Reduce the Storage of Incremental MQDF Classifier for Chinese Writer Adaptation

Yan Gao[1], Huiyun Mao[2], Lanwen Jin[1]
[1]*School of Electronic and Information Engineering,*
[2]*School of Computer Science,*
South China University of Technology,
Guangzhou, China
thegoldfishwang@163.com, huiyun.mao01@gmail.com, lianwen.jin@gmail.com

Abstract

In this paper, we propose a method to reduce the storage of the incremental MQDF classifier [1] for writer adaptation, through updating the dominant eigenvectors and eigenvalues of MQDF in the incremental discriminant subspace. Comparing to previous incremental MQDF algorithm, the proposed Compact Discriminant IMQDF (CDIMQDF) classifier only need to store the dominant eigenvectors and eigenvalues instead of covariance matrices, therefore save the storage cost greatly. The experimental results on SCUT COUCH2009 dataset show that, comparing with previous method, the new CDIMQDF can reduce storage significantly, while keeps comparable or even higher recognition accuracies.

KeyWords: Handwriting character recognition, Incremental learning, IMQDF, Writer adaptation

1. Introduction

Writer-adaptation is the process that adjusts a generic handwriting recognition system to fit the special writing style of a particular writer [2]. Such adaptation has the potential of significantly increasing accuracy for a particular user and can be very useful in building a high performance personalized handwriting character recognition system [2-4]. Modified quadratic discriminant function (MQDF) classifier has been widely used in handwriting Chinese character recognition due to its high classification performance [5,6,9]. Ding and Jin [1] proposed a basic Incremental MQDF algorithm (IMQDF) for writer adaptation, which can improve the recognition accuracy for a particular writer significantly. However, Ding and Jin's method needs to store the covariance matrix of each class and such storage requirement is huge for large scale classification task such as handwritten Chinese character recognition, which usually involved with thousands of classes. So this method is unpractical for real-world applications.

In this paper, we propose a new method to reduce the storage requirement of the IMQDF algorithm of [1]. The proposed method uses the dominant eigenvectors and eigenvalues for the incremental learning of covariance matrices instead of re-computing the covariance matrices directly, so that the storage requirement can be reduced significantly. Furthermore, based on the Incremental Linear Discriminative Analysis (ILDA) algorithm proposed in [7], we proposed a Compact Discriminant IMQDF (CDIMQDF) algorithm that can incrementally learn the MQDF in the discriminative feature space. Experiments carried on SCUT COUCH2009 dataset [8] show that, the proposed CDIMQDF classifier can reduce storage significantly, while keeps similar or even higher recognition accuracies compared with the previous DIMQDF method.

2. A Brief Introduction of MQDF and IMQDF

2.1. MQDF Classifier

The MQDF proposed by Kimura et al. [5] is obtained by smoothing eigenvalue of quadratic discriminant function (QDF), which achieves very good results in the field of Chinese characters recognition [6,9]. Under the assumption of equal a priori class probabilities, the discriminant function of a QDF classifier can be expressed as:

$$g_0(\mathbf{x},\omega_i) = (\mathbf{x}-\mathbf{m_i})^T \mathbf{\Sigma_i^{-1}}(\mathbf{x}-\mathbf{m_i}) + \log|\mathbf{\Sigma_i}| \quad (1)$$

where ω_i represents the i^{th} class, $\mathbf{m_i}$ and $\mathbf{\Sigma_i}$ denote the mean vector and covariance matrix of class ω_i, respectively.

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

$$\mathbf{\Sigma_i} = \mathbf{\Phi_i \Lambda_i \Phi_i^T} \quad (2)$$

where $\mathbf{\Lambda} = diag[\lambda_{i1}, \ldots, \lambda_{iD}]$ with λ_{ij}, $j = 1, \ldots, D$, being the eigenvalues (ordered in decreasing order) of $\mathbf{\Sigma_i}$, and $\mathbf{\Phi_i} = [\boldsymbol{\varphi}_{\mathbf{i1}}, \ldots, \boldsymbol{\varphi}_{\mathbf{iD}}]$ with $\boldsymbol{\varphi}_{\mathbf{ij}}$, $j = 1, \ldots, D$, being the ordered eigenvectors. $\mathbf{\Phi_i}$ is orthonormal (unitary) such that $\mathbf{\Phi_i^T \Phi_i = I}$, and D is the dimension of $\mathbf{x}$.

