Journal of Fiber Bioengineering and Informatics 8:2 (2015) 365–372
doi:10.3993/jfbim00121

Modular Two-dimensional Principal Component Regression for Robust Face Recognition ⋆

Zhenyue Zhang, Mingyan Jiang*, Xianye Ben, Fei Li

School of Information Science and Engineering, Shandong University, Jinan 250100, China

Abstract

To improve robustness of Linear Regression (LR) for face recognition, a novel face recognition framework based on modular two-dimensional Principal Component Regression (2DPCR) is proposed in this paper. Firstly, all face images are partitioned into several blocks and the approach performs 2DPCA process to project the blocks onto the face spaces. Then, LR is used to obtain the residuals of every block by representing a test image as a linear combination of class-specific galleries. Lastly, three minimum residuals of every block and fuzzy similarity preferred ratio decision method are applied to make a classification. The proposed framework outperforms the state-of-the-art methods and demonstrates strong robustness under various illumination, pose and occlusion conditions on several face databases.

Keywords: Face Recognition; Block 2DPCR; Liner Regression; Fuzzy Similarity Preferred Ratio Decision

1 Introduction

Face Recognition (FR) has been a concerned issue because of its wide application in information science field in recent years. One of the most challenging problems about FR is how to suppress the influence of illumination, gesture variation and occlusion.

Pattern recognition methods based on subspace have developed rapidly in the past few years. Subspace methods also have been used to achieve successful face recognition which is based on the theory that face images from a specific class lie on a subspace. Images are transformed linearly or non-linearly to low-dimensional vectors with subspace methods in the face space. These approaches could be categorized into two categories, namely, reconstructive and discriminative methods. Classical reconstructive subspace methods are robust to noise including Principal Component Analysis (PCA), Independent Component Analysis (ICA) and Kernel PCA (KPCA). Discriminative methods, such as Linear Discriminant Analysis (LDA) and kernel LDA (KDA) are efficient

⋆Project supported by the Natural Science Foundation of Shandong Province (No. ZR2014FM039), the National Natural Science Foundation of China (No. 61201370) and the Doctoral Scientific Fund Project of the Ministry of Education of China (No. 20120131120030).

*Corresponding author.
Email address: jiangmingyan@sdu.edu.cn (Mingyan Jiang).

1940–8676 / Copyright © 2015 Binary Information Press & Textile Bioengineering and Informatics Society
June 2015

in pure conditions. Although the following discriminative methods, such as Generalized Discriminant Analysis (GDA) [1], the Kernel Direct LDA (KDDA) [2] and Kernel Subspace LDA (KSLDA) [3], can achieve a good performance under illumination and gesture variation, the performance is still not satisfactory for severe variations.

Recently, a few reconstructive approaches were proposed to tackle illumination variations. In 2010, Linear Regression Classification (LRC) [4] has been introduced by assuming that face images from a specific class lie on a linear subspace. Test samples are reconstructed in square sense by training samples of every class, so the regression coefficients can be estimated with the least square method, and then the decision is made in favor of the class with the minimum reconstruction error. In 2012, Huang incorporated LRC with PCA and proposed Principal Component Regression (PCR) [5] by utilizing PCA's ability of data expression. Moreover, Lu proposed Kernel Linear Regression Classification (KLRC) [6] based on LRC and kernel trick in 2014. Experimental results show that the three methods are robust to illumination variations. However, it still could not withstand severe illumination variations.

The good robustness under various illumination of LRC is derived from its reconstructive ability. PCR maintains LRC's advantages and utilizes PCA's strong ability of data expression. Two-dimensional Principal Component (2DPCA) has a stronger expression ability than PCA, so two-dimensional Principal Component Regression (2DPCR) is proposed in this paper. Although strong reconstructive capacity can lower the influence of illumination variations, it will reduce the robustness to pose and occlusion. Modular presentation of face images is helpful to eliminate the error caused by local pollution. Therefore, modular LRC is a valid method to cope with pose and occlusion. Usually, minimum distance or vote is adopted to make decisions in modular approach, but both of them lack effectiveness since few informations are exploited. Final decision can be treated as a fuzzy decision problem due to the existence of residual errors of every class after modular linear regression. Li et al. [7] only use fuzzy membership degree to make decisions and does not make full use of the advantages of fuzzy decision. In this paper, fuzzy similarity preferred ratio decision method [8] is adopted to make decisions.

Based on the above analysis, we proposed a novel face recognition framework based on modular 2DPCR with fuzzy decision.

