



Online Handwriting Tibetan Character Recognition Based on Two-Dimensional Discriminant Locality Alignment

Zhengqi Cai^{1,2} and Weilan Wang^{1(✉)}

¹ Key Laboratory of China's Ethnic Languages and Information Technology of Ministry of Education, Northwest Minzu University, Lanzhou 730000, Gansu, People's Republic of China

caizhengqi@l26.com, wangweilan@xbmu.edu.cn

² College of Mathematics and Computer Science, Northwest Minzu University, Lanzhou 730000, Gansu, People's Republic of China

Abstract. Discriminant Locality Alignment (DLA) has been successfully applied in handwriting character recognition. In this paper, a new manifold based subspace learning algorithm, which is called Two-dimensional Discriminant Locality Alignment (2DDLDA) algorithm, is proposed for online handwriting Tibetan character recognition (OHTCR). The proposed algorithm integrates the idea of DLA and two-dimensional feature extraction algorithm. At first, extracting direction feature matrix and edge feature matrix of Tibetan character respectively, they are together formed original feature matrix. Then, in part optimization stage, for each character sample, a local patch is built by the given sample and its neighbors, and an object function is designed to preserve local discriminant information. Third, in whole alignment stage, the alignment trick is used to align all part optimizations to the whole optimization. The projection matrix can be obtained by solving a standard eigen-decomposition problem. Finally, a SMQDF classifier is used training and recognition. Experimental results demonstrate that 2DLDA is superior to LDA and IMLDA in terms of recognition accuracy. In addition, 2DLDA can overcome the matrix singular problem and small sample size problem in OHTCR.

Keywords: Online handwriting recognition · Tibetan character recognition
Two-dimensional discriminant locality alignment (2DDLDA) · Subspace learning

1 Introduction

With the acceleration of Tibetan information process, the demand of Tibetan character recognition system is becoming more and more prominent. At present, handwriting Tibetan character recognition (HTCR) has made great progress in both research and practical application [1–6]. However, the recognition of Tibetan character is different from handwriting recognition of other languages, it poses a special challenge due to a complex structure, wide varieties in writing style, a large character set and many instances of highly similar characters. Figure 1 illustrates some samples of handwriting

Tibetan character. Unconstrained online HTCR is still an open problem remaining to be solved, for it is still challenging to reach high recognition rate.

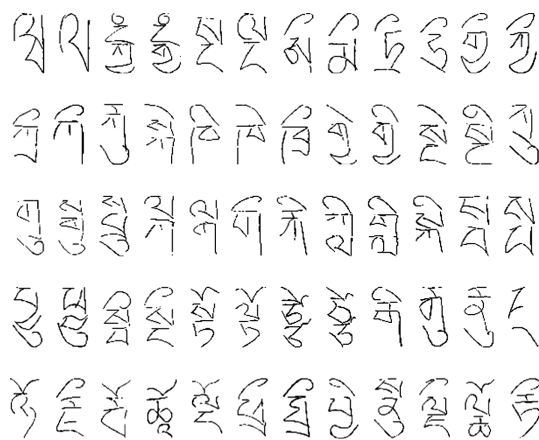


Fig. 1. Some samples of handwriting Tibetan character

As shown in Fig. 1, many Tibetan characters are almost identical to other characters except for only a small difference. However, the small difference can be lost during the feature extraction process. So the discriminate information extraction is crucial for improvement of the recognition performance. In OHTCR, dimensionality reduction is the process of transforming data from a high dimensional space to a low dimensional space to reveal the intrinsic structure of the distribution of data. It plays a crucial role in the field of computer vision and pattern recognition as a way of dealing with the “curse of dimensionality”. In past decades, a large number of dimensionality reduction algorithms have been proposed and studied. Among them, principal components analysis (PCA) [7] and Fisher’s linear discriminant analysis (LDA) [8] are two of the most popular linear dimensionality reduction algorithms.

PCA maximizes the mutual information between original high dimensional Gaussian distributed data and projected low dimensional data. PCA is optimal for reconstruction of Gaussian distributed data. However, it is not optimal for classification problems. LDA overcomes this shortcoming by utilizing the class label information. It finds the projection directions that maximize the trace of the between-class scatter matrix and minimize the trace of the within-class scatter matrix simultaneously. However, LDA is only a suboptimal model which suffers from the class separation problem. The objective of LDA can be formulated as maximizing the sum of all the pairwise distances between different classes, which will overemphasize the large distance of the already well-separated classes, and confuse the small distance classes that are close in the original feature space. Li and Yuan [9] proposed a new method of feature extraction using two-dimensional linear discriminant analysis (2DDLDA), and directly uses the matrix to extract the discriminant feature without a vectorization

procedure, it has a great advantage over the one-dimensional method in calculation and processing efficiency.

