International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:13, No:11, 2019

Bidirectional Discriminant Supervised Locality Preserving Projection for Face Recognition

Yiqin Lin, Wenbo Li

Abstract—Dimensionality reduction and feature extraction are of crucial importance for achieving high efficiency in manipulating the high dimensional data. Two-dimensional discriminant locality preserving projection (2D-DLPP) and two-dimensional discriminant supervised LPP (2D-DSLPP) are two effective two-dimensional projection methods for dimensionality reduction and feature extraction of face image matrices. Since 2D-DLPP and 2D-DSLPP preserve the local structure information of the original data and exploit the discriminant information, they usually have good recognition performance. However, 2D-DLPP and 2D-DSLPP only employ single-sided projection, and thus the generated low dimensional data matrices have still many features. In this paper, by combining the discriminant supervised LPP with the bidirectional projection, we propose the bidirectional discriminant supervised LPP (BDSLPP). The left and right projection matrices for BDSLPP can be computed iteratively. Experimental results show that the proposed BDSLPP achieves higher recognition accuracy than 2D-DLPP, 2D-DSLPP, and bidirectional discriminant LPP (BDLPP).

Keywords—Face recognition, dimension reduction, locality preserving projection, discriminant information, bidirectional projection.

I. Introduction

The problem of dimension reduction has received a lot of attention in areas such as face recognition [20], [24], micro-array data analysis [1], [12], text classification [19], information retrieval [21], and pattern recognition [3], [11], [14], where high dimensional data are required to deal with. Dimensionality reduction transforms the high dimensional data into a low dimensional subspace, and effectively reduces data dimensionality for efficient data processing tasks.

Many dimensionality reduction methods, such as principal component analysis (PCA) [25], linear discriminant analysis (LDA) [14], and locality preserving projection (LPP) [16], [35], [34], [36], have been developed over the past few decades. PCA aims to find the projection directions by maximizing variance of features in the low dimensional subspace. LDA has been one of the popular techniques in classification. The basic idea of LDA is to calculate the optimal discriminant vectors so that the ratio of the between-class scatter and the within-class scatter is maximized. The optimal discriminant vectors of LDA can be computed by solving a generalized eigenvalue problem involving scatter matrices. In contrast to PCA and LDA, locality preserving projection (LPP) [16] aims to preserve the local structure information of the original data, which is more important in many classification applications where the nearest neighbor classifier is used.

Y. Lin is with the Institute of Computational Mathematics, School of Science, Hunan University of Science and Engineering, Yongzhou 425199, P. R. China (Corresponding author; e-mail: yqlin@aliyun.com).
W. Li is with the School of Mathematics and Computational Science, Xiangtan University, Xiangtan 411105, P. R. China (e-mail: wenboli88@aliyun.com).

For two-dimensional data such as images, several bidirectional projection methods, such as the generalized low-rank approximation method (GLRAM) [30], tensor subspace analysis (TSA), [17], and discriminant TSA (DTSA)[26], have been proposed. These methods aim to find two subspaces for two-sided projection. GLRAM computes the left and right projections by minimizing the reconstruction error, and it only preserves the global Euclidean structure of the image data. However, TSA and DTSA can preserve the local structure information of the original data.

In this paper, we propose the bidirectional discriminant supervised LPP (BDSLPP) by combining the discriminant supervised LPP with the bidirectional projection. Similarly to other bidirectional projection methods, BDSLPP also iteratively computes the left and right projection matrices. Since BDSLPP exploits the bidirectional projection, it yields higher compression ratio than 2D-DSLPP. Two experiments on face recognition are conducted to evaluate the effectiveness of BDSLPP. Experimental results show that the BDSLPP proposed in this paper achieves higher recognition accuracy than 2D-DLPP, 2D-DSLPP, and BDLPP.

Throughout this paper, we adopt the following notations: I_l denotes an identity matrix of order l, and $\otimes$ represents the Kronecker product of the matrices. $\|\cdot\|$ denotes the Frobenius norm for a matrix, i.e., $\|A\| = \sqrt{\sum_i \sum_j A_{ij}^2}$.

