

Design and Implementation of Robust 2D Face Recognition System for Illumination Variations

Kiran P.Gaikwad
GHRIET-Pune, India

V.M.Wadhai
MAE-Pune, India

Prasad S.Halgaonkar
MITCOE-Pune, India

Santosh Kumar
SITRC-Nashik, India

ABSTRACT

Illumination variation is a challenging problem in face recognition research area. Same person can appear greatly different under varying lighting conditions. This paper consists of Face Recognition System which is invariant to illumination variations. Face recognition system which uses Linear Discriminant Analysis (LDA) as feature extractor have Small Sample Size (SSS). It consists of implementation of Feature Extraction Module using Two Dimensional Maximum Margin Criteria which removes “Small Sample Size (SSS)” problem present in existing Face Recognition System.

In statistical pattern recognition, high dimensionality is a major cause of the practical limitations of many pattern recognition technologies. Large number of features degrades the performance of classifiers if the number of training samples is small relative to the number of features. This is referred to as the “peaking phenomenon”, is caused by the “curse of dimensionality”. The dimensionality of images after feature extraction for storing feature database is reduced in this paper. The input for the system is images from standard database. Features are extracted of given images using Two Dimensional Maximum Margin Criteria from row as well as column direction. Finally matching is done using Euclidean Distance for test image and stored image features.

Index terms - Face Recognition, Feature Extraction, Small Sample Size (SSS), Automatic Feature Selection Method.

I. INTRODUCTION

Feature extraction is the key step in face recognition [1], [2], [3]. Principle component analysis (PCA) [4], [5] and linear discriminant analysis (LDA) [6] are well-known feature extraction and data representation techniques widely used in the areas of pattern recognition for feature extraction and dimensionality reduction. As an unsupervised method, PCA looks for a subspace where the samples have the minimum reconstruction error. In contrast to PCA, LDA takes the class labels into consideration and searches for a subspace where the samples from the same class are as compact as possible, and meanwhile, the samples from the different classes are as far as possible.

However, PCA and LDA are both based on vectors. In other words, if we use them to deal with the image recognition problem, the first step is to transform 2-D face image matrices into 1-D image vectors column by column or row by row. As a result, the data vectors lie

in a high-dimensional space, and in general, it far exceeds the number of data samples. In machine learning terminology, it is said to have high dimension and small sample size (SSS) [7], [8], [9].

In face recognition, LDA usually encounters the SSS problem in which the within-class covariance matrix becomes singular, and thus, the traditional LDA algorithm fails. To address this problem, a number of approaches have been suggested, including regularized discriminant analysis [10], [11], inverse Fisher [12], [13], weighted piecewise LDA [14], pseudo-inverse LDA [15], null space method [16], direct LDA (DLDA) [17], and maximal margin criterion (MMC) [18]. Among them, the most popular one is PCA plus LDA approach (PCA + LDA) which uses PCA for dimension reduction prior to performing LDA. However, this method cannot completely solve the SSS problem, and the discarded null space of the within-class scatter matrix S_w may cause the performance deterioration.

2-D discriminant analysis (2DDA) has become an interesting technique in face recognition. Compared with 1-D discriminant analysis (1DDA), 2DDA has four significant advantages: 1) It is easier to evaluate the between class, within-class, and total scatter matrices. 2) It consumes less time to extract discriminative features. 3) It preserves the underlying data structures for data analysis while 1DDA ignores the underlying structure. 4) In real applications, the within-class scatter matrix is usually non-singular [19], so that the SSS problem does not exist here.

The methods for 2DDA can be categorized as one-directional feature extraction approaches and two-directional feature extraction approaches. Two-dimensional PCA (2DPCA) [20], [21] and 2-D LDA (2DLDA) [22] are two representative methods which extract features from one direction. 2DPCA and 2DLDA significantly reduce the feature extraction time than traditional PCA and LDA; however, they generally use more discriminative features and take longer to test than 1-D environments. 2DPCA and 2DLDA are essentially working in the row direction of images. By simultaneously combining row and column directions, two-directional 2-D PCA ((2D)2PCA) [23] and two-directional 2-D LDA((2D)2LDA) [24], [25], [26] are proposed. These methods can make more efficient representation than one-directional feature extraction approaches. Similar to PCA, 2DPCA and (2D)2PCA are unsupervised learning algorithms, thus, not always optimal for classification task. In the aforementioned 2-D methods, they cannot provide an automatic strategy to choose discriminant vectors and some parameters must be validated.