Zhang [10] proposed a local linear dimensionality reduction algorithm called discriminative locality alignment (DLA). The DLA takes into account the locality of samples, deals with the nonlinearity of the samples distribution, and preserves the discriminability of classes as well. However, the DLA algorithm is based on the vector space, and the data must be vectorized during calculation, which destroys the spatial distribution characteristics and structure information of the data. Based on the stability and effectiveness of DLA algorithm in recognition performance, in this paper, we combine the idea of DLA algorithm with two-dimensional feature extraction algorithm, and proposes a two-dimensional discriminative locality alignment (2DDLA) algorithm to improve the recognition performance in OHTCR.

The rest of paper is organized as follows. Section 2 introduces two-dimensional discriminant locality alignment (2DDLA) algorithm for extracting discriminative features for OHTCR and details the basic formulation. Section 3 introduces SMQDF classifier. We perform experiments in Sect. 4 to show the effectiveness of the proposed method and Sect. 5 gives concluding remark.

2 Two-Dimensional Discriminative Locality Alignment

2DDLA aims to extract discriminative information from patches. To achieve this goal, one patch is first built for each sample. Each patch includes a sample and its within-class nearest samples and its between-class nearest samples. Then an objective function is designed to preserve the local discriminative information of each patch. Finally, all the part optimizations are integrated together to form a global coordinate according to the alignment trick. The projection matrix can be obtained by solving a standard Eigen decomposition problem.

2.1 Part Optimization

Suppose we have a set of samples $X = [X_1, X_2, \dots, X_N]$ from C different classes, $X_i \in \mathbb{R}^{m \times n}$. For a given sample X_i , we select k_1 nearest neighbors from the samples of the same class with X_i and name them as the neighborhoods of a same class: $X_{i^1}, \dots, X_{i^{k_1}}$, we also select k_2 nearest neighbors from samples of different classes with X_i , and name them as neighborhoods of different classes: $X_{i_1}, \dots, X_{i_{k_2}}$. By putting them together, we can build the local patch for X_i as $\Pi_i = [X_i, X_{i^1}, \dots, X_{i^{k_1}}, X_{i_1}, \dots, X_{i_{k_2}}]$. For each patch, the corresponding output in the low-dimensional space is denoted by $\Gamma_i = [Y_i, Y_{i^1}, \dots, Y_{i^{k_1}}, Y_{i_1}, \dots, Y_{i_{k_2}}]$.

In the low-dimensional space, we expect that distances between the given sample and its within-class samples are as small as possible, while distances between the given sample and its between-class samples are as large as possible. So we have

$$\arg \min_{Y_i} \sum_{j=1}^{k_1} \|Y_i - Y_{i_j}\|_F^2 \quad (1)$$

$$\arg \max_{Y_i} \sum_{p=1}^{k_2} \|Y_i - Y_{i_p}\|_F^2 \quad (2)$$

where $\|\cdot\|_F$ denotes the Frobenius norm for a matrix.

Since the patch formed by the local neighborhood can be regarded approximately linear, we formulate the part discriminator by using the linear manipulation as follows:

$$\arg \min_{Y_i} \left(\sum_{j=1}^{k_1} \|Y_i - Y_{i_j}\|_F^2 - \beta \sum_{p=1}^{k_2} \|Y_i - Y_{i_p}\|_F^2 \right) \quad (3)$$

where β is a scaling factor ($\beta \in [0,1]$). The coefficients vector is defined as:

$$W_i = \left[\overbrace{1, \dots, 1}^{k_1}, -\overbrace{\beta, \dots, \beta}^{k_2} \right] \quad (4)$$