The structure of the paper is as follows. In Section II, we briefly review 2D-DSLPP. In Section III, we propose the BDSLPP by combining the discriminant supervised LPP with the bidirectional projection. Section IV is devoted to numerical experiments. Some concluding remarks are provided in Section V.

II. Two-Dimensional Discriminant Supervised LPP

Given a set of N image data

$$\mathcal{X} = \{X_1, X_2, \cdots, X_N\},$$

where $X_i \in \mathbb{R}^{L_1 \times L_2}$.

In two-dimensional methods for dimension reduction and feature extraction of facial image data, we wish to find a low-dimensional analogue

$$\mathcal{Y} = \{Y_1, Y_2, \cdots, Y_N\},$$

where $Y_i \in \mathbb{R}^{l_1 \times L_2}$ with $l_1 \leq L_1$. We hope that $\mathcal{Y}$ is a faithful representation of $\mathcal{X}$ in some sense. Formally, we are seeking a projection matrix $U \in \mathbb{R}^{L_1 \times l_1}$ so that

$$Y_i = U^T X_i, \quad i = 1, 2, \cdots, N.$$

We assume that the given data set $\mathcal{X}$ is partitioned into C classes as

$$\mathcal{X} = \Big\{X_1^{(1)}, X_2^{(1)}, \cdots, X_{N_1}^{(1)}, X_1^{(2)}, X_2^{(2)}, \cdots, X_{N_2}^{(2)}, \cdots, X_1^{(C)}, X_2^{(C)}, \cdots, X_{N_C}^{(C)}\Big\},$$

where $X_i^{(c)}$ means the ith sample in the cth class, N_c is the number of samples in the cth class, and $\sum_{c=1}^{C} N_c = N$ is satisfied.

The Locality Preserving Projections (LPP) [16] is a graph-based projective technique. It projects the data so as to preserve a certain affinity graph constructed from the data. By extending LPP for 2D data, two-dimensional discriminant supervised LPP (2D-DSLPP) [27] has been proposed for dimension reduction and feature extraction of face recognition.

In order to construct the objective function, we need to define two similarity matrices $W, T \in \mathbb{R}^{N \times N}$. The entry of W, T, is, respectively, defined as follows:

$$W_{ij} = \begin{cases} \exp(-\|X_i - X_j\|^2/t), & X_i, X_j \text{ in same class}, \\ 0, & \text{otherwise}, \end{cases}$$

$$T_{ij} = \begin{cases} \exp(-\|X_i - X_j\|^2/t), & \text{otherwise}, \\ 0, & X_i, X_j \text{ in same class}, \end{cases}$$

where t is a positive parameter which can be determined empirically.

It is easy to see that W is a block diagonal matrix with C blocks and the size of the cth block being the number N_c of samples in the cth class i.e., $W = \mathrm{diag}(W_1, W_2, \cdots, W_C)$.

With two similarity matrices W, T, then $\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}\|Y_i - Y_j\|$ can be used to measure the within-class closeness in the projected data space, while $\sum_{i=1}^{N}\sum_{j=1}^{N} T_{ij}\|Y_i - Y_j\|$ measures the between-class separation. Ideally, the optimal transformation U should minimize the within-class distance and maximize simultaneously, by which the low dimensional data $Y_i = U^T X_i$ are easier to be distinguished.

In 2D-DSLPP [27], the optimal transformation matrix U is determined by the following optimization problem

$$\max_U \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} T_{ij}\|Y_i - Y_j\|^2}{\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}\|Y_i - Y_j\|^2}. \quad (1)$$

Define two diagonal matrices $D_T = \mathrm{diag}(\alpha_1, \alpha_2, \cdots, \alpha_N)$ and $D_W = \mathrm{diag}(\beta_1, \beta_2, \cdots, \beta_N)$ with

$$\alpha_i = \sum_{j=1}^{N} T_{ij} \quad \text{and} \quad \beta_i = \sum_{j=1}^{N} W_{ij}.$$

Let $L_T = D_T - T$ and $L_W = D_W - W$ be Laplacian matrices. Then, the above optimization problem (1) can be equivalently formulated as

$$\max_U \frac{\mathrm{tr}(U^T(X(L_T \otimes I_{L_2})X^T)U)}{\mathrm{tr}(U^T(X(L_W \otimes I_{L_2})X^T)U)}. \quad (2)$$

The optimal projection matrix U can be obtained by computing the l_1 eigenvectors of the generalized eigenvalue problem

$$\left[X(L_T \otimes I_{L_2})X^T\right] u = \lambda \left[X(L_W \otimes I_{L_2})X^T\right] u$$

corresponding to the largest l_1 eigenvalues, see [27].

