

Face Recognition based on Modular 2DPCA and Contextual Constraints based Kernel Discriminant Analysis

Hua-Li Feng

Educational Informatization Centre

Wuxi Institute of Commerce

Wuxi, China

fhl0515@163.com

Abstract—In this paper, an improved face recognition algorithm is proposed based on the combination of modular 2DPCA and contextual constraints based kernel discriminant analysis (CCKDA) because of the disadvantages of CCLDA. CCLDA first transforms an image matrix to a vector which caused high dimensionality and computational complexity and not considers the local feature. While our method first extracts the local features with the original images which are divided into modular sub-images, then CSKDA is utilized, which incorporates the contextual information into kernel discriminant analysis. Experimental results obtained on ORL and XM2VTS databases show the effectiveness of the new method.

Keywords- M2DPCA; CCLDA; CCKDA; face recognition

I. INTRODUCTION

Over the past decades, human face recognition has been an extremely attractive research subjects. It is motivated by the wide range of commercial and law enforcement applications. A large number of algorithms have been developed.

Due to its efficacy and efficiency, subspace methods, like PCA [1] and LDA [2] have achieved great success in pattern recognition and computer vision area, such as face recognition and image retrieval, etc. PCA uses the Karhunen-Loeve transform to produce the most expressive subspace for face representation and recognition by minimizing the residual of the reconstruction. However, it does not utilize any class information and so it may drop some important clues for classification. To avoid the disadvantage, the LDA is proposed and the objective of LDA is to find a subspace that maximizes the sample distance from different classes and meanwhile minimizes the samples distance from the same class. Thus the derived subspace is discriminative to classify different samples correctly. However, in real applications, due to the high dimension of feature and usually small number of samples, the classical LDA always fails; this is called small sample size (SSS) problem. To address this problem, a lot of work has been investigated and many LDA-variants have been proposed.

All the above methods based face recognition technique, the 2D face image matrices must be previously transformed into 1D image vectors. The resulting image vectors of faces usually lead to a high dimensional image vector space, where it is difficult to evaluate the covariance matrix accurately due to its large size and the relatively small number of training

samples. Fortunately, the two dimensional methods for image feature extraction, called 2DPCA [3] and 2DLDA [4], is developed. 2DPCA and 2DLDA are based on 2D matrices rather than 1D vector. As a result, 2DPCA and 2DLDA has two important advantages over PCA and LDA. First, it is easier to evaluate the covariance matrix accurately. Second, less time is required to computational and the LDA's singularity problem can be avoided. However it has been seen that the facial changes due to variations of pose, illumination, expression, etc. are appeared only some regions of the whole image. Therefore, the conventional face recognition methods, which use whole image for feature extraction and recognition, do not result much success. To cope with the above facial changes, M2DPCA is proposed.

However, all the above methods consider the pixels in image independently, not taking into account their spatial relationship. It is well known that images with certain pattern occupy specific manifold, which is constrained by contextual information, in high-dimension feature space. Therefore, contextual constraint in image is important for image understanding and will provide useful information for classification. One of the most successful work to model the contextual information is the Markov Random Fields (MRFs) [5-6] which derive the results by maximizing the posterior probability in Bayesian deduction framework. However, the optimization by MRF is somewhat computational expensive and is easy to converge into local minima that limits its application. Wang [7] proposed a novel image matching distance considering the spatial information. However, they did not demonstrate how to integrate the contextual information into dimensionality reduction problem. Lei propose contextual constraints based linear discriminant analysis (CCLDA) [8] which incorporates the contextual information into linear discriminant analysis.

However, CCLDA is a linear method and does not take account of the high dimensional information of images. Meanwhile it first need transform an image matrix to a vector which caused high dimensionality and computational complexity and not considers the local feature. In this paper, we proposed an improved face recognition algorithm based on the combination of modular 2DPCA and contextual constraints based kernel discriminant analysis (CCKDA).

The rest of this paper is organized as follows. Section 2 details the feature extraction based on M2DPCA and CCKDA. Experimental results on ORL and XM2VTS

databases are proposed in Section 3. Finally, conclusions are drawn in Section 4.

II. FEATURE EXTRACTION BASED ON M2DPCA AND CCKDA

To reduce the computational time and preserve the local information of images, we proposed an improved face recognition method based M2DPCA and CCKDA. First we adopt M2DPCA to reduce the dimensions of samples and extract the local features with the original images which are divided into modular sub-images. Then the CCKDA is utilized which can provide more effective information for classification.

