Incremental MQDF Learning for Writer Adaptive Handwriting Recognition[1]

Kai Ding, Lianwen Jin*
School of Electronic and Information Engineering, South China University of Technology, Guangzhou, P.R.China
**E-Mail: lianwen.jin@gmail.com*

Abstract

Writer adaptation has been proved to be an effective approach to improve the recognition performance of the writer-independent recognizer for a particular writer. In this paper, we propose a writer adaptive handwriting recognition approach by incremental learning the Modified Quadratic Discriminant Function (MQDF) classifier. We derived the solution of Incremental MQDF (IMQDF) and then present a Discriminative IMQDF (DIMQDF) by deriving the solution of IMQDF in the updated discriminative feature space. Based on IMQDF or DIMQDF, the writer adaptation is finally performed by updating the MQDF recognizer adaptively. The experimental results for recognizing handwriting Chinese characters indicate that the proposed IMQDF and DIQMDF approaches can reduce as much as 52.71% and 45.38% error rate respectively on the writer-dependent dataset while only have less than 0.18% accuracy loss on the writer-independent dataset. In other words, the proposed IMQDF and DIMQDF based writer adaptation approaches can significantly increase the recognition accuracy on writer-dependent dataset while only have limited negative influence for general writer.

1. Introduction

With the emergence of Personal Digital Assistants (PDA) and of Tablet PCs using pen based interfaces, the handwriting recognition accuracy is the key factor in determining the acceptability of a handwriting recognition system and the whole application in which it is implemented. Motivated by these, many researchers devoted themselves to the field of handwriting character recognition and achieved great progress during the past 40 years [1-3]. However, recent researches [4-6] have shown that the recognition accuracy using state-of-the-art techniques cannot satisfy user's expectations [7].

The training set of a writer-independent recognizer is typically composed of large sets of data produced by different writers to achieve a good performance for general writer, but it is not optimal with respect to any particular writer. A straightforward idea to increase the recognition accuracy for a specific user is adapting the writer-independent recognizer to the specific writing style, which is also known as writer adaptation [8-12].

Many the previous works either design for the small scale (small classes, small vocabulary) recognition problem [8-10] or focus on feature transformation module and ignore the classifier design module during the writer adaptation process [11-12]. In this paper, we propose a writer adaptation approach for large scale online handwriting Chinese character recognition by adaptively updating the Modified Quadratic Discriminant Function (MQDF) recognizer, which has been successful in achieving state-of-the-art performance for handwriting character recognition [13-14]. However, the assumption used in typical implementations of the traditional MQDF, that a complete dataset for training is given in advance, conflicts with the principles of the writer adaptation.

To solve this problem, we have derived a solution of the MQDF model using incremental learning (henceforth referred to as the IMQDF), which enables the MQDF technique to be applied to the writer adaptation for the first time. On the other hand, in practical, the raw feature is usually transformed to a discriminative feature space to reduce the feature dimension as well as improve the recognition accuracy. In this situation, we can directly apply the IMQDF approach by assuming the discriminative feature space is preserved during the incremental learning process.

1 This work is supported in part by the research funding of NSFC (no. U0735004, 60772216), GDSTP (no. 07118074, 2007B010200048, 2008A050200004, 2009B090300394)

978-0-7695-4221-8/10 $26.00 © 2010 IEEE
DOI 10.1109/ICFHR.2010.92

However, this assumption is not reasonable and may cause recognition accuracy decrease due to the discriminative feature space is actually changed when the incremental samples are provided. To solve this problem, we proposed a Discriminative IMQDF (DIMQDF) to adaptively learn the MQDF model in the updated discriminative feature space. Based on these, the writer adaptation is carried out by adaptive updating MQDF recognizer to learn the specific writer's writing styles. Experimental results on the writer-dependent dataset verify that both of the IMQDF and DIMQDF approaches are very effective to improve the recognition accuracy for a particular writer, and DIMQDF achieves better performance than IMQDF approach. Moreover, experimental results on the independent-dataset show that both of the proposed writer adaptation approaches have limited accuracy loss for the general writer.

The remainder of this paper is organized as follows: Section2 first reviews the MQDF approach and then derives the solution of IMQDF and DIMQDF. The WIMQDF/WDIMQDF approach is proposed in Section3. And the proposed writer adaptation method is presented in Section4. Section5 gives experimental results and Section6 concludes this paper.