We outline the procedure of 2D-DSLPP in Algorithm 1, which is used for computing the optimal transformation matrix U.

III. Bidirectional Discriminant Supervised LPP

In this section, we will improve the 2D-DSLPP by using the bidirectional projection technique. In bidirectional methods for dimension reduction and feature extraction of facial image data matrices, it aims to find two projection matrices $U \in \mathbb{R}^{L_1 \times l_1}, V \in \mathbb{R}^{L_2 \times l_2}$ with $l_1 \le L_1$ and $l_2 \le L_2$ so that the original data matrices X_i are transformed into

$$Y_i = U^T X_i V.$$

In bidirectional discriminant supervised LPP (BDSLPP), we seek to find the left and right transformation matrices U, V by solving the following optimization problem

$$\max_{U,V} \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} T_{ij}\|Y_i - Y_j\|^2}{\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}\|Y_i - Y_j\|^2}. \quad (3)$$

As pointed out in [27], the numerator part of the objective function in (3) denotes the global variance on the manifold in low-dimensional subspace, while the denominator part of the objective function is a measure of nearness of samples from the same class. Therefore, by maximizing the objective function, the samples from the same class are transformed into data points close to each other and samples from different classes are transformed into data points far from each other.

Let

$$Y = [Y_1, Y_2, \cdots, Y_N], \quad \widetilde{Y} = [Y_1^T, Y_2^T, \cdots, Y_N^T].$$

Then, it is easy to verify that

$$\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}\|Y_i - Y_j\|^2 = \mathrm{tr}(Y(L_W \otimes I_{l_2})Y^T),$$

or

$$\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}\|Y_i - Y_j\|^2 = \mathrm{tr}(\widetilde{Y}(L_W \otimes I_{l_1})\widetilde{Y}^T).$$

Define

$$P_U = [X_1^T U, X_2^T U, \cdots, X_N^T U],$$
$$P_V = [X_1 V, X_2 V, \cdots, X_N V].$$

Then,

$$Y = U^T P_V, \quad \widetilde{Y} = V^T P_U.$$

So, we obtain

$$\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}\|Y_i - Y_j\|^2 = \mathrm{tr}(U^T(P_V(L_W \otimes I_{l_2})P_V^T)U), \quad (4)$$

International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:13, No:11, 2019

Algorithm 1: **2D-DSLPP**

Input: a set of N sample matrices $\{X_i\}_{i=1}^N$ with class label information and the dimension l_1
Outputtransformation matrix U

1) Form the matrix $M_T = X(L_T \otimes I_{L_2})X^T$;
2) Form the matrix $M_W = X(L_W \otimes I_{L_2})X^T$;
3) Compute the l_1 eigenvectors $\{u_i\}_{i=1}^{l_1}$ of the pencil (M_T, M_W) corresponding to the largest l_1 eigenvalues;
4) Set $U = [u_1, u_2, \cdots, u_{l_1}]$.

or

$$\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}\|Y_i - Y_j\|^2 = \mathrm{tr}(V^T(P_U(L_W \otimes I_{l_1})P_U^T)V). \tag{5}$$

By (4) or (5), the optimization problem (3) can be equivlently rewritten as the following optimization problem

$$\max_{U,V} \frac{\mathrm{tr}\left(V^T P_U(L_T \otimes I_{l_1})P_U^T V\right)}{\mathrm{tr}\left(V^T P_U(L_W \otimes I_{l_1})P_U^T V\right)}, \tag{6}$$

or

$$\max_{U,V} \frac{\mathrm{tr}\left(U^T P_V(L_T \otimes I_{l_2})P_V^T U\right)}{\mathrm{tr}\left(U^T P_V(L_W \otimes I_{l_2})P_V^T U\right)}. \tag{7}$$

Clearly, from the equivalence between the maximization problem (3) and the optimization problem (6) or (7), we have the following results.

