V.anastassopousls. Fusion of correlated decisions for writer verification. *Pattern Recognition*, vol.33, no.10:1821–1829, 1999.

[7] Yasushi Yamazaki, Naohisa Komatsu. A proposal for text-indicated writer verification method. *the 4th International Conference Document Analysis and Recognition*, pages 333–342, Ulm, German, August 18 - 20 1997.

[8] G.Van de Wouwer, P.Scheunders. Statistical texture characterization from discrete wavelet representations. *IEEE transactions on image processing*, 8:4:592–598, 1999.

[9] M.N.Do, M.Vetterli. Wavelet-based texture retrieval using generalized gaussian density and kullback-leibler distance. *IEEE Transactions on Image Processing*, vol. 11:146–158, 2002.

1520-5263/05 $20.00 © 2005 IEEE

IEEE COMPUTER SOCIETY