2. Incremental learning of MQDF

2.1 MQDF classifier

Based on a Bayesian decision rule that classifies the input pattern to the class as the maximum a posteriori (MAP) probability of all the classes, the quadratic discriminant function (QDF) is obtained assuming a multivariate Gaussian density and an equal prior probability for each class. The MQDF proposed by Kimura et al. [13] incorporates a modification to the QDF through a K-L transform and smoothing the minor eigenvalues to improve the computation efficiency and classification performance.

The discriminant function of a QDF classifier can be represented as [14]:

$$g_0(\boldsymbol{x},\omega_i)=(\boldsymbol{x}-\bar{\boldsymbol{x}}_i)^T\textstyle\sum_i^{-1}(\boldsymbol{x}-\bar{\boldsymbol{x}}_i)+\log|\Sigma_i|,\quad i=1,\ldots,M \tag{1}$$

where M is the number of classes and ω_i represents the i^{th} class, and $\bar{\boldsymbol{x}}_i$ denotes the mean vector of class ω_i, $\boldsymbol{D}$ is the dimension of $\bar{\boldsymbol{x}}_i$.

By K-L transform, the covariance matrix can be diagonalized as:

$$\Sigma_i=\Phi_i\Lambda_i\Phi_i^T \tag{2}$$

where $\Lambda_i=diag[\lambda_{i1},\ldots,\lambda_{iD}]$ with λ_{ij}, $j=1,\ldots,D$, being the eigenvalues (ordered in decreasing order) of Σ_i, and $\Phi_i=[\phi_{i1},\ldots,\phi_{iD}]$ with ϕ_{ij}, $j=1,\ldots,D$, being the ordered eigenvectors, Φ_i is orthonormal (unitary) such that $\Phi_i^T\Phi_i=I$.

According to (2), the QDF can be rewritten in the form of eigenvectors and eigenvalues:

$$\begin{aligned}g_0(\boldsymbol{x},\omega_i)&=[\Phi_i^T(\boldsymbol{x}-\bar{\boldsymbol{x}}_i)]^T\Lambda_i^{-1}\Phi_i^T(\boldsymbol{x}-\bar{\boldsymbol{x}}_i)+\log|\Lambda_i|\\&=\sum_{j=1}^{D}\frac{1}{\lambda_{ij}}[\varphi_{ij}^T(\boldsymbol{x}-\bar{\boldsymbol{x}}_i)]^2+\sum_{j=1}^{D}\log\lambda_{ij}\end{aligned} \tag{3}$$

By replacing the minor eigenvalues with a constant δ_i, the MQDF classifier can be expressed as:

$$\begin{aligned}g_1(\boldsymbol{x},\omega_i)&=\frac{1}{\delta_i}\left\{\left\|\boldsymbol{x}-\bar{\boldsymbol{x}}_i\right\|^2-\sum_{j=1}^{K}(1-\frac{\delta_i}{\lambda_{ij}})[\varphi_{ij}^T(\boldsymbol{x}-\bar{\boldsymbol{x}}_i)]^2\right\}\\&+\sum_{j=1}^{K}\log\lambda_{ij}+(D-K)\log\delta_i\end{aligned} \tag{4}$$

where K denotes the number of dominant eigenvectors.

Since the training of the QDF classifier always underestimate the patterns' eigenvalues by limited sample set, the minor eigenvalues become some kind of unstable noises and affect the classifier's robustness. By smoothing them in the MQDF classifier, not only the classification performance is improved, but also the computation time and storage for the parameters are saved.

2.2 Incremental learning of MQDF (IMQDF)

Suppose $\boldsymbol{X}$ and $\boldsymbol{Y}$ are two sets of observations in the feature space, where $\boldsymbol{X}$ is the given observation set with N samples $\boldsymbol{X}=\{\boldsymbol{x}_i\}$ $(i=1,\ldots,N)$ in M classes, and $\boldsymbol{Y}$ is a set of new observations with L incremental samples $\boldsymbol{Y}=\{\boldsymbol{y}_j\}$ $(j=1,\ldots,L)$ in P classes. It is worthwhile noting that some of classes in $\boldsymbol{Y}$ may be not available in $\boldsymbol{X}$, in other words, some new classes may be introduced. Thus, the mixed observation set $\boldsymbol{Z}=\boldsymbol{X}\cup\boldsymbol{Y}=\{\boldsymbol{z}_k\}$ $(k=1,\ldots,L+N)$ has $L+N$ samples in C classes, where $C\geqslant M$ and $C\geqslant P$. Without loss of generality, we assume that n_i of the N original samples and l_i of the L incremental samples belong to class C_i $(i=1,\ldots,C)$, therefore, in the updated observation set, the sample number belonging to the C_i class is $s_i=n_i+l_i$. Let $\bar{x}_i,\Sigma_{xi},\bar{y}_i,\Sigma_{yi}$ represent the i^{th} class's mean vector and covariance matrix of the given observation set and the new observation set respectively. Then we have