Theorem 1: Let U and V be the solution of the maximization problem (3). Then,

1) For a given U, V consists of the l_2 eigenvectors of the generalized eigenvalue problem

$$\left[P_U(L_T \otimes I_{l_1})P_U^T\right] v = \lambda \left[P_U(L_W \otimes I_{l_1})P_U^T\right] v$$

corresponding to the largest l_2 eigenvalues.

2) For a given V, U consists of the l_1 eigenvectors of the generalized eigenvalue problem

$$\left[P_V(L_T \otimes I_{l_2})P_V^T\right] u = \lambda \left[P_V(L_W \otimes I_{l_2})P_V^T\right] u$$

corresponding to the largest l_1 eigenvalues.

From Theorem 1, an iterative algorithm for the computation of the transformation matrices U and V results. The algorithm is outlined in Algorithm 2.

IV. Experimental Results

In order to evaluate the performance of the proposed BDSLPP algorithm, two well-known face image databases, i.e., ORL[1] and Yale[2], are used in the experiments. We compare the BDSLPP algorithm with 2D-DLPP [33], 2D-DSLPP [27], and BDLPP [26]. In the experiments, the nearest neighbor classifier is used to classify the transformed results of samples obtained using different methods.

[1]http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html
[2]http://cvc.yale.edu/projects/yalefaces/yalefaces.html

TABLE I
Recognition Accuracy (%) on ORL Database (mean±std)

	2D-DLPP	2D-DSLPP	BDLPP	BDSLPP
2 Train	76.78±3.02	76.88±3.06	79.22±3.19	80.22±3.14
3 Train	84.61±2.42	84.82±1.86	88.21±2.64	89.54±2.59
4 Train	88.42±1.05	89.00±1.54	91.29±2.56	93.04±1.42
5 Train	90.55±1.89	91.75±1.77	93.30±2.08	95.40±1.80
6 Train	91.75±2.44	92.63±2.49	94.56±2.75	96.06±1.59
7 Train	92.36±2.98	93.08±1.98	95.64±2.19	96.42±2.19
8 Train	93.50±1.75	95.13±1.45	97.16±1.41	97.88±1.32

A. Experiment on the ORL Database of Face Images

The ORL database contains 400 images of 40 individuals. Each individual has 10 images, which were taken at different time, different lighting conditions, different facial expressions, and different accessories (glasses/no glasses). The sample images of one individual from the ORL database are shown in Fig 1.

We randomly select i ($i = 2, 3, \cdots, 7, 8$) samples of each individual for training, and the remaining ones are used for testing. Based on the training set, the project matrices are obtained by 2D-DLPP, 2D-DSLPP, BDLPP, and BDSLPP. Then all the testing samples are projected to generate the low-dimensional samples, which will be recognized by using the nearest neighbor classifier. We repeat the process 10 times and calculate the mean and standard deviation of recognition rates.

In our experiments, the parameters l_1 and l_2 in BDLPP and BDSLPP are set to be 10, and l_1 in 2D-DLPP and 2D-DSLPP are also set to be 10. The parameter t in defining the similar matrices T and W is set to 1. The mean and standard deviation of recognition accuracy of 10 runs of tests of four algorithms are presented in Table I. It shows that for all methods, the recognition increases with the increase in training sample size. Moreover, the bidirectional methods have higher recognition accuracy than the one-directional methods, and BDSLPP outperforms 2D-DLPP, 2D-DSLPP, and BDLPP.

Since the project matrices U, V in BDLPP and BDSLPP are iteratively computed. To investigate the effect of the number of iteration, we present their recognition accuracy curves versus the number of iteration in Fig 2, 3, and 4. It is shown that to obtain high recognition accuracy, only one iteration step is required, and more iteration would lower recognition rate.

