An Effective Writer Verification Algorithm Using Negative Samples

Xianliang Wang, Xiaoqing Ding

State Key Laboratory of Intelligent Technology and Systems,
Department of Electronic Engineering,
Tsinghua University, Beijing 100084, P. R. China
wxl@ocrserv.ee.tsinghua.edu.cn

Abstract

Writer verification is the process of deciding on the claimed identity of a writer by comparing some attributes of test handwriting with those of his reference handwriting, which has great applications in forensic justification. However, automatic writer verification is very difficult, and at present most writer verification tasks are fulfilled by document examiners, which is very time-consuming and tend to be subjective. In this paper, we proposed an effective writer verification algorithm using negative samples. This algorithm uses handwriting Chinese characters as the object. In the algorithm, the directional element features are first extracted from the handwriting Chinese character scripts, then negative handwriting samples are introduced, and the decision is made based on the combination value of similarity measure between test handwriting and reference handwriting and that between test handwriting and negative handwriting. Experiments demonstrated the effectiveness of the proposed algorithm.

1. Introduction

Writer verification[1] is a kind of biometric based identity identification method, which has a wide variety of potential applications, from security, forensics, financial activities to archeology. Compared with other biometric based identity identification method, writer verification has lots of merits, such that handwritings are easy to acquire, the verification is non-intrusive and is widely accepted by people. Currently, most writer verification tasks are fulfilled by specially trained document examiners, which is very time-consuming and tend to be subjective. As more and more writer verification tasks need to be handled, automatic writer verification by computer is getting more and more attention by researchers[1][2][3]. However, due to the great difficulty of writer verification, current methods for writer verification are usually not effective enough.

Writer verification is very difficult. The first reason is that handwritings are usually not stable enough and have vast variability, so it is difficult to obtain discriminant handwriting features. The second reason is that in real world writer verification problems the number of writers to be verified is often unlimited and each writer has only few reference handwriting samples, extremely, only one reference handwriting sample served as training samples.

Generally, writer verification can be classified into close set writer verification[2][3] and open set one. In close set writer verification, the number of writers to be verified is known in advance, and the training handwriting samples are relative large, so we can design verification algorithm specific to this group of writers. While in open set one, the number of writers to be verified is unlimited and each writer has few training handwriting samples, extremely, only one sample. This kind of verification task is more difficult than the close set one and there is few research related to this problem. However, many real world writer verification problems, especially forensic justification cases, can be classified into this category, and it is more adaptable and more convenient than the close set one in practical applications. In [2], writer verification was performed on a given database and the dissimilarity measure was constructed based on this database, if the dissimilarity measure was smaller than a given threshold, then the test handwriting was accepted, otherwise, it was rejected. However, this close set writer verification method was designed only for the specific database. If we want to compare whether two handwritings were written by an individual or not and the writer or the writers of these handwritings aren't in the enrolled database, this method will fail. A common method to this sort of open set writer verification problem is to directly define a kind of similarity measure between these two handwritings, if the similarity measure is

Proceedings of the 9th Int'l Workshop on Frontiers in Handwriting Recognition (IWFHR-9 2004)
0-7695-2187-8/04 $20.00 © 2004 **IEEE**

smaller than a predefined threshold, these two handwritings are regarded as written by one individual. However, this similarity measure based method only considered the influence of the client's reference handwriting and didn't take negative handwritings into account. In fact, according to Bayesian theory, it is necessary to consider the influence of the negative handwriting samples, which are defined as samples written by other individuals.

In this paper, we proposed an open set writer verification algorithm using negative samples. First, directional element features (DEFs) are extracted from the handwriting characters, and then we point out the drawback of the similarity measure based method and deduce our writer verification algorithm from Bayesian theory. In the algorithm, both the similarity measure between the test handwriting and the reference one and that between the test handwriting and the negative one are calculated, and they are combined to make a joint decision.