$$\bar{\boldsymbol{x}}_i=\frac{1}{n_i}\sum_{k=1}^{n_i}\boldsymbol{x}_{ik} \tag{5}$$

$$\bar{\boldsymbol{y}}_i=\frac{1}{l_i}\sum_{k=1}^{l_i}\boldsymbol{y}_{ik} \tag{6}$$

$$\Sigma_{xi} = \sum_{k=1}^{n_i} (\boldsymbol{x}_{ik} - \bar{\boldsymbol{x}}_i)(\boldsymbol{x}_{ik} - \bar{\boldsymbol{x}}_i)^T = \frac{1}{n_i} \sum_{k=1}^{n_i} \boldsymbol{x}_{ik} \boldsymbol{x}_{ik}^T - \bar{\boldsymbol{x}}_i \bar{\boldsymbol{x}}_i^T \quad (7)$$

$$\Sigma_{yi} = \sum_{k=1}^{l_i} (\boldsymbol{y}_{ik} - \bar{\boldsymbol{y}}_i)(\boldsymbol{y}_{ik} - \bar{\boldsymbol{y}}_i)^T = \frac{1}{l_i} \sum_{k=1}^{l_i} \boldsymbol{y}_{ik} \boldsymbol{y}_{ik}^T - \bar{\boldsymbol{y}}_i \bar{\boldsymbol{y}}_i^T \quad (8)$$

According to the above definitions, each class's mean vector $\bar{z}_i$ and covariance matrix Σ_{zi} of the mixed dataset can be updated as follows:

$$\bar{\boldsymbol{z}}_i = \frac{1}{s_i} \sum_{k=1}^{s_i} \boldsymbol{z}_{ij} = \frac{1}{n_i + l_i} \left(\sum_{k=1}^{n_i} \boldsymbol{x}_{ik} + \sum_{j=1}^{l_i} \boldsymbol{y}_{ij} \right) = \frac{n_i \bar{\boldsymbol{x}}_i + l_i \bar{\boldsymbol{y}}_i}{n_i + l_i} \quad (9)$$

$$\begin{aligned}
\Sigma_{zi} &= \frac{1}{s_i} \sum_{k=1}^{s_i} \boldsymbol{z}_{ik} \boldsymbol{z}_{ik}^T - \bar{\boldsymbol{z}}_i \bar{\boldsymbol{z}}_i^T \\
&= \frac{1}{n_i + l_i} \sum_{k=1}^{n_i} \boldsymbol{x}_{ik} \boldsymbol{x}_{ik}^T + \frac{1}{n_i + l_i} \sum_{k=1}^{l_i} \boldsymbol{y}_{ik} \boldsymbol{y}_{ik}^T - \left(\frac{n_i \bar{\boldsymbol{x}}_i + l_i \bar{\boldsymbol{y}}_i}{n_i + l_i}\right)\left(\frac{n_i \bar{\boldsymbol{x}}_i + l_i \bar{\boldsymbol{y}}_i}{n_i + l_i}\right)^T \\
&= \frac{n_i}{n_i + l_i} \left(\frac{1}{n_i} \sum_{k=1}^{n_i} \boldsymbol{x}_{ik} \boldsymbol{x}_{ik}^T - \bar{\boldsymbol{x}}_i \bar{\boldsymbol{x}}_{ik}^T \right) + \frac{l_i}{n_i + l_i} \left(\frac{1}{l_i} \sum_{k=1}^{l_i} \boldsymbol{y}_{ik} \boldsymbol{y}_{ik}^T - \bar{\boldsymbol{y}}_i \bar{\boldsymbol{y}}_i^T \right) \\
&+ \frac{n_i l_i}{(n_i + l_i)^2} (\bar{\boldsymbol{y}}_i \bar{\boldsymbol{y}}_i^T + \bar{\boldsymbol{x}}_i \bar{\boldsymbol{x}}_i^T - \bar{\boldsymbol{x}}_i \bar{\boldsymbol{y}}_i^T - \bar{\boldsymbol{y}}_i \bar{\boldsymbol{x}}_i^T) \\
&= \frac{n_i}{n_i + l_i} \Sigma_{xi} + \frac{l_i}{n_i + l_i} \Sigma_{yi} + \frac{n_i l_i}{(n_i + l_i)^2} (\bar{\boldsymbol{y}}_i - \bar{\boldsymbol{x}}_i)(\bar{\boldsymbol{y}}_i - \bar{\boldsymbol{x}}_i)^T
\end{aligned} \quad (10)$$