Then the Eq. (3) reduces to:

$$\begin{aligned} & \arg \min_{Y_i} \left(\sum_{j=1}^{k_1} \|Y_i - Y_{i_j}\|_F^2 W_i(j) + \sum_{p=1}^{k_2} \|Y_i - Y_{i_p}\|_F^2 W_i(p + k_1) \right) \\ &= \arg \min_{Y_i} \left(\sum_{j=1}^{k_1+k_2} \|Y_{F_i\{j\}} - Y_{F_i\{j+1\}}\|_F^2 W_i(j) \right) \\ &= \arg \min_{Y_i} \text{tr}(\Gamma_i(L \otimes I_n)(\text{diag}(W_i) \otimes I_n)(R \otimes I_n)(\Gamma_i)^T) \\ &= \arg \min_{\Gamma_i} \text{tr}(\Gamma_i T_i \Gamma_i^T) \end{aligned} \quad (5)$$

where $\text{tr}(\cdot)$ denotes the trace operator, the operator \otimes denotes the Kronecker product of matrix, $F_i = \{i, i^1, \dots, i^{k_1}, i_1, \dots, i_{k_2}\}$ is the index set for the i th patch, $e_{k_1+k_2} = [1, \dots, 1]^T \in \mathbb{R}^{k_1+k_2}$, $I_{k_1+k_2}$ is a $(k_1+k_2) \times (k_1+k_2)$ identity matrix, $R = [-e_{k_1+k_2}, I_{k_1+k_2}]^T$, $L = \begin{bmatrix} -e_{k_1+k_2}^T \\ I_{k_1+k_2} \end{bmatrix}$, and

$$T_i = (L \otimes I_n)(\text{diag}(W_i) \otimes I_n)(R \otimes I_n) \quad (6)$$

2.2 Whole Alignment

After the part optimization step, we unify the optimizations together as a whole one by assuming that the coordinate for the i 'th patch $\Gamma_i = [Y_i, Y_{i_1}, \dots, Y_{i^{k_1}}, Y_{i_1}, \dots, Y_{i_{k_2}}]$ is selected from the global coordinate $\Gamma = [Y_1, Y_2, \dots, Y_N]$, such that $\Gamma_i = \Gamma S_i$, where $S_i \in \mathbb{R}^{(N \times n) \times ((k_1+k_2+1) \times n)}$ is the selection matrix and an entry is defined as follows:

$$(S_i)_{pq} = \begin{cases} I_n & \text{if } p = F_i(q) \\ 0_n & \text{else} \end{cases} \quad (7)$$

Then Eq. (7) can be rewritten as

$$\arg \min_{\Gamma} \text{tr}(\Gamma S_i T_i S_i^T \Gamma^T) \quad (8)$$

By summing over all the part optimizations described as Eq. (8), we can obtain the whole alignment as

$$\begin{aligned} & \arg \min_{\Gamma} \sum_{i=1}^N \text{tr}(\Gamma S_i T_i S_i^T \Gamma^T) \\ &= \arg \min_{\Gamma} \text{tr}((\Gamma \sum_{i=1}^N S_i T_i S_i^T) \Gamma^T) \\ &= \arg \min_{\Gamma} \text{tr}(\Gamma L \Gamma^T) \end{aligned} \quad (9)$$

where $L = \sum_{i=1}^N S_i T_i S_i^T \in R^{N \times N}$ is the alignment matrix.

To obtain the linear and orthogonal projection matrix W , such as $Y = W^T X$, Eq. (9) is deformed as follows:

$$\arg \min_W \text{tr}(W^T X L X^T W), \text{ s.t. } W^T W = I \quad (10)$$

The transformation matrix W that minimizes the objective function is given by the minimum eigenvalue solution to the standard eigenvalue problem,

$$X L X^T P = \lambda P \quad (11)$$

3 SMQDF

3.1 MQDF

MQDF [11] classifier's discriminate function is formulated as

$$\begin{aligned} f(Y, \omega_j) &= \sum_{i=1}^k \frac{[(Y - \mu_j)^T \zeta_i^{(j)}]^2}{\lambda_i^j} + \sum_{i=k+1}^m \frac{[(Y - \mu_j)^T \zeta_i^{(j)}]^2}{\lambda} \\ &+ \sum_{i=1}^k \log \lambda_i^{(j)} + \sum_{i=k+1}^m \log \lambda \quad j = 1, 2, \dots, C \end{aligned} \quad (12)$$

where, Y is input feature vector, m is line number of feature matrix, μ_j denotes the mean vector of class ω_j , $\lambda_i^{(j)}$ and $\zeta_i^{(j)}$ denote the i th larger eigenvalue and the corresponding eigenvector of the covariance matrix of class ω_j , respectively. k is the number of dominant principal eigenvectors that are kept in MQDF, λ is experiment parameter.