How to rapidly extract efficient discriminant features is a research emphasis for face recognition. Jing et al. [27] proposed a nonparametric two-directional feature extraction method: 2-D Fisher-face which extracted face discriminative features by automatically selecting 2-D principal components and discriminant vectors. However, their method may lose some important classification information in feature selection stage.

2DDA is essentially working in the row direction, or column direction, or both directions of images while 1DDA is working in the each pixel of images. 2DDA methods extract features based on the rows or the columns of all images; thus, it is possible that the extracted features still contain some redundant information.

To further reduce the dimension by 2DDA, naturally, the combination of 2DDA and 1DDA is studied. Ye et al. developed a 2DLDA plus LDA approach (2DLDA + LDA), where LDA is preceded by 2DLDA. Zuo et al. [28] also put forward a bidirectional PCA plus LDA method (BDPCA + LDA), where LDA is performed in the BDPCA subspace. However, as most of the 2-D methods, in their approaches, two parameters must be validated by manual.

In this paper, motivated by the success and the power of 2DDA in pattern classification tasks, we propose a new 2DDA method: Automatic Feature Selection Method which is based on MMC. Feature extraction method based on MMC criterion is robust, stable, and efficient.

II. FEATURE EXTRACTION USING TWO DIRECTIONAL MAXIMUM MARGIN CRITERIA

Given a face image sample set $A = \{A_1, A_2, \dots, A_N\}$, where each element represents a face image expressed by an $m \times n$ matrix and N is the total number of training samples. Suppose that $\omega_1, \omega_2, \dots, \omega_c$ are c known pattern classes, and n_i ($i = 1, \dots, c$) is the number of the samples of the i^{th} class. Suppose that M_i ($i = 1, 2, \dots, c$) denotes the mean matrix of the i^{th} class and M denotes the total mean matrix of A . The 2-D between-class scatter matrix R_b , the 2-D within-class scatter matrix R_w in row direction are defined as,

$$R_b = 1/N \sum_{i=1}^c n_i (M_i - M)(M_i - M)^t \dots \dots \dots 1$$

$$R_w = 1/N \sum_{i=1}^c \sum_{j \in \omega_i} n_i (A_j - M_i)(A_j - M_i)^t \dots \dots \dots 2$$

In another direction, define the 2-D between-class scatter matrix C_b , the 2-D within-class scatter matrix C_w

$$C_b = 1/N \sum_{i=1}^c n_i (M_i - M)(M_i - M)^t \dots \dots \dots 3$$

$$C_w = 1/N \sum_{i=1}^c \sum_{j \in \omega_i} n_i (A_j - M_i)(A_j - M_i)^t \dots \dots \dots 4$$

A. Algorithm for Feature Extraction:

Step 1: Calculate the Mean M_i of each class separately for Red(R), Green (G), Blue (B) components

Step 2: Calculate the Global Mean M for Red(R), Green (G), Blue (B) components separately.

Step 3: Calculate R_b and R_w which are $n \times n$ matrices and C_b and C_w which are $m \times m$ matrices for R, G, and B components.

Step 4: Compute two projectors from two directions.

a) Compute the row projector W_{MMC_R} by solving the following optimization problem for R, G, and B components:

$$W_{MMC_R} = \arg \max \text{Tr} (W^t R_b W - W^t R_w W)$$

Where, $W \in \mathbb{R}^{n \times n_r}$, $W^t W = \text{Inr}$
 $= [w_1^r, w_2^r, \dots, w_{n_r}^r]$

Such that,
 $(w_i^r)^t (R_b - R_w) w_i^r > \max (\text{Tr} (R_b - R_w) / n, 0)$
 $i = 1, \dots, n_r$

b) Compute the column projector W_{MMC_C} by solving the following optimization problem for R, G and B components:

$$W_{MMC_C} = \arg \max \text{Tr} (W^t C_b W - W^t C_w W)$$

Where, $W \in \mathbb{R}^{n \times n_c}$, $W^t W = \text{Inc}$
 $= [w_1^c, w_2^c, \dots, w_{n_c}^c]$

Such that,
 $(w_i^c)^t (C_b - C_w) w_i^c > \max (\text{Tr}(C_b - C_w) / m, 0)$,
 $i = 1, \dots, n_c$.