After updating the mean vector and covariance matrix of each class, the covariance matrix of each class is diagonalized by K-L transformation to obtain the dominant eigenvalues and eigenvectors. Finally the MQDF classifier can be updated according to Eq. (4).

2.3 Discriminative IMQDF (DIMQDF)

In practical, before using MQDF, a discriminative transform approach, such as Linear Discriminative Analysis (LDA) [15], Principal Component Analysis (PCA)[14] and General Tensor Discriminative Analysis (GTDA)[16], is usually employed to transform the raw feature to a discriminative feature space to reduce the feature dimension as well as improve the recognition performance. After then, the MQDF classifier is performed in the discriminative feature space. Due to the proposed IMQDF only can be used in the preserved feature space, therefore, when the discriminative transform approach is used, the IMQDF approach can't be directly employed without the assumption that the discriminative feature space is preserved during the incremental learning process, but this assumption may cause performance loss.

To solve this problem, we proposed a Discriminative IMQDF (DIMQDF) to improve performance of IMQDF by adaptively learning the MQDF classifier in the updated discriminative feature space. Since the LDA, also known as Fisher Discriminative Analysis (FDA), is widely used approach in the character recognition applications [1,3,4,11,12,14], in this section, we focus on how to incremental learning MQDF classifier in the LDA feature space.

Suppose $\bar{x}_i, \Sigma_{xi}, \bar{y}_i, \Sigma_{yi}$ represents the i^{th} class's mean vector and covariance matrix of the given and new observation sets respectively in the raw feature space. And W_{orglda} denotes the LDA transform matrix, which is trained by the given observation data. When the incremental samples are provided, the LDA transformation matrix is updated as W_{inclda} according to the ILDA approach [12].

After then, each class's mean vector $\bar{y}_{i-inclda}$ and covariance matrix $\Sigma_{yi-inclda}$ of the incremental data can be computed in the updated discriminative feature space according to:

$$\bar{\boldsymbol{y}}_{i\text{-}inclda} = W_{inclda}^T \bar{\boldsymbol{y}}_i \quad (11)$$

$$\Sigma_{yi-inclda} = \sum_{k=1}^{l_i} (W_{inclda}^T \boldsymbol{y}_{ik} - \bar{\boldsymbol{y}}_{i\text{-}inclda})(W_{inclda}^T \boldsymbol{y}_{ik} - \bar{\boldsymbol{y}}_{i\text{-}inclda})^T \quad (12)$$

Similar with above steps, each class's mean vector $\bar{x}_{i-inclda}$ and covariance matrix $\Sigma_{xi-inclda}$ of the given data can be computed in the new discriminative feature space as:

$$\bar{\boldsymbol{x}}_{i\text{-}inclda} = W_{inclda}^T \bar{\boldsymbol{x}}_i \quad (13)$$

$$\begin{aligned}
\Sigma_{xi-inclda} &= \sum_{k=1}^{l_i} (W_{inclda}^T \boldsymbol{x}_{ik} - W_{inclda}^T \bar{\boldsymbol{x}}_i)(W_{inclda}^T \boldsymbol{x}_{ik} - W_{inclda}^T \bar{\boldsymbol{x}}_i)^T \\
&= W_{inclda}^T \sum_{k=1}^{l_i} (\boldsymbol{x}_{ik} - \bar{\boldsymbol{x}}_i)(\boldsymbol{x}_{ik} - \bar{\boldsymbol{x}}_i)^T W_{inclda} = W_{inclda}^T \Sigma_{xi} W_{inclda}
\end{aligned} \quad (14)$$

Since all the class's mean vector and the covariance matrix of the given data and incremental data are obtained, each class's updated mean vector and covariance matrix of mixed dataset can be computed by inserting Eq. (11)~(14) to Eq. (9) and Eq. (10) respectively. After obtaining the dominant eigenvalues and eigenvectors of each class's covariance matrix by K-L transformation, the MQDF classifier can be updated in the updated discriminative feature space according to Eq. (4).