We can obtain the classified result based on the following criterion: If $f(Y, \omega_i) = \min_{1 \leq j \leq C} f(Y, \omega_j)$ (C is class number), then we believe that input pattern Y belongs to the ω_i class.

3.2 SMQDF

MQDF classifier is widely used in the area of character recognition. However, it only applies to feature vector, and it is not appropriate for feature matrix. For this reason, SMQDF (second modified quadratic discriminate function) classifier is generated by improving MQDF classifier, its discriminate function as shown in the follow formula (13). We take it as a baseline classifier.

$$f(Y, \omega_j) = \sum_{i=1}^{m-1} \frac{((Y - \mu_j)^T \zeta_i^{(j)})^T ((Y - \mu_j)^T \zeta_i^{(j)})}{\lambda_i^{(j)}} + \frac{((Y - \mu_j)^T \zeta_m^{(j)})^T ((Y - \mu_j)^T \zeta_m^{(j)})}{\lambda} + \sum_{i=1}^{m-1} \log \lambda_i^{(j)} + \log \lambda \quad j = 1, 2, \dots, C \quad (13)$$

where, Y is feature matrix, m is positive integer, When classifies, Y belongs to the class whose $f(Y, \omega_i)$ is minimum. To compensate for the estimation error of parameters on limited training samples, the minor eigenvalues are replaced with a constant λ . It can be set to a class-independent constant or class-dependent constant. Here we set λ to be class-independent for its superior performance. λ is computed by

$$\lambda = \frac{1}{c * d} \sum_{j=1}^c \sum_{i=1}^d \lambda_i^{(j)} \quad (14)$$

4 Experiment Results

4.1 Experiment Data

We evaluated the recognition performance of 2DDLA on a databases of handwritten Tibetan characters, collected by our group, contains the handwriting samples of 7240 characters, 5000 samples per class [12]. In order to reduce the computing cost, we only selected 562 frequently used characters, 2000 samples per class for training and 500 samples per class for testing.

For character image pre-processing and feature extraction, we adopt the same methods as in [6]. Each character image is normalized to 48×96 pixels, the

directional features and edge features are extracted. The resulting 60×12 feature matrix is projected onto a 12×12 subspace learned by 2DDLA, then the baseline classifier SMQDF is designed on this 12×12 subspace.

4.2 Choice of Parameters for 2DLDA

Since the parameters setup for 2DDLA is essential for its performance, we carried out the 2DDLA parameter optimization experiments before for OHTCR. In the model of 2DDLA, there are three parameters: k_1 , k_2 and β , where k_1 is the number of the samples from identical class in the given patch, k_2 is the number of the samples from other classes in the same given patch, and parameter β is the scale parameter. In order to find a proper range for the dominant parameters k_1 , k_2 and β in 2DDLA, we will investigate the effects of the three model parameters on the recognition rates in the validation phase based on our collected database.

Suppose n is the training sample number in each class ($n = 2000$), N is the total training sample number ($N = 562 \times 2000 = 1024000$), and C is class number ($C = 562$). Then, k_1 and k_2 could be chosen in the range of $[1, n-1]$ and $[0, N-n]$ respectively. Therefore, $1 \leq k_1 \leq 1999$, $0 \leq k_2 \leq 1022000$, and $0 \leq \beta \leq 1$.

To evaluate the effects of the three model parameters, firstly, we analyze the effect of the scale parameter β , by fixing patch building parameters k_1 and k_2 to arbitrary values. For a given pair parameters k_1 and k_2 , we can obtain the recognition rate curve with respect to β , as shown in Fig. 1. Base on the figures, we observe that the best recognition rates are obtained when β is neither too small nor too larger.

Secondly, we analyze the effects of patch building parameters k_1 and k_2 , by fixing scale parameter $\beta = 0.1$. When we vary k_1 and k_2 simultaneously, the best recognition rate with the corresponding to β can be acquired. Table 1 shows that the details of the best recognition rate. Figure 2 shows that best recognition rate with the corresponding $k_1 = 50$ and $k_2 = 300$ in this experiment (Fig. 3).

Table 1 shows that, the best combination of k_1 , k_2 , and β are $k_1 = 50$, $k_2 = 300$, $\beta = 0.1$ and $k_1 = 100$, $k_2 = 300$, $\beta = 0.3$, with the corresponding accuracy 99.38%. Considering the computing cost, in the following experiments, we use the best combination of k_1 , k_2 , and β is 50, 300, 0.1 respectively.