The paper is organized as follows. In Sec. 2, DEFs extraction is briefly introduced. In Sec. 3, the conventional similarity measure based method is introduced and its disadvantages are identified. While in Sec. 4, we first explain why it is necessary to introduce negative handwriting samples, and then introduce our open set writer verification algorithm in Sec. 4.2. Sec. 5 gives the construction of negative handwriting character samples. The experimental results based on a single character and on the combination of several characters are first given in Sec. 6.1 & 6.2, respectively. Then the verification results on some real cases provided by the Second Research Institute of Police in China (SRIPC) are given in Sec. 6.3. Conclusions are drawn in Sec. 7.

2. DEFs extraction

The first problem of writer verification is to choose a set of effective discriminant handwriting features. In our algorithm, we use DEFs[4][5][6] that have been proved very effective for writer identification[6] as our handwriting features. Different from DEFs extraction in character recognition, where characters are first normalized by nonlinear normalization method in order to alleviate the writing style changes, we use gravity-center linear normalization method in order to keep the different writing styles of different writers. After normalization, contour extraction is done. Then each contour pixel is assigned a 4-dimensional vector to measure four types of directional element attribute (horizontal, vertical, and two diagonal, which is very similar to the stroke styles of Chinese characters) according to its neighboring contours. Divide the script into $N_1 \times N_1$ sub-blocks, count up the contour pixels which have the same type of directional element attribute in each block and these sums constitute the 4-dimension vector of this sub-block. We then divide the script into $N_2 \times N_2$ sub-areas, each sub-area contains several sub-blocks, and the DEFs are extracted from these sub-areas. More details can be found in [6].

3. Similarity measure based writer verification method and its drawbacks

Assume $\boldsymbol{X}$ is the test handwriting feature vector, $\boldsymbol{C}$ the claimed writer's reference handwriting feature vector, $D(\boldsymbol{X},\boldsymbol{C})$ the similarity measure between two feature vectors. The writer verification algorithm can be expressed as follows:

if $D(\boldsymbol{X},\boldsymbol{C}) < t$, then accept $\boldsymbol{X}$

$D(\boldsymbol{X},\boldsymbol{C}) > t$, then reject $\boldsymbol{X}$

where t is the threshold. Although it is very simple and very straightforward, it has the following drawbacks. Firstly, the threshold t is determined by subjective experience, can't be adapted to different situations. Secondly, it doesn't consider the influence of negative handwriting samples. In fact, the decision boundary is determined not only by claimed writer's reference samples but also by negative samples. It is apparent not suitable to use just the similarity measure between the test sample and the reference samples to make such a decision.

4. Writer verification using negative samples

4.1. Necessity of using negative samples

Let Ω_c be the writer of the reference sample's feature vector $\boldsymbol{C}$, $l(\boldsymbol{X}) = \dfrac{p(\boldsymbol{X} \mid \Omega_c)}{p(\boldsymbol{X} \mid \overline{\Omega}_c)}$ likelihood ratio, according to Bayesian decision rule[7],

$$\text{if } l(\boldsymbol{X}) > \frac{p(\overline{\Omega}_c)}{p(\Omega_c)}, \text{then } \boldsymbol{X} \in \Omega_c$$

$$l(\boldsymbol{X}) < \frac{p(\overline{\Omega}_c)}{p(\Omega_c)}, \text{then } \boldsymbol{X} \in \overline{\Omega}_c$$

where $\overline{\Omega}_C$ denotes other writers. Although the rule is the optimal decision making method and can obtain the minimum classification error rate in principle, it needs to estimate class conditional probability density function (cPDF). As there are only few reference samples in open set writer verification problem, and the form of the cPDF often can't be determined in advance, the rule is of little practical usage. An alternative method is to define a kind of discriminant function $h(\boldsymbol{X},\theta)$ based on some assumptions. The parameters θ of the discriminant function can be determined by reference samples. So the decision rule based on disciminant function is as follows:

if $h(\boldsymbol{X},\theta_{\Omega_c}) < h(\boldsymbol{X},\theta_{\overline{\Omega}_c})$, then $\boldsymbol{X} \in \Omega_c$

$h(\boldsymbol{X},\theta_{\Omega_c}) > h(\boldsymbol{X},\theta_{\overline{\Omega}_c})$, then $\boldsymbol{X} \in \overline{\Omega}_c$.