4. Writer adaptation

Figure 1 depicts a flow diagram of the proposed WDIMQDF/DIMQDF based writer adaptation approach. The initial MQDF modeling is trained using the full writer-independent training set. The preprocessing and feature extraction are performed firstly. After then, the feature space is transformed to a discriminative feature space to reduce the feature dimension and improve the recognition accuracy by LDA. Thereafter, to generate the MQDF recognizer for the writer-independent dataset, the adaptation

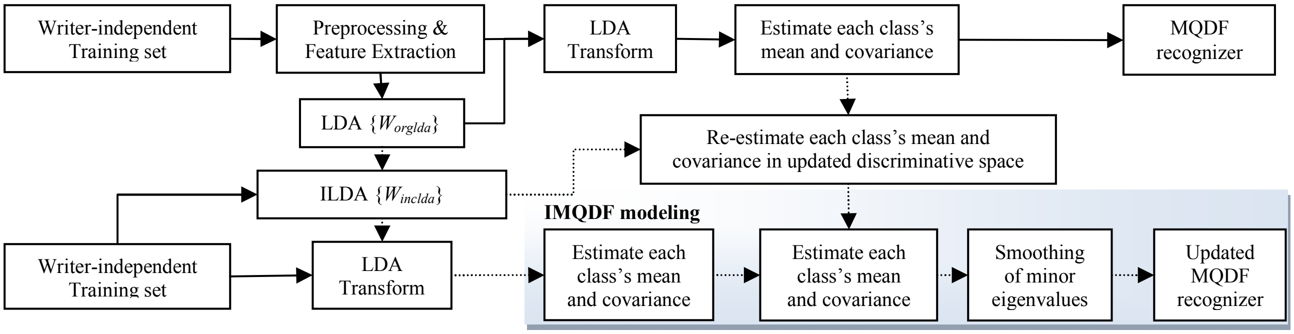


Figure 1 An overview of writer adaptation process (dashed lines show the flow of events for writer adaptation)

approach tries to find a statistical model and estimates the model's parameters to fit the feature distribution of each class. When additional samples for a particular writer are provided, the LDA feature space is updated first, and then the parameters of each class's feature distribution are re-estimated in the updated discriminative feature space to learn the writing styles of the particular writer adaptively. Finally, the updated MQDF recognizer is generated to improve the recognition accuracy for the particular writer.

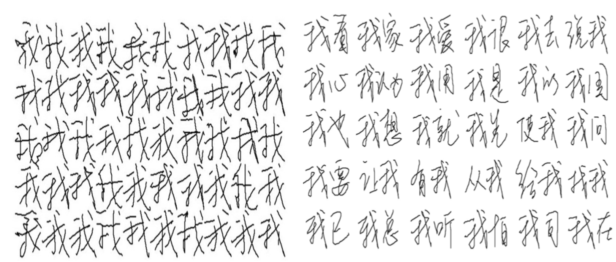
In the writer-independent recognizer training process, the feature of the independent training dataset is first extracted using an 8-directional feature extraction method [17]. And then, the parameters of each class's distribution are estimated to generate the writer-independent MQDF recognizer according to the methods described in Section 2.1.

The accuracy of a recognition system for a particular writer can be improved by adjusting the recognizer's model to match the distribution of the particular writer's handwriting more closely, rather than using the more generalized models trained for a large number of writers. Therefore, our writer adaptation approach begins by adaptive updating the new LDA feature space using the incremental samples. After transforming the given samples' distribution parameters into updated discriminative feature space, we also compute the distribution models of the writer-dependent dataset in the update discriminative feature space. In this way, the distribution models of both the writer-independent and writer-dependent datasets are obtained in the updated discriminative feature space. Thereafter, the distribution models of the mixed dataset can be generated according to Eq. (9) and (10). Finally, the updated MQDF recognizer is developed according to Eq. (4) using the distribution models of the mixed dataset.