4.3 Evaluation Experiments

To evaluation the performance of 2DDLA in SHCCR, we compare the performance of 2DDLA, LDA, IMLDA [13] and 2DLDA in terms of recognition rate over SMQDF classifier [6]. The experimental results are summarized in Table 2. We can see that the proposed method obtains higher top 1 and top 10 recognition rate than other method.

From Table 2, it is shown that the recognition rates of 2DDLA are significantly higher than that of IMLDA and 2DLDA respectively. It also shows the discriminate information extraction performance is very competitive in OHTCR.

To illustrate the effects of the 2DDLA, Fig. 4 shows some sample that are mis-recognized by IMLDA and 2DLDA, but can be corrected by 2DDLA.

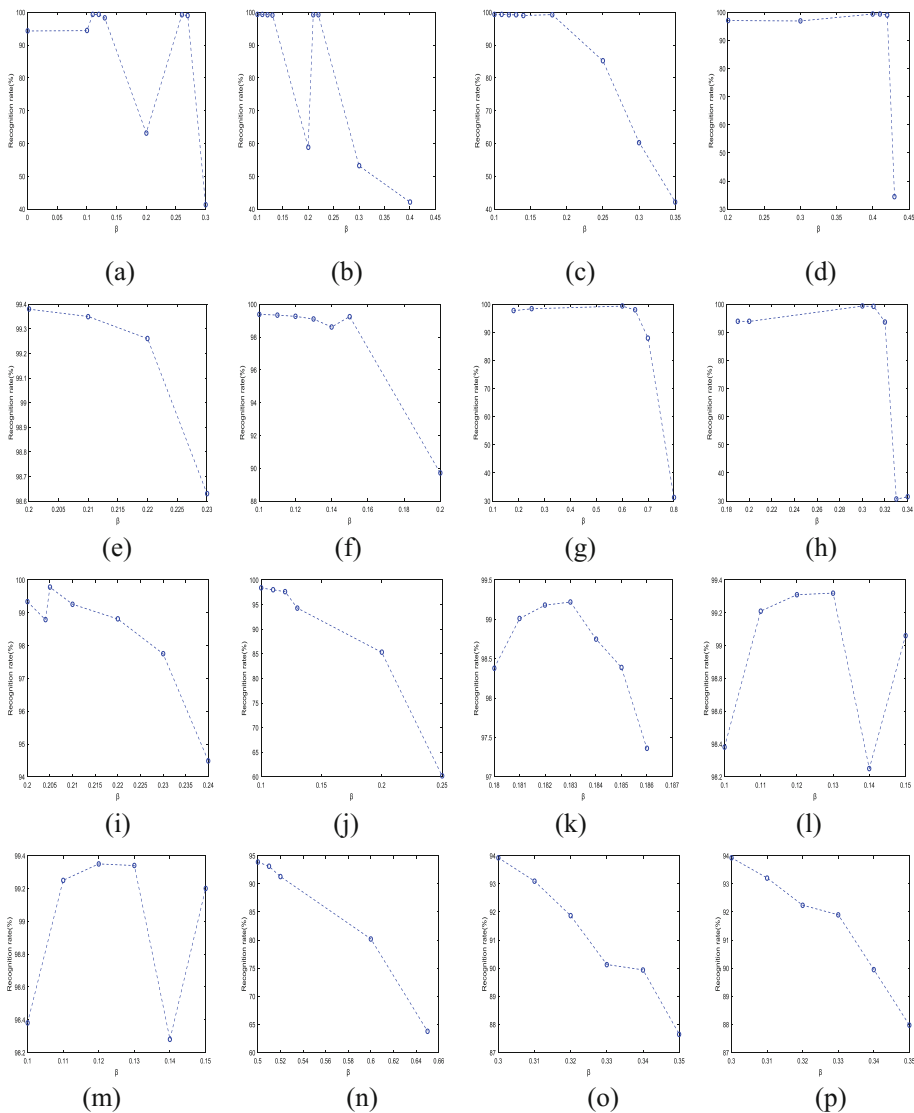


Fig. 2. For a given pair parameters k_1 and k_2 , the best recognition rate with the corresponding to β . (a) $k_1 = 32, k_2 = 100$, (b) $k_1 = 32, k_2 = 200$, (c) $k_1 = 32, k_2 = 300$, (d) $k_1 = 50, k_2 = 100$, (e) $k_1 = 50, k_2 = 200$, (f) $k_1 = 50, k_2 = 300$, (g) $k_1 = 80, k_2 = 100$, (h) $k_1 = 80, k_2 = 200$, (i) $k_1 = 80, k_2 = 300$, (j) $k_1 = 100, k_2 = 300$, (k) $k_1 = 100, k_2 = 500$, (l) $k_1 = 100, k_2 = 800$, (m) $k_1 = 100, k_2 = 990$, (n) $k_1 = 300, k_2 = 500$, (o) $k_1 = 300, k_2 = 800$, (p) $k_1 = 300, k_2 = 990$, (q) $k_1 = 500, k_2 = 800$, (r) $k_1 = 500, k_2 = 1000$, (s) $k_1 = 500, k_2 = 1200$, (t) $k_1 = 500, k_2 = 1500$, (u) $k_1 = 800, k_2 = 1500$, (v) $k_1 = 800, k_2 = 2500$, (w) $k_1 = 800, k_2 = 3800$, (x) $k_1 = 1000, k_2 = 300$

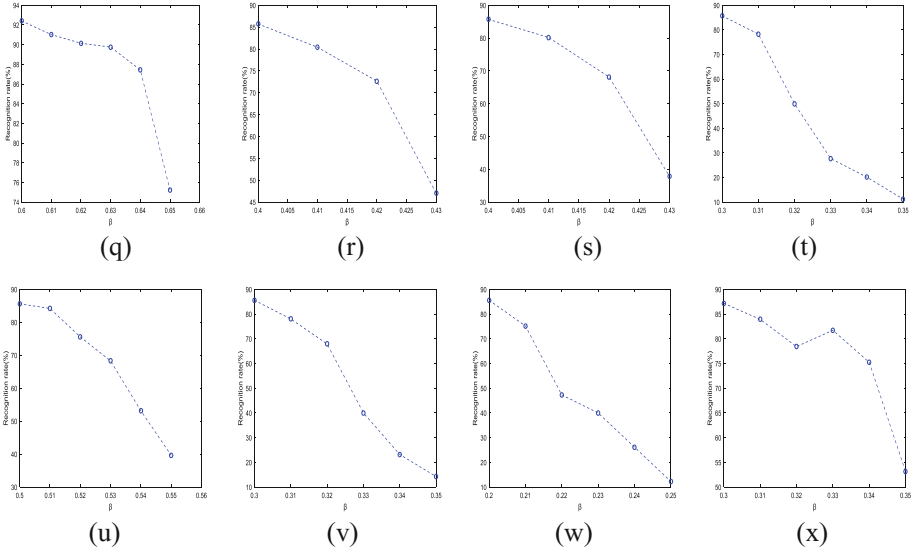


Fig. 2. (continued)

Table 1. Best recognition rate (%) with the corresponding k_1 and k_2 .

K_1	K_2	β	Recognition rates
32	100	0.11	99.36
32	200	0.11	99.36
32	300	0.11	99.36
50	100	0.41	99.37
50	200	0.21	99.37
50	300	0.10	99.38
80	100	0.61	99.34
80	200	0.31	99.34
80	300	0.21	99.34
100	300	0.30	98.38
100	500	0.18	99.18
100	800	0.12	99.31
100	990	0.12	99.35
300	500	0.50	93.93
300	800	0.30	93.93
300	990	0.30	93.93
500	800	0.60	92.41
500	1000	0.40	85.73
500	1200	0.40	85.72
500	1500	0.30	85.73
800	1500	0.50	85.62
800	2500	0.3	85.62
800	3800	0.2	85.62
1000	3000	0.3	87.19

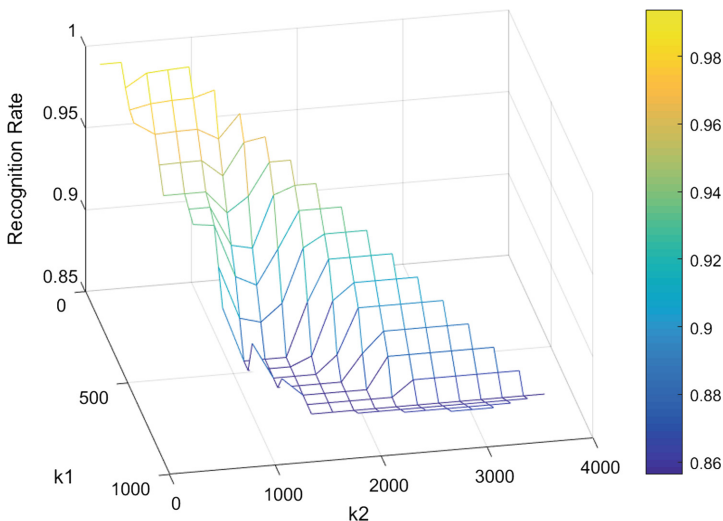


Fig. 3. Recognition rate vs. k_1 and k_2

Table 2. Best recognition rates (%) of three methods.