Step 5: Then, applying the projectors W_{MMC_R} of R, G, B components with the projectors W_{MMC_C} of R, G, B components respectively, the dimensionality reduction of face image can be performed,

$$W_{MMC_R} \Phi W_{MMC_C} : X \in \mathbb{R}^{m \times n} \rightarrow W_{MMC_C}^t X$$

$$W_{MMC_R} \in \mathbb{R}^{n \times n_r}$$

III. EXPERIMENTAL RESULTS

Experiments are designed using two standard databases: one from Georgia Technology (GTech) [29] and another Indian Database. Georgia Technology Database contains 50 classes and each class contains 15 color images taken at different lighting conditions. Indian Database contains 60 classes and each class contains 10 color images.

All images are cropped in such a way that it contains only face region and are resized to 92 x 112 resolution. For training purpose we selected randomly two images (for GTech database) and three images (for Indian database) from each class. Using above mentioned algorithm, features are extracted and stored in image size of 4 x 6. For testing purpose, we selected remaining images from respective class and applied same algorithm for feature extraction. Then using Euclidean distance, we matched test images with stored feature database. Code is designed with JAVA platform which is independent of language.

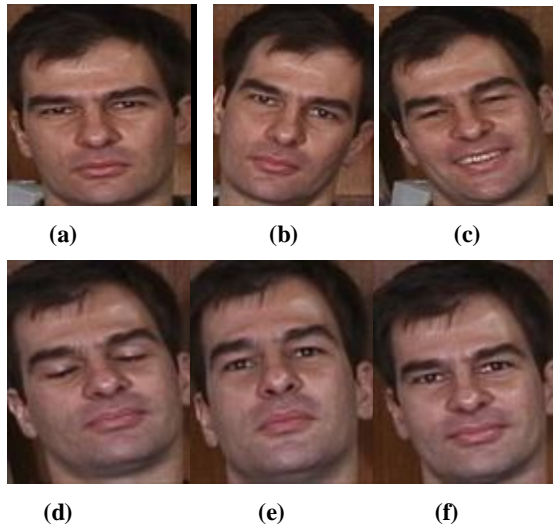
A. Results for GTech Database:

Training Database contains 50 classes and each class contains two images.

Case 1: Training Database for class 1 contains following images which are taken at different lighting conditions.



And test images are as follows,



Case 2: Training Database for class 1 contains following images.



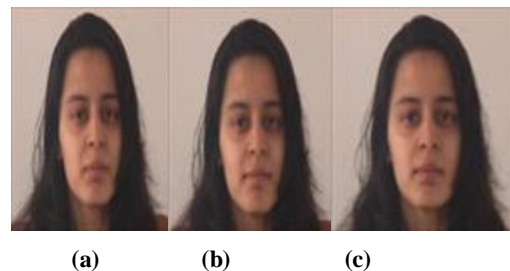
And test images are as follows,



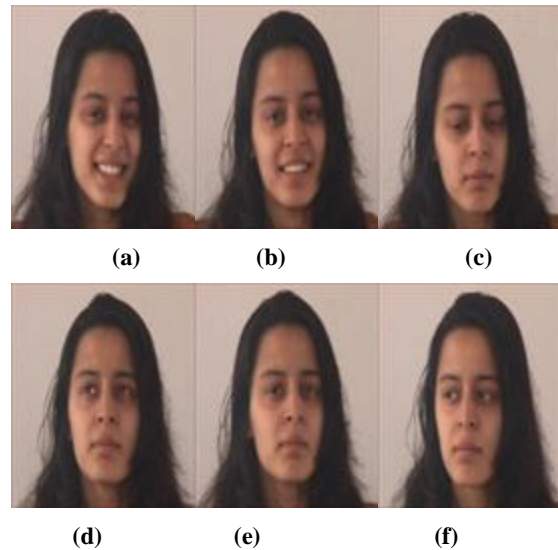
B) Results of Indian Database

Training Database contains 59 classes and each class contains three images.

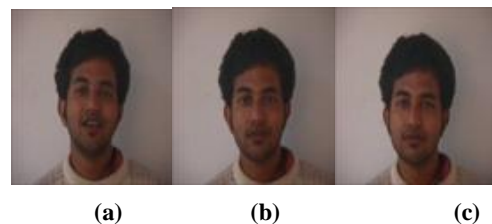
Case 1: Training Database for class 1 contains following images,



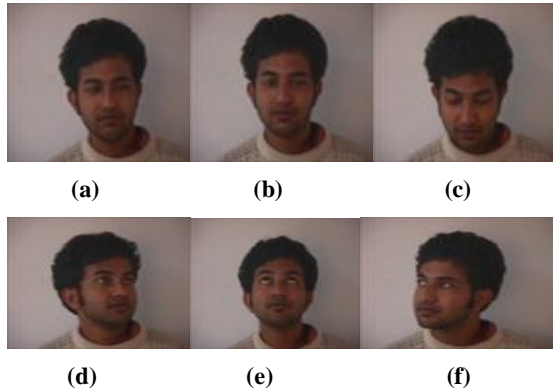
And test images are as follows,



Case 2: Training Database for class 58 contains following images.