Or equivalently, let $f(\boldsymbol{X}) = \dfrac{h(\boldsymbol{X},\theta_{\Omega_c})}{h(\boldsymbol{X},\theta_{\overline{\Omega}_c})}$, then

if $f(\boldsymbol{X}) < 1$, then $\boldsymbol{X} \in \Omega_c$

$f(\boldsymbol{X}) > 1$, then $\boldsymbol{X} \in \overline{\Omega}_c$.

where θ_{Ω_c}, $\theta_{\overline{\Omega}_c}$ denote the parameters of discriminant function of Ω_c and $\overline{\Omega}_c$, respectively.

From above discussion, it is easy to know that in order to verify a test handwriting, it is necessary to combine both the value of the claimed writer's discriminant function and that of other writers' discriminant function. This is the reason that we should introduce negative samples.

4.2. The verification algorithm

According to different assumption of the probability distribution of the claimed writer's handwriting and that of negative handwriting, the discriminant function can have different forms. The number of training samples also influences the form of discriminant function. In this paper, we select $h(\boldsymbol{X},\theta_{\Omega_c}) = \|\boldsymbol{X} - C\|^2$ as Euclidean distance measure due to limited reference samples.

From the discussion of Sec. 4.1, we propose our verification algorithm as follows:

a) Construct the negative handwriting samples;
b) Calculate the similarity measure between test handwriting and the claimed writer's reference handwriting DT, $DT = h(\boldsymbol{X},\theta_{\Omega_c}) = \|\boldsymbol{X} - \overline{\boldsymbol{C}}\|^2$, where $\overline{\boldsymbol{C}}$ is the mean vector of the claimed writer's reference handwriting feature vectors. When there is only one reference sample, $\overline{\boldsymbol{C}}$ is the reference feature vector;
c) Calculate the similarity measure between test handwriting and negative handwritings DF.

Assume there are m negative handwriting samples $F_i (i = 1,2,...,m)$, then

$$DF_i = h(\boldsymbol{X},\theta_{\overline{\Omega}_c}) = \|\boldsymbol{X} - \boldsymbol{F}_i\|^2$$

The value DF is the combination of these values,

$$DF = \sum_{i=1}^{m} w_i \cdot DF_i$$

where w_i is the weight, $w_i > 0$, $\sum_{i=1}^{m} w_i = 1$.

d) Combine DT and DF to give a final score V, $V = f(DT, DF)$, where $f(x, y)$ is a combination function. We select $f(x, y) = x / y$;
e) Make decision. If $V < t$, then accept $\boldsymbol{X}$; else, reject $\boldsymbol{X}$, where t is a threshold.

Compared with conventional similarity measure based algorithm, our algorithm introduced negative handwriting samples and the decision threshold is a relative value. Experiments proved it is very effective.

5. Construction of negative handwriting samples

The most important part of our algorithm is to effectively construct the negative handwriting samples. These samples should be representative enough to contain all kinds of handwriting variations, and shouldn't depend on a specific writer. For these reasons, we use handwriting Chinese character samples collected by our lab as negative samples. These samples were written by many different individuals and contain lots of variations, so they are good representative of negative samples. For each character, there are 1806 samples in total, which is a very large number. We first cluster the 1806 samples into 40 clusters using *K*-means cluster algorithm, and treat the 40 cluster centers as new negative samples.

6. Experimental results

We first give writer verification result based on a single character, then we give the verification result based on the combination of several characters, at last we give the verification result on some real cases provided by SRIPC.