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

$$g_0(\mathbf{x}, \omega_i) = [\mathbf{\Phi_i^T (x - m_i)}]^T \mathbf{\Lambda_i^{-1} \Phi_i^T (x - m_i)} + \log|\mathbf{\Lambda_i}|$$

$$= \sum_{j=1}^{D} \frac{1}{\lambda_{ij}} [\boldsymbol{\varphi}_{\mathbf{ij}}^T (\mathbf{x} - \mathbf{m_i})]^2 + \sum_{j=1}^{D} \log \lambda_{ij} \quad (3)$$

By replacing the minor eigenvalues with a constant δ_i, the MQDF is obtained as

$$g_1(\mathbf{x}, \omega_i) = \frac{1}{\delta_i} \left\{ \|\mathbf{x} - \mathbf{m_i}\|^2 - \sum_{j=1}^{K} (1 - \frac{\delta_i}{\lambda_{ij}}) [\boldsymbol{\varphi}_{\mathbf{ij}}^T (\mathbf{x} - \mathbf{m_i})]^2 \right\}$$

$$+ \sum_{j=1}^{K} \log \lambda_{ij} + (D - K) \log \delta_i \quad (4)$$

where K denotes the number of dominant eigenvectors.

2.2. Incremental MQDF Classifier

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 the classes in $\boldsymbol{Y}$ may not be available in $\boldsymbol{X}$; in other words, some new classes may be introduced. Thus, the mixed observation set $\boldsymbol{Z}=\boldsymbol{X} \cup \boldsymbol{Y} = \{\mathbf{z_k}\}$ ($k=1,\ldots,L+N$) has $L+N$ samples in C classes, where $C \geq M$ and $C \geq P$. Without loss of generality, we assume that n_i of the original N 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 number of samples belonging to the C_i class is $s_i=n_i+l_i$. Let $\mathbf{m}_{x\mathbf{i}}$ and $\mathbf{\Sigma}_{x\mathbf{i}}$ represent the i^{th} class's mean vector and covariance matrix for the given observation set, and $\mathbf{m}_{y\mathbf{i}}$ and $\mathbf{\Sigma}_{y\mathbf{i}}$ represent the i^{th} class's mean vector and covariance matrix for the new observation set, respectively. Then the incremental learning of MQDF can be achieved by updating the class's mean vector $\mathbf{m}_{z\mathbf{i}}$ and covariance matrix $\mathbf{\Sigma}_{z\mathbf{i}}$ as follows [1]:

$$\mathbf{m}_{z\mathbf{i}} = \frac{1}{s_i} \sum_{k=1}^{s_i} \mathbf{z}_{\mathbf{i}k} = \frac{n_i \mathbf{m}_{x\mathbf{i}} + l_i \mathbf{m}_{y\mathbf{i}}}{n_i + l_i} \quad (5)$$

$$\mathbf{\Sigma_{zi}} = \frac{1}{s_i} \sum_{k=1}^{s_i} (\mathbf{z_{ik} z_{ik}^T} - \mathbf{m_{zi} m_{zi}^T})$$

$$= \frac{n_i}{n_i + l_i} \mathbf{\Sigma}_{x\mathbf{i}} + \frac{l_i}{n_i + l_i} \mathbf{\Sigma}_{y\mathbf{i}}$$

$$+ \frac{n_i l_i}{n_i + l_i} (\mathbf{m}_{y\mathbf{i}} - \mathbf{m}_{x\mathbf{i}})(\mathbf{m}_{y\mathbf{i}} - \mathbf{m}_{x\mathbf{i}})^T \quad (6)$$

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 updated dominant eigenvalues and eigenvectors. Finally, the MQDF classifier can be updated according (4).

3. Compact Discriminant Incremental MQDF (CDIMQDF)

The Incremental MQDF (IMQDF) proposed by [1] is simple easy to be implemented, yet can improve the accuracy for writer adaptation significantly. However, this method needs to store the original covariance matrix of each class, which results in involving with huge storage requirement. In this section, we present a new method to reduce the storage of IMQDF.

According to (2), suppose $\mathbf{\Lambda_{ik}} = diag[\lambda_{i1}, \ldots, \lambda_{iK}]$, where K denotes the dominant eigenvector number of MQDF classifier, and $\mathbf{\Lambda_{il}} = diag[\lambda_{i(K+1)}, \ldots, \lambda_{iD}]$, then $\mathbf{\Lambda_i}$ can be expressed by $\mathbf{\Lambda_{ik}}$ and $\mathbf{\Lambda_{il}}$ as:

$$\mathbf{\Lambda_i} = \begin{bmatrix} \mathbf{\Lambda_{ik}, 0} \\ \mathbf{0, \Lambda_{il}} \end{bmatrix} \quad (7)$$

Similarly, $\mathbf{\Phi_i}$ can be expressed as:

$$\mathbf{\Phi_i} = [\mathbf{\Phi_{ik}}, \mathbf{\Phi_{il}}] \quad (8)$$

where $\mathbf{\Phi_{ik}} = [\boldsymbol{\varphi}_{\mathbf{i1}}, \ldots, \boldsymbol{\varphi}_{\mathbf{iK}}]$ and $\mathbf{\Phi_{il}} = [\boldsymbol{\varphi}_{\mathbf{i(K+1)}}, \ldots, \boldsymbol{\varphi}_{\mathbf{iD}}]$.