A. The Mian Idear of Modular 2DPCA

Let $A = \{A_1, A_2, \dots, A_N\}$ denote N face images belonging to C classes: $\omega_1, \omega_2, \dots, \omega_C$ and the size of each image A_i is $m \times n$, and each class has N_i samples ($N = \sum_{i=1}^C N_i$). We can partition each face image A_i into $S = p \times q$ equally sized sub-images in a non-overlapping way and the size of each sub-image is $\frac{m}{p} \times \frac{n}{q}$:

$$A_i = \begin{bmatrix} (A_i)_{11} & (A_i)_{12} & \dots & (A_i)_{1q} \\ (A_i)_{21} & (A_i)_{22} & \dots & (A_i)_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ (A_i)_{p1} & (A_i)_{p2} & \dots & (A_i)_{pq} \end{bmatrix} \quad (1)$$

Then the covariance matrix G_t of the sub-image can be defined as:

$$G_t = \frac{1}{M} \sum_{i=1}^N \sum_{k=1}^p \sum_{l=1}^q ((A_i)_{kl} - \mu)((A_i)_{kl} - \mu)^T \quad (2)$$

where $M = Npq$ denotes the total number of the training sub-images and the mean image matrix of the all the training sub-image is denoted by $\mu = \frac{1}{M} \sum_{i=1}^N \sum_{k=1}^p \sum_{l=1}^q (A_i)_{kl}$.

Like 2DPCA, the optimal projection matrix $W = [w_1, w_2, \dots, w_d]$ is used for feature extraction. For a given image sample A_i , an image feature is obtained by the following transformation:

$$Y_i = \begin{bmatrix} W^T(A_i)_{11} & W^T(A_i)_{12} & \dots & W^T(A_i)_{1q} \\ W^T(A_i)_{21} & W^T(A_i)_{22} & \dots & W^T(A_i)_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ W^T(A_i)_{p1} & W^T(A_i)_{p2} & \dots & W^T(A_i)_{pq} \end{bmatrix} \quad (3)$$

To carry out the following steps, we need to transform the matrix Y_i into a column vector y_i , it is easy to know that the size of y_i is $R^{d \times m \times n}$.

B. Feature Extraction of CCKDA

Then we conduct the CCKDA method with the transformed sample vector y_i . Moreover, to avoid the complexity of KDA, our work adopts the framework of KPCA+LDA [9].

For a given nonlinear mapping Φ , the input data space R^M can be mapped into the feature space F :

$$\Phi : R^M \rightarrow F \quad (4)$$

$$y \in R^M \rightarrow \Phi(y) \in F$$

Then the training set mapped in the feature space is denoted by $Q = [\Phi(y_1), \Phi(y_2), \dots, \Phi(y_N)]$, so the covariance operator on the feature space F can be constructed by

$$S_t^\Phi = \frac{1}{N} \sum_{i=1}^N (\Phi(y_i) - m^\Phi)(\Phi(y_i) - m^\Phi)^T \quad (5)$$

where $m^\Phi = \frac{1}{N} \sum_{i=1}^N \Phi(y_i)$. Let β be the eigenvector of S_t^Φ , and assume that β can be linearly expressed by

$$\beta = \sum_{i=1}^N a_i \Phi(y_i) \quad (6)$$

To obtain the expansion coefficients, let us denote an $N \times N$ Gram matrix $R = Q^T Q$, whose elements can be determined by virtue of kernel tricks:

$$\tilde{R}_{ij} = \Phi(y_i)^T \Phi(y_j) = (\Phi(y_i) \cdot \Phi(y_j)) = k(y_i, y_j) \quad (7)$$

Centralize \tilde{R} by

$$R = \tilde{R} - I_N \tilde{R} - \tilde{R} I_N + I_N \tilde{R} I_N \quad (8)$$

where the matrix $I_N = (1/N)_{N \times N}$.

Calculate the orthonormal eigenvectors $\gamma_1, \gamma_2, \dots, \gamma_K$ of R corresponding to the K largest positive eigenvalues, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_K$, then

$$\beta_i = \frac{1}{\sqrt{\lambda_i}} Q \gamma_i, i = 1, 2, \dots, K \quad (9)$$

Thus, we can obtain the KPCA-transformed feature vector $Z = (z_1, z_2, \dots, z_K)^T$ by

$$Z = P^T \Phi(y) \quad (10)$$

where $P = [\beta_1, \beta_2, \dots, \beta_K]$.