6.1. Verification result based on single character

For the experiments, we use two databases to evaluate the effectiveness of the proposed algorithm. One database consists of 20 Chinese characters as scripts, each was written 16 times by 27 people. We call this database LABSet1 (Figure 3). The other is composed of 16 repeated writings of a Chinese text with 33 characters written by 25 different individuals. We call this database LABSet2. For each character script of a writer, we randomly select a sample as reference character script, the remaining 15 samples as test handwritings. The samples of the same character written by other individuals are served as negative test handwritings. So there are 405 true verification tests and 11232 negative verification tests for each character in LABSet1, and 375 true verification tests and 9600 negative verification tests for each character in LABSet2. First DEFs are extracted, then we use the method described in Sec. 4.2 to perform writer verification. The threshold is selected when FRR (False Rejection Rate) equals to FAR (False Acceptance Rate).

Table 1. Writer verification result for LABSet1

Script	生	的	无	难	别
EER-SM(%)	24.40	20.00	25.00	17.10	19.92
EER-PS(%)	17.36	13.59	15.90	11.55	18.96
Script	花	成	但	月	此
EER-SM(%)	18.02	20.28	20.00	23.20	20.00
EER-PS(%)	13.91	14.57	19.96	17.48	17.73
Script	去	为	中	有	天
EER-SM(%)	23.02	21.00	24.00	22.00	24.10
EER-PS(%)	13.97	14.27	22.86	19.15	17.33
Script	不	是	和	在	人
EER-SM(%)	22.80	26.70	23.72	19.80	29.60
EER-PS(%)	16.13	16.94	11.55	18.34	25.24
Aver. EER-SM(%)		22.23			
Aver. EER-PS(%)		16.84			

Note:

EER-SM: Equal error rate using conventional similarity measure based algorithm

EER-PS: Equal error rate using the proposed algorithm

Aver. EER-SM: average equal error rate for similarity measure based algorithm

Aver. EER-PS: average equal error rate for the proposed algorithm

Table 2. Writer verification result for LABSet2

Script	把	慢	坪	讲	河
EER-SM(%)	19.51	25.56	17.74	22.96	17.29
EER-PS(%)	15.75	20.11	16.39	18.01	14.26
Script	细	妹	骆	辩	苹
EER-SM(%)	25.97	21.91	19.55	19.89	20.11
EER-PS(%)	18.66	17.96	12.92	17.71	18.04
Script	害	感	禁	常	粟
EER-SM(%)	19.18	19.21	17.11	17.89	20.21
EER-PS(%)	14.95	14.41	14.54	15.22	15.61
Script	延	这	骨	适	爬
EER-SM(%)	21.60	24.30	20.29	19.85	16.41
EER-PS(%)	17.10	19.30	15.66	14.84	11.64
Script	起	座	原	病	屏
EER-SM(%)	20.33	20.71	16.57	16.04	18.31
EER-PS(%)	15.54	16.72	13.02	14.59	16.08
Script	氧	虎	凤	国	问
EER-SM(%)	16.52	22.45	20.44	18.10	20.96
EER-PS(%)	13.33	19.74	16.51	13.49	18.06
Script	表	首	鬼	Average	
EER-SM(%)	22.25	23.27	15.45	19.94	
EER-PS(%)	16.83	18.06	12.54	15.99	

Note:

EER-SM: Equal error rate using conventional similarity measure based algorithm

EER-PS: Equal error rate using the proposed algorithm

Average: average equal error rate for the two algorithms

The experimental results for LABSet1 and LABSet2 are shown in Table 1 and Table 2, respectively. For comparison, we also listed the experimental results using conventional similarity measure based algorithm.