And test images are as follows,



IV. PERFORMANCE ANALYSIS

A. Performance analysis for Georgia Technology Database:-

These results are compared with small number of training samples. Each training class contains two images. There are 50 known pattern classes. Therefore total number of training face image files is 100. Test image database contains remaining images from different classes. Average recognition rate when $n_r = 4$, $n_c = 10$ and $n_r = 4$, $n_c = 6$ is shown in table 1.

TABLE 1
 RECOGNITION RATE FOR DIFFERENT N_R AND N_C VALUES IN GTECH DATABASE IMAGES

	$n_r = 4, n_c = 10$	$n_r = 4, n_c = 6$
Average Recognition Rate (in %)	93 %	98%

Here average CPU time for training and testing on the Georgia Technology Database is computed. Training database contains 100 files and maximum test face image files selected are 42. Then average CPU time for training and testing is shown in table 2.

TABLE 2
 AVERAGE CPU TRAINING AND TESTING TIME FOR DIFFERENT N_R AND N_C VALUES IN GTECH DATABASE IMAGES

	Average training time(Sec)	Average testing time(Sec)
When $n_r = 4, n_c = 10$	17	2.6
When $n_r = 4, n_c = 6$	15	1.2

B. Performance analysis for Indian Database:-

These results are compared with small number of training samples. Each training class contains three images. There are 59 known pattern classes. Therefore total number of training face image files is 177. Test image database contains remaining images from different classes. Average recognition rate when $n_r = 4$, $n_c = 10$ and $n_r = 4$, $n_c = 6$ is shown in table 3.

TABLE 3
 RECOGNITION RATE FOR DIFFERENT N_R AND N_C VALUES IN INDIAN DATABASE IMAGES

	$n_r = 4, n_c = 10$	$n_r = 4, n_c = 6$
Average Recognition Rate (in %)	95 %	99%

Here average CPU time for training and testing on the Indian Database is computed. Training database contains 177 files and maximum test face image files selected are 68. Then average CPU time for training and testing is shown in table 4.

TABLE 4
 AVERAGE CPU TRAINING AND TESTING TIME FOR DIFFERENT N_R AND N_C VALUES IN INDIAN DATABASE IMAGES

	Average training time(Sec)	Average testing time(Sec)
When $n_r = 4, n_c = 10$	21	2.8
When $n_r = 4, n_c = 6$	18	1.3

V. CONCLUSION

The original size of the images is 92×112 and when features are extracted using Automatic Feature Selection Method, feature images are represented by image size 4×6 . So we can represent image in reduced size with extracted feature. Our experimental results are superior as compared to prior work implemented on MMC by others. This shows that the number of eigenvectors in row as well as column direction is reduced to 24 features (4×6) from 40 features (10×4). Automatic Feature Selection Method has less memory requirements and lower time complexity than MMC, which is desirable for large face data sets. Automatic Feature Selection Method is an extension of MMC. The key difference between Automatic Feature Selection Method and MMC is that Automatic Feature Selection Method works on the matrix representation of images directly, while MMC uses a vector representation. Finally, in this method, an entire automatic strategy is established for feature selection.