Algorithm 2: **BDSLPP**

Input: a set of N sample matrices $\{X_i\}_{i=1}^{N}$ with class label information, l_1, l_2
Outputleft and right transformation matrices U and V

1) Initialize U with an identity matrix;
2) Until convergence Do:
   2.1 Form the matrix $M_T^{(U)} = P_U(L_T \otimes I_{l_1})P_U^T$;
   2.2 Form the matrix $M_W^{(U)} = P_U(L_W \otimes I_{l_1})P_U^T$;
   2.3 Compute the l_2 eigenvectors $\{v_i\}_{i=1}^{l_2}$ of the pencil $(M_T^{(U)}, M_W^{(U)})$ corresponding to the largest l_2 eigenvalues.
   2.4 Set $V = [v_1, v_2, \cdots, v_{l_2}]$;
   2.5 Form the matrix $M_T^{(V)} = P_V(L_T \otimes I_{l_2})P_V^T$;
   2.6 Form the matrix $M_W^{(V)} = P_V(L_W \otimes I_{l_2})P_V^T$;
   2.7 Compute the l_1 eigenvectors $\{u_i\}_{i=1}^{l_1}$ of the pencil $(M_T^{(V)}, M_W^{(V)})$ corresponding to the largest l_1 eigenvalues;
   2.8 Set $U = [u_1, u_2, \cdots, u_{l_1}]$.
   End Do

Fig. 1 Sample images for one individual of the ORL database

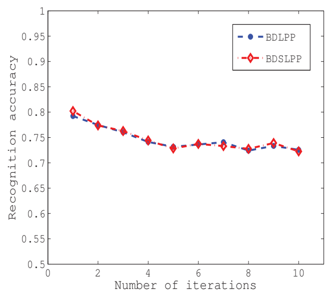

Fig. 2 Recognition accuracy versus the number of iteration for the ORL database with 2 training samples

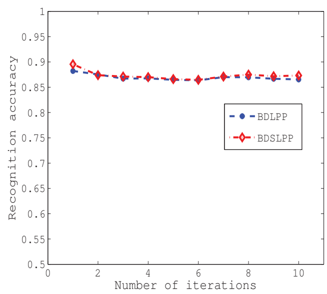

Fig. 3 Recognition accuracy versus the number of iteration for the ORL database with 3 training samples

B. Experiment on the Yale Database

The Yale face database contains 165 gray-scale images from 15 individuals where each individual has 11 images.

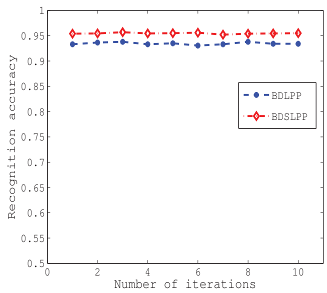

Fig. 4 Recognition accuracy versus the number of iteration for the ORL database with 5 training samples

These facial images have variations in lighting conditions (left-light, center-light, right-light), facial expressions (normal, happy, sad, sleepy, surprised, and wink), and with/without glasses. The 11 sample images of one individual from the Yale database are shown in Fig 5.

Fig. 5 Sample images for one individual of the Yale database

As in the previous experiments, the parameters l_1 and l_2 are set to 10, and t is set to 1. The mean and standard deviation of recognition accuracy of 10 runs of tests for the Yale database are presented in Table II. Clearly, BDSLPP performs better than 2D-DLPP, 2D-DSLPP, and BDLPP for the Yale database. The recognition accuracy curves versus the number of iteration are shown in Figs. 6, 7, and 8. Clearly, one iteration step produces the highest recognition rate.