According to (7) and (8) and (2), the covariance matrix $\mathbf{\Sigma_i}$ can be rewritten as:

$$\mathbf{\Sigma_i} = [\mathbf{\Phi_{ik}}, \mathbf{\Phi_{il}}] \begin{bmatrix} \mathbf{\Lambda_{ik}, 0} \\ \mathbf{0, \Lambda_{il}} \end{bmatrix} [\mathbf{\Phi_{ik}}, \mathbf{\Phi_{il}}]^{\mathbf{T}}$$

$$= \mathbf{\Phi_{ik} \Lambda_{ik} \Phi_{ik}^T} + \mathbf{\Phi_{il} \Lambda_{il} \Phi_{il}^T} \quad (9)$$

Inspired by the idea of MQDF [5], the minor eigenvalues become some kind of unstable noises and affect the classifier's robustness. By smoothing them, the covariance matrix $\mathbf{\Sigma_i}$ could be approximated as follow:

$$\mathbf{\Sigma_i} \approx \mathbf{\Phi_{ik}\Lambda_{ik}\Phi_{ik}}^T \quad (10)$$

According to (6) and (10), the incremental learning of the covariance matrix of IMQDF can be obtained as follow:

$$\mathbf{\Sigma_{zi}} = \frac{n_i}{n_i+l_i}\mathbf{\Phi_{ik}\Lambda_{ik}\Phi_{ik}}^T + \frac{l_i}{n_i+l_i}\mathbf{\Sigma}_{y\mathbf{i}} + \frac{n_i l_i}{n_i+l_i}(\mathbf{m}_{y\mathbf{i}}-\mathbf{m}_{x\mathbf{i}})(\mathbf{m}_{y\mathbf{i}}-\mathbf{m}_{x\mathbf{i}})^T \quad (11)$$

Equation (11) shows that the proposed method (we call it CIMQDF) uses the dominant eigenvectors and eigenvalues for incremental learning covariance matrix $\mathbf{\Sigma_{zi}}$ instead of using the original covariance matrix $\mathbf{\Sigma}_{x\mathbf{i}}$ directly, so that the storage requirement can be reduced significantly.

In Chinese handwriting recognition applications, Linear Discriminant Analysis (LDA) is usually used to transform features to a discriminative space [7,10], in order to reduce the feature dimension as well as to improve the recognition accuracy. In [1], a discriminant IMQDF (DIMQDF) algorithm was proposed to adaptively learn the MQDF recognizer in the updated LDA feature space based on ILDA [7]. Again, the DIMQDF algorithm suffers the same huge storage problem as IMQDF. Here, we will present a new method,CDIMQDF, to learn the eigenvectors and eigenvalues of the covariance matrix of MQDF in the LDA discriminant subspace incrementally.

Suppose the sample vector in original feature space is $\mathbf{x}$, and $\mathbf{x}_{lda}$ in LDA feature space, i.e., $\mathbf{x}_{lda} = \mathbf{W}_{lda}^T\mathbf{x}$. Then, the covariance matrix $\mathbf{\Sigma}_{LDA}$ in LDA feature space can be expressed as follow.

$$\begin{aligned}\mathbf{\Sigma_{LDA}} &= \sum(\mathbf{x}_{lda}-\mathbf{m}_{lda})(\mathbf{x}_{lda}-\mathbf{m}_{lda})^T \\ &= \mathbf{W}_{lda}^T(\sum(\mathbf{x}-\mathbf{m})(\mathbf{x}-\mathbf{m})^T)\mathbf{W}_{lda} \\ &= \mathbf{W}_{lda}^T\mathbf{\Sigma}\mathbf{W}_{lda}\end{aligned} \quad (12)$$