Methods	Top 1	Top 10
IMLDA	55.56	92.73
2DLDA	83.73	98.21
2DDLDA	85.9	99.38

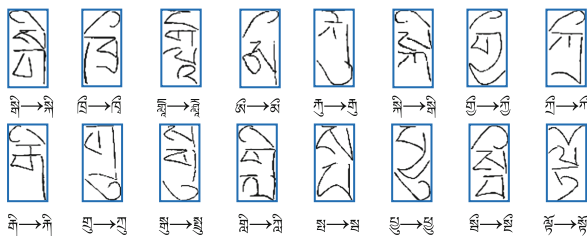


Fig. 4. Some misrecognized by SMQDF, but corrected by compound distance method

5 Conclusion

In this paper, we present a new method, called two-dimensional discriminative locality alignment (2DDLDA). Compare with IMLDA and 2DLDA, the proposed method has better recognition rate. It inherits all the advantages of DLA and can overcome the matrix singular problem and small sample size problem in OHTCR.

Acknowledgment. This work was supported by the Fundamental Research Funds for the Central University of Northwest Minzu University (No. 31920170142), the Program for Leading Talent of State Ethnic Affairs Commission, the National Science Foundation (No. 61375029), and supported by the Gansu Provincial first-class discipline program of Northwest Minzu University.

References

1. Wang, W., Ding, X., Qi, K.: Study on similar character in Tibetan character recognition. *J. Chin. Inf. Process.* **16**(4), 60–65 (2002)
2. Wang, H., Ding, X.: Multi-font printed Tibetan character recognition. *J. Chin. Inf. Process.* **17**(6), 47–52 (2003)
3. Wang, W., Qian, J., Duojie, Z., Ma, M., Qi, K., Duo, L., et al.: A method of online handwritten Tibetan characters recognition. State Intellectual Property Office of the Peoples Republic of China (2011). ZL200910128595.8
4. Huang, H., Da, F., Hang, X.: Wavelet transform and gradient direction based feature extraction method for offline handwritten Tibetan letter recognition. *J. Southeast Univ.* **30**(1), 27–31 (2014)
5. Ma, L.L., Wu, J.: A Tibetan component representation learning method for online handwritten Tibetan character recognition. In: Proceedings of the 14th ICFHR, pp. 317–322 (2014)
6. Wang, W., Qian, J., Wang, D., Duojie, Z.: Online handwriting recognition of Tibetan characters based on the statistical method. *J. Commun. Comput.* **8**(2011), 188–200 (2011)
7. Jolliffe, I.T.: *Principal Component Analysis*. Springer, New York (1986). <https://doi.org/10.1007/978-1-4757-1904-8>
8. Fisher, R.A.: The use of multiple measurements in taxonomic problems. *Ann. Eugen.* **7**, 179–188 (1936)
9. Li, M., Yuan, B.: 2D-LDA: a statistical linear discriminant analysis for image matrix. *Pattern Recognit. Lett.* **26**(5), 527–532 (2005)
10. Zhang, T., Tao, D., Li, X., et al.: Patch alignment for dimensionality reduction. *IEEE Trans. Knowl. Data Eng.* **10**(2), 433–439 (2009)
11. Kimura, F., Takashina, K., Tsuruoka, S., Miyake, Y.: Modified quadratic discriminant functions and its application to Chinese character recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **9**(1), 149–153 (1987)
12. Wang, W., Lu, X., Cai, Z., Shen, W., Fu, J., Caike, Z.: Online handwritten sample generated based on component combination for Tibetan-Sanskrit. *J. Chin. Inf. Process.* **31**(5), 64–73 (2017)
13. Qian, J., Wang, J.: A novel approach for online handwriting recognition of Tibetan characters. In: Proceedings of IMECS 2010, pp. 1–4 (2010)