International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:13, No:11, 2019

TABLE II
RECOGNITION ACCURACY (%) ON YALE DATABASE (MEAN±STD)

	2D-DLPP	2D-DSLPP	BDLPP	BDSLPP
2 Train	41.78±5.04	41.56±4.46	41.41±5.97	42.81±6.10
3 Train	48.25±4.43	47.33±6.12	56.17±3.40	57.17±4.09
4 Train	58.14±3.90	60.52±4.54	63.10±4.00	64.76±4.20
5 Train	64.22±3.40	64.67±3.90	65.78±4.51	66.78±4.63
6 Train	70.00±4.54	70.27±4.99	70.13±5.59	73.33±3.84
7 Train	72.17±3.43	72.50±3.95	74.29±4.65	75.00±2.70
8 Train	75.78±4.38	76.00±5.62	76.89±5.87	77.64±6.27

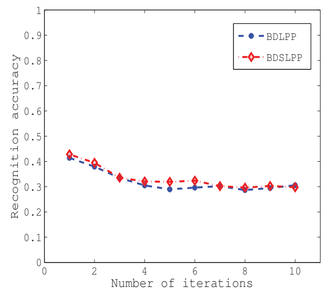

Fig. 6 Recognition accuracy versus the number of iteration for the Yale database with 2 training samples

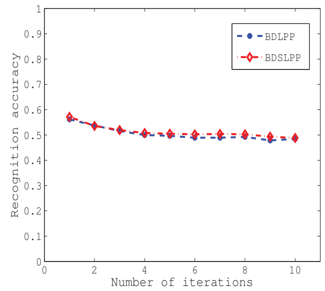

Fig. 7 Recognition accuracy versus the number of iteration for the Yale database with 3 training samples

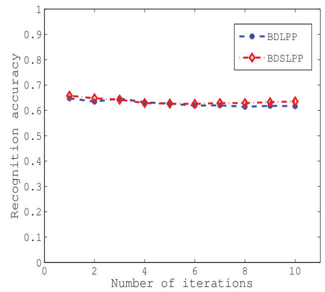

Fig. 8 Recognition accuracy versus the number of iteration for the Yale database with 5 training samples

V. CONCLUSION

In this paper, we propose a bidirectional discriminant supervised locality preserving projection (BDSLPP) method for face recognition. The left and right projection matrices of the proposed method can be iteratively computed. Experimental results show that BDSLPP has higher recognition accuracy than 2D-DLPP, 2D-DSLPP, and BDLPP. Moreover, it is shown that only one iteration step is required to obtain high recognition accuracy, and more iteration would lower recognition rate.

ACKNOWLEDGMENTS

Yiqin Lin is supported by the Natural Science Foundation of Hunan Province under grant 2017JJ2102, the Academic Leader Training Plan of Hunan Province in University, the Talent Support Plan of Hunan University of Science and Engineering, and the Construct Program of the Key Discipline in Hunan University of Science and Engineering.

REFERENCES

[1] P. Baldi and G. W. Hatfield, *DNA Microarrays and Gene Expression: From Experiments to Data Analysis and Modeling*, Cambridge, 2002.
[2] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, *Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 19 (1997), pp. 711–720.
[3] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
[4] L. Chen, H. M. Liao, M. Ko, J. Lin, and G. Yu, *A new LDA-based face recognition system which can solve the small sample size problem*, Pattern Recognition, 33 (2000), pp. 1713–1726.
[5] S. B. Chen, H. F. Zhao, M. Kong, and B. Luo, *2D-LPP: a two-dimensional extension of locality preserving projections*, Neurocomputing, 70 (2007), pp. 912–921.
[6] W. K. Ching, D. Chu, L. Z. Liao, and X. Wang, *Regularized orthogonal linear discriminant analysis*, Pattern Recognition, 45 (2012), pp. 2719–2732.
[7] D. Chu and S. T. Goh, *A new and fast implementation for null space based linear discriminant analysis*, Pattern Recognition, 43 (2010), pp. 1373–1379.
[8] ———, *A new and fast orthogonal linear discriminant analysis on undersampled problems*, SIAM Journal on Scientific Computing, 32 (2010), pp. 2274–2297.
[9] D. Chu, S. T. Goh, and Y. S. Hung, *Characterization of all solutions for undersampled uncorrelated linear discriminant analysis problems*, SIAM Journal on Matrix Analysis and Applications, 32 (2011), pp. 820–844.
[10] D. Q. Dai and P. C. Yuen, *Regularized discriminant analysis and its application to face recognition*, Pattern Recognition, 36 (2003), pp. 845–847.
[11] R. Q. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, John Wiley and Sons, second ed., 2001.
[12] S. Dudoit, J. Fridlyand, and T. P. Speed, *Comparison of discrimination methods for the classification of tumors using gene expression data*, Journal of the American Statistical Association, 97 (2002), pp. 77–87.
[13] J. H. Friedman, *Regularized discriminant analysis*, Journal of the American Statistical Association, 84 (1989), pp. 165–175.
[14] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, CA: Academic, San Diego, second ed., 1990.
[15] Y. Guo, T. Hastie, and R. Tibshirani, *Regularized linear discriminant analysis and its application in microarray*, Biostatistics, 8 (2007), pp. 86–100.
[16] X. He and P. Niyogi, *Locality preserving projections*, Advances in Neural Information Processing Systems, 16 (2004), pp. 153–160.
[17] ———, *Tensor subspace analysis*, Advances in Neural Information Processing Systems, 18 (2005).
[18] P. Howland and H. Park, *Generalizing discriminant analysis using the generalized singular value decomposition*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 26 (2004), pp. 995–1006.
[19] A. K. JAIN AND R. C. DUBES, *Algorithms for Clustering Data*, Prentice Hall, 1988.
[20] Z. Jin, J. Y. Yang, Z. S. Hu, and Z. Lou, *Face recognition based on the uncorrelated discriminant transformation*, Pattern Recognition, 34 (2001), pp. 1405–1416.
[21] G. Kowalski, *Information Retrieval Systems: Theory and Implementation*, Kluwer Academic Publishers, 1997.