Suppose the covariance matrix in the original LDA space is $\mathbf{\Sigma_{LDA}}$, and $\mathbf{\Sigma'_{LDA}}$ in the updated LDA space. According to (13), $\mathbf{\Sigma'_{LDA}}$ can be rewritten as:

$$\begin{aligned}\mathbf{\Sigma'_{LDA}} &= \mathbf{W'}_{lda}^T\mathbf{\Sigma}\mathbf{W'}_{lda} \\ &= \mathbf{W'}_{lda}^T(\mathbf{W}_{lda}^T)^{-1}\mathbf{W}_{lda}^T\mathbf{\Sigma}\mathbf{W}_{lda}(\mathbf{W}_{lda})^{-1}\mathbf{W'}_{lda} \\ &= \mathbf{W'}_{lda}^T(\mathbf{W}_{lda}^T)^{-1}\mathbf{\Sigma_{LDA}}(\mathbf{W}_{lda})^{-1}\mathbf{W'}_{lda} \\ &= \mathbf{S}^T\mathbf{\Sigma_{LDA}}\mathbf{S}\end{aligned} \quad (13)$$

where $\mathbf{S} = (\mathbf{W}_{lda})^{-1}\mathbf{W'}_{lda}$.

Equation (13) shows that the covariance matrix $\mathbf{\Sigma'_{LDA}}$ can be obtained from $\mathbf{\Sigma_{LDA}}$ directly. Furthermore, according to (2), $\mathbf{\Sigma_{LDA}}$ and $\mathbf{\Sigma'_{LDA}}$ can be diagonalized as:

$$\begin{cases}\mathbf{\Sigma}_{LDA} = \mathbf{\Phi}_{LDA}\mathbf{\Lambda}_{LDA}\mathbf{\Phi}_{LDA}^T \\ \mathbf{\Sigma}'_{LDA} = \mathbf{\Phi}'_{LDA}\mathbf{\Lambda}'_{LDA}\mathbf{\Phi}'^T_{LDA}\end{cases} \quad (14)$$

Then, (13) can be rewritten as:

$$\begin{aligned}\mathbf{\Sigma}'_{LDA} &= \mathbf{S}^T\mathbf{\Sigma}_{LDA}\mathbf{S} \\ \mathbf{\Phi}'_{LDA}\mathbf{\Lambda}'_{LDA}\mathbf{\Phi}'^T_{LDA} &= \mathbf{S}^T\mathbf{\Phi}_{LDA}\mathbf{\Lambda}_{LDA}\mathbf{\Phi}^T_{LDA}\mathbf{S} \\ \mathbf{\Phi}'_{LDA}\mathbf{\Lambda}'_{LDA}\mathbf{\Phi}'^T_{LDA} &= (\mathbf{S}^T\mathbf{\Phi}_{LDA})\mathbf{\Lambda}_{LDA}(\mathbf{S}^T\mathbf{\Phi}_{LDA})^T\end{aligned} \quad (15)$$

According to (15), we can obtain the following Equations:

$$\begin{cases}\mathbf{\Lambda}'_{LDA} = \mathbf{\Lambda}_{LDA} \\ \mathbf{\Phi}'_{LDA} = \mathbf{S}^T\mathbf{\Phi}_{LDA}\end{cases} \quad (16)$$

Equation (16) indicates the incremental updating of the eigenvectors and eigenvalues respectively. We can see that when the LDA transformation matrix is updated, the eigenvalues of the MQDF are not changed, while the eigenvectors are transformed by $\mathbf{S}^T$.

4. Experiment

The dataset used is SCUT-COUCH2009 [8], which is consists of 11 subsets. We use the GB1 subset for training a baseline MQDF classifier, and segment isolated characters of 10 writers from the Word8888 subset for training/testing of the incremental models.

We compare the performance between the proposed CIMQDF/CDIMQDF algorithm and IMQDF/DIMQDF method of [1] in terms of classifier storage and recognition accuracy for particular writer.

There are 3755 classes of Chinese characters in GB1 subset, the dimension of original 8-directional features [9] we extract is 512, and the LDA dimension is 160, the storage requirements of the proposed method against that of [1] are given in table 1. It can be seen that the storage cost of our method is much smaller than that of [1]. It is worthwhile to note that storage of both IMQDF/DIQMDF of [1] is not changed when the dominant eigenvector number *K* of MQDF classifier reduced, because methods in [1] need to store the covariance matrix of every class, while the proposed method only need to store the dominant eigenvectors.

We examined the performance of the proposed CDIMQDF against DIMQDF of [1] using 10 sets of the incremental data. The dominant eigenvector number K of MQDF is set to 12 and 30 respectively. The results are given in Table 2 and Table 3.

From Table 2 and Table 3, it can be seen that when K is set to 12, recognition rate of the proposed Compact DIMQDF is slightly lower than DIMQDF of [1]. However, when K set to 30, the proposed method is slightly higher. It shows that when K is large enough, the recognition rate of the proposed method is even higher than previous method, while the storage is significantly smaller.