VI. REFERENCES

- [1] Y. H. Liu and Y. T. Chen, "Face recognition using total margin-based adaptive fuzzy support vector machines," *IEEE Trans. Neural Netw.*, vol. 18, no. 1, pp. 178–192, Jan. 2007.
- [2] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 1, pp. 4–20, Jan. 2004.
- [3] W. Zhao, R. Chellappa, and P. J. Phillips et al., "Face recognition: A literature survey," *ACM Comput. Surv.*, vol. 35, no. 4, pp. 399–459, Dec. 2003.
- [4] I. T. Jolliffe, *Principal Component Analysis*. New York: Springer-Verlag, 1986.
- [5] M. A. Turk and A. P. Pentland, "Eigenfaces for recognition," *J. Cogn. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.
- [6] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.
- [7] P. Howland, J. Wang, and H. Park, "Solving the small sample size problem in face recognition using generalized discriminant analysis," *Pattern Recognit.*, vol. 39, no. 2, pp. 277–287, Feb. 2006.
- [8] W. J. Krzanowski, P. Jonathan, W. V. McCarthy, and M. R. Thomas, "Discriminant analysis with singular covariance matrices: Methods and applications to spectroscopic data," *Appl. Stat.*, vol. 44, no. 1, pp. 101–115, 1995.
- [9] S. J. Raudys and A. K. Jain, "Small sample size effects in statistical pattern recognition: Recommendations for practitioners," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 3, pp. 252–264, Mar. 1991.
- [10] D. Q. Dai and P. C. Yuen, "Regularized discriminant analysis and its application to face recognition," *Pattern Recognit.*, vol. 36, no. 3, pp. 845–847, Mar. 2003.
- [11] D. Q. Dai and P. C. Yuen, "Face recognition by regularized discriminant analysis," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 37, no. 4, pp. 1080–1085, Aug. 2007.
- [12] X. S. Zhuang and D. Q. Dai, "Inverse Fisher discriminate criteria for small sample size problem and its application to face recognition," *Pattern Recognit.*, vol. 38, no. 11, pp. 2192–2194, Nov. 2005.
- [13] X. S. Zhuang and D. Q. Dai, "Improved discriminate analysis for high-dimensional data and its application to face recognition," *Pattern Recognit.*, vol. 40, no. 5, pp. 1570–1578, May 2007.
- [14] M. Kyperountas, A. Tefas, and I. Pitas, "Weighted piecewise LDA for solving the small sample size problem in face verification," *IEEE Trans. Neural Netw.*, vol. 18, no. 2, pp. 506–519, Mar. 2007.
- [15] S. Raudys and R. P. W. Duin, "Expected classification error of the Fisher linear classifier with pseudo-inverse covariance matrix," *Pattern Recognit. Lett.*, vol. 19, no. 5/6, pp. 385–392, Apr. 1998.
- [16] L. Chen, H. Liao, M. Ko, J. Lin, and G. Yu, "A new LDA-based face recognition system which can solve the small sample size problem," *Pattern Recognit.*, vol. 33, no. 10, pp. 1713–1726, Oct. 2000.
- [17] H. Yu and J. Yang, "A direct LDA algorithm for high-dimensional data - With application to face recognition," *Pattern Recognit.*, vol. 34, no. 10, pp. 2067–2070, Oct. 2001.
- [18] H. Li, T. Jiang, and K. Zhang, "Efficient and robust feature extraction by maximum margin criterion," *IEEE Trans. Neural Netw.*, vol. 17, no. 1, pp. 157–165, Jan. 2006.
- [19] M. Li and B. Yuan, "2D-LDA: A statistical linear discriminant analysis for image matrix," *Pattern Recognit. Lett.*, vol. 26, no. 5, pp. 527–532, Apr. 2005.
- [20] J. Yang and J. Y. Yang, "From image vector to matrix: A straightforward image projection technique—IMPCA vs. PCA," *Pattern Recognit.*, vol. 35, no. 9, pp. 1997–1999, Sep. 2002.
- [21] 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 Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 131–137, Jan. 2004.
- [22] J. Yang, D. Zhang, X. Yong, and J. Y. Yang, "Two-dimensional discriminant transform for face recognition," *Pattern Recognit.*, vol. 38, no. 7, pp. 1125–1129, Jul. 2005.
- [23] D. Zhang and Z. H. Zhou, "(2D)2PCA: Two-directional two-dimensional PCA for efficient face representation and recognition," *Neuro-computing*, vol. 69, no. 1–3, pp. 224–231, Dec. 2005.
- [24] S. Noushath, G. Hemantha Kumar, and P. Shivakumara, "(2D)2LDA: An efficient approach for face recognition," *Pattern Recognit.*, vol. 39, no. 7, pp. 1396–1400, Jul. 2006.
- [25] J. Ye, "Generalized low rank approximations of matrices," *Mach. Learn.*, vol. 61, no. 1–3, pp. 167–191, Nov. 2005.
- [26] J. Ye, R. Janardan, and Q. Li, "Two-dimensional linear discriminant analysis," in *Advances in Neural Information Processing Systems*, vol. 17. Cambridge, MA: MIT Press, 2005, pp. 1569–1576.
- [27] X. Y. Jing, H. S. Wong, and D. Zhang, "Face recognition based on 2D Fisherface approach," *Pattern Recognit.*, vol. 39, no. 4, pp. 707–710, Apr. 2006.
- [28] W. M. Zuo, D. Zhang, J. Yang, and K. Wang, "BDPCA plus LDA: A novel fast feature extraction technique for face recognition," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 36, no. 4, pp. 946–953, Aug. 2006.