5. Experiments and Analysis

5.1 Data preparation and experimental setup

The evaluation of the proposed writer adaptation approach was conducted using two subsets of SCUT-COUCH handwriting dataset [18], GB1 subset (henceforth referred to as ***CouchGB1*** dataset) and ***Word8888*** subset. The GB1 subset contains 168 writers' samples of 3755 categories of Chinese characters, whereas the ***Word8888*** subset consists of 30 writers' samples of 8888 categories of most frequently used handwritten words. Figure2 shows some samples of the character "我" from the GB1 subset and the words contains character "我" in ***Word8888*** subset. It is worthwhile noticing that many different words contain same character. For example, all of the words in figure2 (b) contain the character "我". This indicates that data collected in this way, in which the same character is collected for many times in accordance with the context of the word corpus, provides us with particularly realistic writer-independent incremental handwritten samples.



(a) Character "我" (b) Words contain "我"

Figure2: Samples from the SCUT-COUCH database

To build a general purpose classifier, we randomly selected 134 sets of data from the ***CouchGB1*** to build a writer-independent baseline classifier, with the remaining 34 sets used to test the performance of the baseline classifier, as well as to evaluate the influence after the adaptation. In order to generate particularly realistic writer-dependent incremental handwritten samples, we first manually segmented all of the handwritten words samples into separate characters, which results in 2078 categories of 19595 isolated Chinese characters, to form a new dataset, the

Figure 4 Performance of different writer adaptation approaches for each particular writer

IncCouchDB dataset. For each particular writer's handwritten samples from the ***IncCouchDB*** dataset, we randomly selected 50% of each category's data for adaptive learning the IMQDF model and then using the remaining 50% to test the writer adaption performance. It is worth noting that the two datasets don't share any common writers.

5.2 Baseline Performance on CouchGB1 and IncCouchDB before writer adaptation

After the classifier was trained by 134 sets of ***CouchGB1*** data, its performance was evaluated on both the testing sets of ***CouchGB1*** dataset and ***IncCouchDB*** dataset respectively. Table 1 shows the average recognition rate of the baseline MQDF classifier on these two datasets. The results demonstrate that, since many of the writing styles of ***IncCouchDB*** are unseen in the training dataset, the recognition accuracy on ***IncCouchDB*** testing datasets is much lower than which on ***CouchGB1*** testing dataset. In other words, the recognition accuracy for the specific writer should be further improved.

Table1 Baseline performance on two datasets

	Top1	Top5	Top10
CouchGB1	96.03%	99.30%	99.49%
IncCouchDB	88.56%	95.93%	96.82%

5.3 Performance comparison with different writer adaptation approaches

In this experiment, we examined the performance of the various writer adaptation approaches based on ILDA [12] and the proposed IMQDF and DMQDF approaches on the writer-dependent ***IncCouchDB*** dataset. It worth noting that since the MQDF classifier is not used in this experiment in previous method [12], we implement an ILDA plus MQDF based writer adaptation approach (we refer it as ILDAMQDF), where the LDA transform matrix is first updated using the new incremental data according to ILDA, and then the MQDF classifier model is transformed to the updated LDA feature space without employing the IMQDF approach.

From the results given in table 2, it can be seen that the ILDA and ILDAMQDF based writer adaptation can only reduce less than 26% error rate and achieve less than 91.10% recognition rate for the particular writer, while the proposed IMQDF and DIMQDF based writer adaptation approaches can reduce above 45% error rate and achieve above 93.75% recognition rate. In other words, they can improve the recognition accuracy from 88.56% to 93.75% and 94.59% respectively. On the other hand, it can be also found that the DIMQDF significantly outperform IMQDF, this indicates that by incrementally learning the MQDF model in the updated discriminative feature space, the performance of the writer adaptation can be further improved. Generally, the experimental results demonstrate that the proposed IMQDF based writer adaptation approach is very effective to improve the recognition accuracy for the specific writer, and by adaptive learning the MQDF recognizer in the updated discriminative feature space, the performance of the writer adaptation can be further improved.

Table2 Recognition performance of various writer adaptation approaches

Approach	Initial	Adapted	Error rate reduction
ILDA[12]	82.64%	87.08%	25.62%
ILDAMQDF	88.56%	91.01%	21.39%
IMQDF	88.56%	93.75%	45.38%
DIMQDF	88.56%	94.59%	52.71%

To examine the influence of writer adaptation approaches for total 30 different writers involved, figure 4 shows the error rate reduction of these four adaptation approaches for each particular writer. From the result given in figure 4, we can see that the proposed writer adaptation approaches significantly outperforms ILDA and ILDAMQDF approach for every particular writer.