2 Modular 2DPCR with Fuzzy Decision

2.1 Linear Regression Classification Algorithm

Assume there are N number of classes with p_i number of training samples from the class, $i = 1, 2, \cdots, N$. All training samples constitute a matrix W and every column is a training sample: $W = [W_1, W_2, ..., W_N] = \left[w_1^1, w_2^1, ..., w_{p_1}^1, ..., w_1^N, w_2^N, ..., w_{p_N}^N\right]$, where W_i represents the ith class, and w_j^i is the jth training sample belonging to the ith class, $j = 1, 2, \cdots, p_i$. LRC reckons that face images from a specific class lie on a linear subspace. So, if y belongs to the ith class, we can use W_i to linearly represent y. As shown in Eq. (1):

$$y = W_i\beta_i + e_i. \tag{1}$$

Vector e_i represents the error of the ith class which is perpendicular to linear subspace spanned by W_i. There are two factors causing errors: one is noise of y, and the other one is that the

subspace is not completely spanned by W_i.

The aim of LR is to find a β_i minimizing the l_2 norm of e_i. The objective function is shown below:

$$\beta_i = \arg\min \|y - W_i\beta_i\|_2, i = 1, 2, \cdots, N. \tag{2}$$

β_i can be estimated using least-square estimation:

$$\beta_i = (W_i^{\mathrm{T}} W_i)^{-1} W_i^{\mathrm{T}} y. \tag{3}$$

The projection of y on the subspace spanned by W_i can be obtained:

$$\tilde{y}_i = W_i\beta_i = W_i(W_i^{\mathrm{T}} W_i)^{-1} W_i^{\mathrm{T}} y = H_i y. \tag{4}$$

Residual r_i corresponding to subspace spanned by W_i is as follows:

$$r_i = \|e_i\|_2 = \|y - \tilde{y}_i\|, i = 1, 2, \cdots, N. \tag{5}$$

Then, y can be classified according to Eq. (6):

$$k = \arg\min r_i, i = 1, 2, \cdots, N. \tag{6}$$

2.2 Modular 2DPCR

Let x be a projection vector, $x \in R^{n\times 1}$ and image A be a random matrix, $A \in R^{n\times n}$. The basic idea of 2DPCA is to project A onto x as shown in Eq. (7),

$$y = Ax, \tag{7}$$

where y is called the projection feature vector. We can use the total scatter of the projected samples to measure discriminatory power of the projection vector x, so the following criterion function is defined:

$$J(x) = \mathrm{tr}(S_x) = x[\mathrm{E}(A - \mathrm{E}A)^{\mathrm{T}}(A - \mathrm{E}A)]x^{\mathrm{T}} = xGx^{\mathrm{T}}, \tag{8}$$

where S_x represents the covariance matrix of the projected feature vectors of the training samples, $\mathrm{tr}(S_x)$ represents the trace of S_x and G represents the covariance matrix. To enhance discriminant ability, we usually need to select a set of orthonormal projection axes $X = [x_1, x_2, \cdots, x_L]$, where L is the number of maximum eigenvalues and satisfies $\sum_{l=0}^{L} \lambda_i / \sum_{l=0}^{n} \lambda_i \geq \alpha$,where λ_l represents the ith maximum eigenvalue and α is a predefined threshold which is usually set as 0.95.

The feature matrices $W_j^{'i}(i = 1, 2, \cdots, N, j = 1, 2, \cdots, p_i)$ can be obtained by projecting the training sample images onto X. Each matrix $W_j^{'i}$ is transformed into a vector w_j^i. then, all vectors w_j^i constitute the matrix of the training samples W. Given a random test image Y, it is projected onto X, and its corresponding feature matrix $Y^{'}$ and vector y can be acquired. Then, Eq. (2)-(6) are used to determine the class of Y.

The modular approach can effectively remove the influences caused by local pollution (various illumination, pose and occlusion conditions) and improve method's robustness. Modular 2DPCR has two steps. Firstly, all images are partitioned into several blocks and 2DPCR is executed on each subimage. Secondly, the fuzzy similarity preferred ratio decision method is used to make a classification.

2.3 Fuzzy Similarity Preferred Ratio Decision

Fuzzy similarity preferred ratio decision method is also a kind of dual contrast decision method. Firstly, fuzzy similarity priority ratio r_{ij} is defined by dual relative comparison, so fuzzy similarity priority ratio matrix is built. Then, all scheduled schemes are sorted by established a λ cut-matrix.