From Table 1 and Table 2, we can see that the EER (Equal Error Rate) differs from one character script to another dependent on the structural complexity of the script (For example, the EER of "难" using the proposed algorithm is 11.55%, while the EER of "人" using the same algorithm is 25.24%). Although different character script has different EER, the EER of the proposed algorithm is much lower than that of conventional similarity measure based algorithm for each character script. The average EER dropped from 22.23% to 16.84% for LABSet1 and from 19.94% to 15.99% for LABSet2. The experiments show that the proposed algorithm is significant better than conventional method.

6.2. Verification result based on the combination of several characters

In order to improve the performance of writer verification, we often need to combine several characters to make a joint decision.

The characters from an individual can be written in different time and can be extracted from different handwriting scripts, so they can be viewed as independent of one another. Use similar methods described in [8] we can further improve the verification result by combining several characters.

In the experiment, we randomly chose n (n from 1 to 20) characters to make a joint verification. First sigmoid function $g(x) = \frac{1}{1+\exp(ax+b)}$ is used to map each character's score V into probability and the parameters are determined by cross-validation (a = 0.7734, b = -0.06889). Then sum rule[8] is applied to fulfill combination task. We repeat the experiment 10 times and the results below are an average of them.

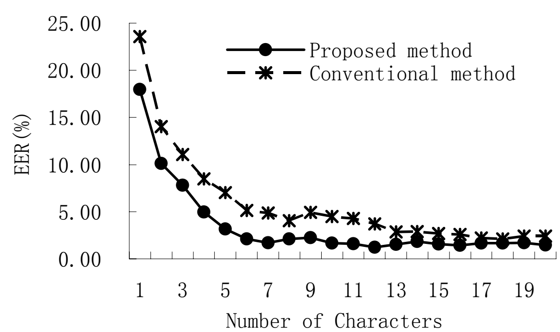


Figure 1. Combination result for LABSet1

Figure 1 & Figure 2 show the combination results for LABSet1 and LABSet2, respectively. As the number of characters increases, the EER first drops rapidly, then it drops slowly as the number of combined characters exceeds 5. One possible explanation is that the characters are actually not independent of one another, and the error rate shouldn't be less than Bayesian error rate, which is the theoretic minimum. However, as long as the number of characters is greater than 5, the EER would be less than 5% (3.17% for LABSet1 and 4.68% for LABSet2) using the proposed algorithm. And the EER of the proposed algorithm is lower than that of conventional similarity measure based method for the same number of characters, which shows the proposed algorithm is very effective.

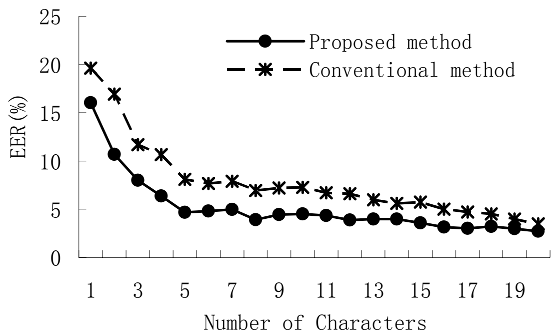


Figure 2. Combination result for LABSet2

6.3. Verification result on real cases

Here we give the writer verification result on some real cases (Figure 4) provided by SRIPC to validate the effectiveness of the proposed algorithm. There are in total 34 cases, the handwritings of which include threat letters, testaments, checks, etc. For each case we can obtain at least 5 different characters, and the final verification result is the combination of these characters. For the similarity measure based algorithm, 22 cases are correctly verified, 6 cases are rejected and 6 cases are wrongly verified. For the proposed algorithm, 31 cases are correctly verified, 3 cases are rejected and none of the cases are wrongly verified. (By "correctly verified" we mean that the verification result is the same as the result by the document examiners, "wrongly verified" means the result is not as same as that by the experts.) These results reveal the proposed algorithm is more effective than similarity measure based algorithm.