Table 1. Storage requirement of the IMQDF algorithm

method		Storage	
		K=12	K=30
The proposed	CIMQDF	88Mb	220Mb
	CDIMQDF	8.6Mb	21.5Mb
Method of [1]	IMQDF	3755Mb	3755Mb
	DIMQDF	367Mb	367Mb

Table 2. The recognition accuracy (%) of IMQDF when K set to 12

Writer	Initial MQDF	The proposed method	Method of [1]
		CompactDIMQDF	DIMQDF
1	94.29%	96.89%	**97.02%**
2	95.92%	98.43%	98.43%
3	88.96%	95.53%	**95.57%**
4	89.84%	97.04%	**97.46%**
5	86.57%	93.22%	**94.13%**
6	90.97%	98.17%	**98.42%**
7	94.88%	97.78%	97.78%
8	92.67%	95.99%	**96.55%**
9	90.19%	95.01%	**95.42%**
10	93.91%	96.32%	**96.69%**
Average	91.82%	96.44%	**96.75%**

Table 3. The recognition accuracy (%) of IMQDF when K set to 30

Writer	Initial MQDF	The proposed method	Method of [1]
		CompactDIMQDF	DIMQDF
1	94.79%	**97.77%**	97.03%
2	96.00%	**98.44%**	98.43%
3	89.12%	**95.70%**	95.60%
4	89.99%	97.48%	97.48%
5	87.27%	**94.62%**	94.25%
6	90.98%	98.44%	98.44%
7	95.35%	**97.79%**	97.78%
8	93.06%	96.61%	96.61%
9	91.92%	**95.58%**	95.54%
10	94.56%	**97.03%**	96.77%
Average	92.30%	**96.95%**	96.79%

5. Conclusion

In the paper, we propose a new Compact Discrimant IMQDF (CDIMQDF) algorithm for Chinese writer adaptation. The proposed CDIMQDF algorithm is based on transforming the eigenvectors and eigenvalues of MQDF incrementally instead of re-computing covariance matrix, therefore, the storage requirement is reduced significantly comparing with previous DIMQDF algorithm. Meanwhile, experiments show that CDIMQDF can keeps similar or even higher recognition accuracies compared with the previous DIMQDF method.

Acknowlegdement

This research was partially supported by NSFC (no. 61075021), GDNSF（No. S2011020000541) and the Fundamental Research Funds for the Central Universities of China(no.2012ZP0002).

References

[1]. K. Ding, L. Jin, *Incremental MQDF Learning for Writer Adaptive Handwriting Recognition*. ICFHR2010. pp. 559-564. 2010.

[2]. S.D. Connel, and A.K. Jain, *Writer Adaptation for Online Handwriting Recognition*, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 3, pp.329-346, March 2002.

[3]. W. Kienzle, and K. Chellapilla, *Personalized handwriting recognition via biased regularization*, Proc. International Conf. Machine learning 2006, pp. 457–464, 2006.

[4]. P. Haluptzok, *Allograph Based Writer Adaptation for Handwritten Character Recognition*, International Workshop Frontiers in Handwriting Recognition 2006, pp. 423–428, 2006

[5]. F. Kimura, K. Takashina, et al., *Modified quadratic discriminant functions and the application to Chinese character recognition*, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 9, no. 1, pp. 149–153, 1987.

[6]. T. Long, L. Jin, *Building compact MQDF classifier for large character set recognition by subspace distribution sharing*, Pattern Recognition, Vol. 41, No. 9, pp. 2916-2925, 2008.

[7]. L. Jin, K. Ding, Z. Huang, *Incremental Learning of LDA model for Chinese Writer Adaptation*, NeuroComputing, vol. 73, no.10-12, pp.1614-1623, 2010.

[8]. L. Jin, Y. Gao, et al, *SCUT-COUCH2009–A Comprehensive Online Unconstrained Chinese Handwriting Database and Benchmark Evaluation*, International Journal on Document Analysis and Recognition", vol. 14, no.1, pp53-64,2011.

[9]. C.-L. Liu, *High Accuracy Handwritten Chinese Character Recognition Using Quadratic Classifiers with Discriminative Feature Extraction*, ICPR2006, vol.2, pp. 942–945, 2006.

[10]. C. Liu, S. Jaeger, M. Nakagawa, Online Recognition of Chinese Characters: The State-of-the-Art, IEEE Trans. Pattern Anal. Machine Intell., vol.26, no.2, pp198-213,2004.