Definition 1 *Let $U = \{x_1, x_2, \cdots, x_n\}$ be a fuzzy domain. For a given pair of elements, if there exists a pair $(f_j(x_i), f_i(x_j))$ which meets $0 \leq f_j(x_i) \leq 1, 0 \leq f_i(x_j) \leq 1$. Comparing x_i with x_j, if x_i has a certain property whose degree is represented by $f_j(x_i)$, then the degree of x_j is $f_i(x_j)$. The pair $(f_j(x_i), f_i(x_j))$ is called dual relative comparison corresponding to x_j and x_i, or comparison for short.*

Fuzzy similarity preferred ratio decision method is introduced in the following.

(1) Determine the fuzzy domain $U = \{x_1, x_2, \cdots, x_n\}$.

(2) Figure out fuzzy similarity priority ratio r_{ij} and build fuzzy similarity priority ratio matrix R. If $(f_j(x_i), f_i(x_j))$ is a dual comparison, then

$$r_{ij} = \frac{f_j(x_i)}{f_j(x_i) + f_i(x_j)}, \quad r_{ji} = \frac{f_i(x_j)}{f_j(x_i) + f_i(x_j)}, \tag{9}$$

where r_{ij} is called fuzzy similarity priority ratio and $R = (r_{ij})_{n\times n}$ represents fuzzy similarity priority ratio matrix. $r_{ij} = 0.5$ illustrates that the priority degree of x_i is equal to its own's.

(3) All scheduled schemes are sorted by establishing a λ cut-matrix (we can also calculate the infimum of each row in R without diagonal elements. The row containing the maximum infimum can be regarded as first priority and repeat above steps to obtain a sequence of all scheduled schemes).

2.4 Steps of Fuzzy Decision

(1) Divide the training images into several subimages and extract the feature matrices of each subimage using 2DPCA.

(2) Transform the feature matrices into the feature vectors and lower feature's dimension with PCA. Therefore, the matrices of the training samples $W^k(k = 1, 2, 3, 4)$ of each block is gained: $W^k = \left[W_1^k, W_2^k, ..., W_N^k\right] = [w_{1,1}^k, w_{1,2}^k, ..., w_{1,p_1}^k, ..., w_{N,1}^k, w_{N,2}^k, ..., w_{N,p_N}^k]$, where $W_{i,j}^k$ represents the jth sample vector of the kth block, ith class.

(3) For a random test image Y, the feature matrices $Y_k^{'}(k = 1, 2, 3, 4)$ and the feature vectors $y_k(k = 1, 2, 3, 4)$ are extracted after is partitioned into several blocks, where k presents the kth subimage.

(4) Work out residuals $r_i^k(i = 1, 2, 3, 4)$ corresponding to the ith class, kth subimage using Eq. (3), (4) and (5).

(5) Classify the test image with fuzzy similarity preferred ratio decision.

a. Pick up three minimum residuals of every block and the different classes are chosen to form domain U, where $U = \{x_1, x_2, ..., x_D\}, D \leq 12$.

b. Define the pair $(f_j(x_i), f_i(x_j))$

$$f_j(x_i) = \sum_{k=1}^{M} \frac{r_j^k}{r_i^k + r_j^k}, \quad f_i(x_j) = \sum_{k=1}^{M} \frac{r_i^k}{r_i^k + r_j^k}. \tag{10}$$

In this paper $M = 4$, which represents the modular quantity. The effect of the definition is twofold. First, it normalizes the residuals of each block and eliminates the influence of local noise further. Second, it is easier to win for a smaller residual in the process of making decisions.

c. Build fuzzy similarity priority ratio matrix $R = (r_{ij})_{nn}$.

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1D} \\ r_{21} & r_{22} & \cdots & r_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ r_{D1} & r_{D2} & \cdots & r_{DD} \end{bmatrix}, \tag{11}$$

where can be calculated by Eq. (9) and (10).

d. Find the infimum of each row in except diagonal elements, then the class corresponding to biggest infimum is final decision.

3 Experimental Results

We will demonstrate the effectiveness of the proposed approach on four publicly available face databases: Yale B [9], CMU-PIE [10], FERET [11] and AR [12]. The threshold involving PCA and 2DPCA is 0.99 on YaleB, CMU-PIE and 0.95 on FERET, AR. In SRC [13], we use eigenfaces as the input facial features, and use all the training samples as the dictionary.