In addition to evaluating the performance improvement of the IMQDF and DIMQDF based writer adaptation approaches on the writer-dependent dataset; we also examined the impact of the proposed writer adaptation approaches for the general purpose writer, since we don't expect the adaptation to the particular writer's handwriting style is at the cost of losing too much generality for other writer styles. The experimental result demonstrates that the accuracy loss on writer-independent ***CouchGB1*** dataset are only 0.02%, 0.18%, for IMQDF and DIMQDF based writer adaptation approaches respectively. It indicates that while the IMQDF and DIMQDF approaches can significantly improve the accuracy for different specific writers, they have very limited negative influence to general writer.

6. Conclusion

In this paper, we proposed a new writer adaption method to improve the recognition of handwritten Chinese character by incremental learning of MQDF model. We proposed the solution of IMQDF for the first time and then presented the DIMQDF approach. Based on IMQDF or DIMQDF, the writer adaptation was finally carried out by updating the MQDF classifier model adaptively. Experimental results demonstrated that both of the proposed IMQDF and DIMQDF based writer adaptation approaches can significantly increase the recognition accuracy for a specific writer, while at the same time, having limited negative impact for the general writer.

References

[1] C. Liu, S. Jaeger, M. Nakagawa, Online Recognition of Chinese Characters: The State-of-the-Art, *IEEE Trans. PAMI*, 26(2): 198-213, 2004.

[2] R. Plamondon and S.N. Srihari, On-line and off-line handwriting recognition: A comprehensive survey, *IEEE Trans. PAMI*, 22(1): 63-84., 2000.

[3] T. Long, L. Jin, A Novel Orientation Free Method for Online Unconstrained Cursive Handwritten Chinese Word Recognition, *ICPR 2008*: 1-4. 2008.

[4] K. Ding, G. Deng, L. Jin, An Investigation of Imaginary Stroke Technique for Cursive Online Handwriting Chinese Character Recognition, *ICDAR 2009*:531-535, 2009.

[5] D. Wang, C. Liu, et al., CASIA-OLHWDB1: A Database of Online Handwritten Chinese Characters, *ICDAR2009*: 1206-1210 2009.

[6] C.Liu, X.Zhou, Online Japanese character recognition using trajectory-based normalization and direction feature extraction, *IWFHR2006*: 217-222, 2006.

[7] R.A. Cole, J. Mariani et al., *Survey of the state of the art in human language technology*, Citeseer, 1997.

[8] S.D.Connel, and A.K. Jain, Writer Adaptation of Online Handwriting Models, *IEEE Trans. PAMI.*, 24(3): pp. 329-346, 2002.

[9] J.J. LaViola, and R.C.Zeleznik, A Practical Approach for Writer-Dependent Symbol Recognition Using a Writer-Independent Symbol Recognizer, *IEEE Trans. PAMI.*, 29(11): 1917-1926, 2007.

[10] Vuori, V. and Teknillinen Korkeakoulu, *Adaptive methods for on-line recognition of isolated handwritten characters*, Helsinki University of Technology, 2002

[11] Z. Huang, K. Ding, L. Jin, Writer Adaptive Online Handwriting Recognition Using Incremental Linear Discriminant Analysis, *ICDAR2009*: 91-95 2009.

[12] L. Jin, K. Ding, Z. Huang, Incremental Learning of LDA model for Chinese Writer Adaptation, *accepted by NeuroComputing 2010*.

[13] F.Kimura, K.Takashina et al, Modified quadratic discriminant functions and the application to Chinese character recognition, *IEEE Trans. PAMI.*, 9: 149-153, 1987.

[14] T. Long and L. Jin, Building Compact MQDF Classifier for Large Character Set Recognition by Subspace Distribution Sharing, *Pattern Recognition*, 41(9):pp. 909-913, 2008.

[15]R.O. Duda, P.E. Hart and D.G. Stork, *Pattern Classification*, 2001

[16] D. Tao, X. Li, et al., General Tensor Discriminant Analysis and Gabor Features for Gait Recognition, IEEE Trans. on PAMI, 9(10) (2007) 1700-1715.

[17] Z. Bai and Q. Huo, A Study On the Use of 8-Directional Features For Online Handwritten Chinese Character Recognition, *ICDAR 2005*: 232-2362005.

[18] SCUT-COUCH dataset, Website: http://www.hcii-lab.net/data/SCUTCOUCH/