7. Conclusion

In this paper, we proposed an effective open set writer verification algorithm using negative samples. The directional element features are first extracted from the handwriting character scripts. We then pointed out the deficiency of the conventional similarity measure based algorithm, and proposed the writer verification algorithm using as decision score the combination value of the similarity measure between test handwriting and the client's reference handwriting and that between test handwriting and negative handwriting. Negative handwriting samples were introduced by clustering handwriting Chinese character samples. Experiments were performed on two handwriting databases gathered by our lab and on some real cases provided by SRIPC, which show the proposed algorithm is more effective than conventional one.

References

[1] R. Plamondon, G. Lorette, "Automatic Signature Verification and Writer Identification – The State of the Art", *Pattern Recognition*, vol. 22, No. 2, pp. 107-131, 1989.

[2] I. Yoshimura, M. Yoshimura, "Off-line Writer Verification Using Ordinary Characters as The Object", *Pattern Recognition*, vol. 24, No. 9, pp. 909-915, 1991.

[3] Hong Yan, Jing Wu, "Writer Verification Using Multiple Neural Networks", *IEEE International Conference on Neural Networks*, vol. 1, 1995, pp. 414-417.

[4] N. Kato, M. Suzuki, S. Omachi, etc, "A Handwritten Character Recognition System Using Directional Element Feature and Asymmetric Mahalanobis Distance", *IEEE Trans. on PAMI*, vol. 21, No. 3, Mar. 1999, pp. 258-262.

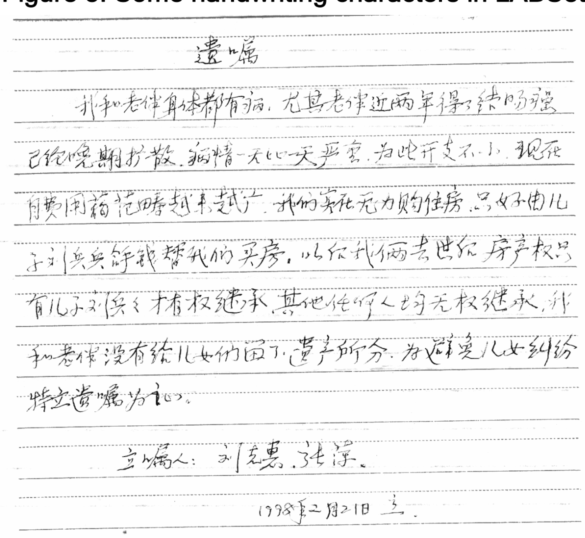
[5] J. Zhang, X. Ding, C. Liu, "Multi-Scale Feature Extraction and Nested-Subset Classifier Design for High Accuracy Handwritten Character Recognition", *Proc. 15th Int. Conf. on Pattern Recognition*, vol. 2, 2000, pp. 581-584.

[6] X. L. Wang, X. Q. Ding, and H. L. Liu, "Writer Identification Using Directional Element Features and Linear Transform", *ICDAR'2003*, pp. 942-945.

[7] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis*. Wiley, New York, 1973.

[8] J. Kittler, M. Hatef, R. P. W. Duin, et al, "On combining classifiers", *IEEE Trans. on PAMI*, 1998, 20(3), pp. 226-239.

Figure 3. Some handwriting characters in LABSet1



遗嘱

我和老伴身体都有病，尤其老伴近两年得了结肠癌，已经晚期扩散，病情一天比一天严重，为此开支不小，现在自费用药花费越来越大。我们实在无力购住房，只好由儿子刘兵兵出钱替我们买房，以后我俩去世后房产权只有儿子刘兵兵才有权继承，其他任何人均无权继承，我和老伴没有给儿女们留下遗产可分，为避免儿女纠纷特立遗嘱为证。

立嘱人：刘志惠、张萍。

1998年2月21日 立。

Figure 4. One real handwriting provided by SRIPC

Proceedings of the 9th Int'l Workshop on Frontiers in Handwriting Recognition (IWFHR-9 2004)
0-7695-2187-8/04 $20.00 © 2004 IEEE