3.1 Yale B Database

Yale B database consists of 640 frontal face images of 10 subjects. All images are cropped and resized to pixels $56 * 64$. We follow the experimental protocol as proposed in [5] and compare the proposed approach with 2DPCA, KPCA, LDA, PCR, 2DPCR, etc. According to various light conditions, all images are divided into 5 subsets: subset 1 consisting of 70 face images under normal light is used as training samples; the others are respectively used for test. Subsets 2-5 characterize different light variations from sight, moderate, severe to very severe. It is shown in Fig. 1.

Results are shown in Table 1. Modular 2DPCR1 depends on the vote to make decision, while modular 2DPCR2 is based on fuzzy decision. All images are partitioned into four blocks, as shown in Fig. 2. Sparse Representation-based Classification (SRC) is added for comparison because of its robustness. The results reveal that the proposed method has strong robustness to severe luminance alterations and is far better than SRC, PCR especially in the condition of very severe light variations. It outperforms PCR by margins of 44 percent for subset 5.

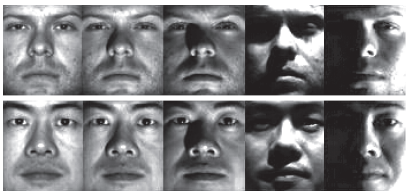


Fig. 1: Samples from five subsets of the Yale B database

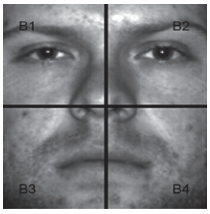


Fig. 2: $2 * 2$ block of a face image

Table 1: Accuracy (%) comparisons on Yale B

Methods	Subset 2	Subset 3	Subset 4	Subset 5	Overall
PCA	95.00	56.67	22.21	11.05	41.05
2DPCA	94.16	99.16	55.00	15.80	59.47
LDA	**100.00**	92.50	33.57	15.79	54.04
KPCA	95.83	64.17	24.29	14.21	47.07
SRC	**100.00**	90.83	44.29	23.68	58.94
PCR	**100.00**	**100.00**	95.00	58.42	84.91
2DPCR	**100.00**	**100.00**	98.57	73.16	90.70
Modular 2DPCR1	99.17	**100.00**	94.29	90.53	95.26
Modular 2DPCR2	**100.00**	**100.00**	**100.00**	**98.42**	**99.47**

3.2 CMU-PIE Database

CMU-PIE database consists of 480 face images of 20 subjects as shown in Fig. 3. All images are cropped and resized to $32 * 32$ pixels. Eighty face images (8, 11, 13 and 22 in per subject) under normal light are picked up as training samples. Results are reported in Table 2. Modular 2DPCR2 shows quite agreeable results with CMU-PIE database as well. It further demonstrates the robustness of modular 2DPCR2. SRC algorithm is not taken as a comparative approach for few training samples in per subject.



Fig. 3: A typical subject from the CMU-PIE database

Table 2: Accuracy (%) comparisons on CMU-PIE, FERET and AR

Methods	CMU-PIE	FERET	AR	Sunglass(AR)	Scarf(AR
PCA	40.00	68.00	82.25	_	_
2DPCA	82.00	71.00	89.25	_	_
LDA	80.00	81.00	**98.50**	_	_
KPCA	45.75	69.00	88.15	_	_
SRC	_	_	98.00	67.50	71.25
PCR	89.50	81.00	97.00	62.50	7.50
2DPCR	91.10	81.00	94.95	75.00	88.75
Modular 2DPCR1	92.00	75.50	93.10	73.75	91.25
Modular 2DPCR2	**93.00**	**82.50**	98.00	**90.00**	**92.50**

3.3 FERET Database

There are 1400 face images of 200 subjects under various illumination and expression conditions in the FERET database as shown in Fig. 4. All images are cropped to $30 * 30$ pixels. Three face images with moderate illumination and expression variations are used as the gallery, while all others are used for validation. Table 2 shows a detailed comparison of modular 2DPCA with a variety of approaches. Modular 2DPCR based on fuzzy similarity preferred ratio decision outperforms modular 2DPCR using voting decision and LDA by a margin of 7 and 1.5 percent respectively. The fuzzy similarity preferred ratio decision method has gained superiority over voting decision.



Fig. 4: A typical subject from the FERET database

3.4 AR Database

3.4.1 Illumination and Expression Variations

In AR database, 40 subjects per 14 face images are picked up for experiments. We crop all the face images to $40 * 50$ pixels. The first 7 as training samples, after the 7 categories as a test sample, correct classification rate statistics. The result displayed by Table 2 reveals that the proposed method achieves 98 percent as well as LDA and SRC approaches.