International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:13, No:11, 2019

[22] M. Li and B. Z. Yuan, *2D-LDA: a statistical linear discriminant analysis for image matrix*, Pattern Recognition Letters, 26 (2005), pp. 527–532.
[23] C. X. Ren and D. Q. Dai, *Bilinear lanczos components for fast dimensionality reduction and feature extraction*, Pattern Recognition, 43 (2010), pp. 3742–3752.
[24] D. L. Swets and J. Weng, *Using discriminant eigenfeatures for image retrieval*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 18 (1996), pp. 831–836.
[25] M. Turk and A. Pentland, *Eigenfaces for recognition*, Journal of Cognitive Neuroscience, 3 (1991), pp. 71–86.
[26] S. J. Wang, C. G. Zhou, N. Zhang, X. J. Peng, Y. H. Chen, and X. Liu, *Face recognition using second-order discriminant tensor subspace analysis*, Neurocomputing, 74 (2011), pp. 2142–2156.
[27] Y. Xu, G. Feng, and Y. Zhao, *One improvement to two-dimensional locality preserving projection method for use with face recognition*, Neurocomputing, 73 (2009), pp. 245–249.
[28] J. Yang, D. Zhang, A. F. Frangi, and J. Y. Yang, *Two-dimensional PCA: a new approach to appearance-based face representation and recognition*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 26 (2004), pp. 131–137.
[29] J. Ye, *Characterization of a family of algorithms for generalized discriminant analysis on undersampled problems*, Journal of Machine Learning Research, 6 (2005), pp. 483–502.
[30] ———, *Generalized low rank approximations of matrices*, Machine Learning, 61 (2005), pp. 167–191.
[31] J. Ye, R. Janardan, C. H. Park, and H. Park, *An optimization criterion for generalized discriminant analysis on undersampled problems*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 26 (2004), pp. 982–994.
[32] J. Ye and Q. Li, *A two-stage linear discriminant analysis via QR-decomposition*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 27 (2005), pp. 929–941.
[33] W. Yu, *Two-dimensional discriminant locality preserving projections for face recognition*, Pattern Recognition Letters, 30 (2009), pp. 1378–1383.
[34] W. Yu, X. Teng, and C. Liu, *Face recognition using discriminant locality preserving projections*, Image and Vision Computing, 24 (2006), pp. 239–248.
[35] Z. Zheng, F. Yang, and W. Tan, *Gabor feature-based face recognition using supervised locality preserving projections*, Signal Processing, 87 (2007), pp. 2473–2483.
[36] L. Zhu and S. Zhu, *Face recognition based on orthogonal discriminant locality preserving projections*, Neurocomputing, 70 (2007), pp. 1543–1546.