3.4.2 Contiguous Occlusion

Contiguous occlusion is usually considered as an obstacle in face recognition. In these experiments, 40 subjects per 9 face images including occlusive and normal conditions are used to verify effectiveness of our approach. The contiguous occlusion is caused by the sunglass and scarf in AR database. The five normal images are used as training samples and the four occlusive images are for testing. Results are shown in Table 2. It is obvious that the proposed method obtains the best performance compared with SRC, 2DPCA and PCR. The modular 2DPCR based on fuzzy similarity preferred ratio decision achieves more than 90 percent, both in the sunglass and the scarf occlusion.

3.5 Interpretation of Results

The experiments carried out on Yale B and CMU-PIE database reveal that Modular 2DPCR based on fuzzy similarity preferred ratio decision is highly robust against illumination variations. Meanwhile, the results on FERET and AR also show that the proposed method is not influenced by gesture and contiguous occlusion. Strong robustness of the proposed method can be attributed to two points: one is that modular presentation of face images and fuzzy similarity preferred ratio decision are helpful to eliminate error caused by local pollution, the other is that powerful reconstructive capacity of 2DPCR is benefit to find the best possible subspace.



Fig. 5: A typical from the AR database

4 Conclusions

In this paper, a novel face recognition framework based on modular 2DPCR with fuzzy decision is proposed. 2DPCR is robust to illumination variations because of combining the reconstructive ability of LRC with 2DPCA's ability of data expression. Meanwhile, the modular presentation and fuzzy decision are also taken to tackle local pollution (various illumination, pose and occlusion conditions). The proposed framework outperforms the state-of-the-art methods and demonstrates good robustness under various illumination, pose and occlusion conditions on Yale B, FERET, CMU-PIE and AR face databases.

References

[1] G. Baudat, F. Anouar, Generalized discriminant analysis using a kernel approach [J], Neural Computation, 12(10), 2000, 2385-2404

[2] J. W. Lu, K. Plataniotis, A. N. Venetsanopoulos, Face recognition using kenel direct discrimininent analysis algorithms [J], IEEE Transactions Neural Network, 14(1), 2003, 117-126

[3] J. Huang, P. C. Yuen, Y. Chen, Choosing parameters of kernel subspace LDA for recognition of face images under pose and illumination variations [J], IEEE Transactions on Systems, Man, and Cybernetics, 37(4), 2007, 847-862

[4] I. Naseem, R. Togneri, M. Bennamoun, Linear regression for face recognition [J], IEEE Transactions on Pattern Analysis and Intelligence, 32(11), 2010, 2106-2112

[5] S. M. Huang, J. F. Yang, Improved principal component regression for face recognition under illumination Variations [J], IEEE Signal Processing Letters, 19(4), 2012, 179-182

[6] Y. W. Lu, X. Z. Fang, B. L. Xie, Kernel linear regression for face recognition [J], Neural Comput. & Applic., 24, 2014, 1843-1849

[7] X. Li, K. J. Wang, X. Y. Ben, MW(2D)-2PCA based face recognition with single training sample [J], Pattern Recognition and Artificial Intelligence, (1), 2001, 77-83

[8] J. J. Xie, C. P. Liu, Fuzzy Mathematics and Its Application [M], Fourth Edition, Wuhan, China, Huazhong University of Science and Technology Press, 2013, 120-141

[9] A. S. Georghiades, P. N. Belhumeur, D. W. Jacobs, From few to many: Illumination cone model for face recognition under variable lighting and pose [J], IEEE Transactions on Pattern Analysis and Intelligence, 23(6), 2001, 630-660

[10] T. Sim, S. Baker, M. Bsat, The CMU pose, illumination, and expression database [J], IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(12), 2003, 1615-1618

[11] P. J. Phillips, H. Moon, A. Rizvi et al., The FERET valuation methodology for face recognition algorithms [J], IEEE Transactions on Pattern Analysis and Intelligence, 22(10), 2000, 1090-1104

[12] A. Martinez, R. Benavente, The AR Face Database, CVC Technical Report 24, June 1998

[13] J. Wright, A. Y. Yang, A. Ganesh et al., Robust face recognition via sparse representation [J], IEEE Transactions on Pattern Analysis and Intelligenc, 31(2), 2009, 210-227