Then CCLDA is utilized on the training samples of KPCA. The principle of LDA is to find the optimal projection and maximize the ratio of between-class to that of within-class scatter. While it merely considers the pixels in image independently, not taking account their spatial relationship. Intuitively, if the pixels are of the similar property or reflect the similar structure, the weights on them would have strong relationship, otherwise the weights on

independent pixels would also be weakly related. Following this rational, we impose a constraint

$J_2(w) = \frac{1}{2} \sum_{i,j} (w_i - w_j)^2 S_{ij}$ on traditional LDA to reformulate the object of discriminant analysis as

$$J = \frac{w^T S_b w}{(w^T S_w w + \eta J_2(w))} \quad (11)$$

where S_{ij} describes the similarity of pixels i and j , and the parameter η is a coefficient to balance the trade-off between the training discriminant power and contextual constraints. For a nature image, one assumption often used is that the pixels in local neighboring region have the consistent properties and reflect the similar image structure. Therefore, one way to define the contextual matrix S among different weights of pixels is

$$S_{ij} = \begin{cases} e^{\|f_i - f_j\|^2 / \sigma^2}, & \text{if } i \text{ and } j \text{ are neighbors} \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

where f_i and f_j are the feature vectors extracted at position i and j , respectively to describe the texture and spatial relationship between positions i and j . In the experiment, the feature vector f in (12) is extracted by grouping the pixel values at corresponding position from all the training images. The parameter σ is empirically set to be the average distance among these feature vectors.

The constraint function $J_2(w)$ gives a high penalty when the weights of related pixels differ too much. Due to the symmetry of S_{ij} in general case, the contextual constraints $J_2(w)$ on weight image can be formulated using the matrix operations further as follows.

$$J_2(w) = \frac{1}{2} \sum_{i,j} (w_i - w_j)^2 S_{ij} \quad (13)$$

$$= w^T D w - w^T S w = w^T L^w w$$

where $L^w = D - S$ is the Laplacian matrix, and D is a diagonal matrix where $D_{ii} = \sum_j S_{ij}$. Thus, the objective of CCLDA can be reformulated as

$$J = \frac{w^T S_b w}{(w^T S_w w + \eta w^T L^w w)} \quad (14)$$

The optimal projection w can be obtained by solving the following generalized eigen-value problem.

$$S_b w = \lambda (S_w + \eta L^w) w \quad (15)$$

III. EXPERIMENTAL RESULTS AND ANALYSIS

The performance of the proposed method has been evaluated on the ORL and XM2VTS face database. And in the experiments the nearest neighbor classifier is used for the classification.

TABLE I. THE PERFORMANCE ON ORL DATABASE

Method	G2	G3	G4	G5	G6	G7	G8
--------	----	----	----	----	----	----	----

CCLDA	77.5	79.29	89.17	92	93.75	95.83	97.5
M2DPC A+CCL DA (2 ×4)	86.56	93.21	96.67	100	100	100	100
M2DPC A+CCL DA (4 ×4)	81.87	90.71	98.33	100	100	100	100
M2DPC A+CCL DA (4 ×8)	82.5	91.43	98.33	100	100	100	100
M2DPC A+CCK DA (2 ×4)	87.49	95.38	97.13	100	100	100	100
M2DPC A+CCK DA (4 ×4)	83.61	92.57	98.47	100	100	100	100
M2DPC A+CCK DA (4 ×8)	85.18	93.75	98.50	100	100	100	100

TABLE II. THE PERFORMANCE ON XM2VTS DATABASE

Method	G2	G3	G4	G5	G6
CCLDA	67.29	91.32	89.92	94.8	95.08
M2DPCA+CCLDA (2×4)	76.33	97.69	100	100	100
M2DPCA+CCLDA (4×4)	76.5	97.83	100	100	100
M2DPCA+CCLDA (4×8)	77.06	97.63	100	100	100
M2DPCA+CCKDA (2×4)	80.12	98.24	100	100	100
M2DPCA+CCKDA (4×4)	80.28	98.46	100	100	100
M2DPCA+CCKDA (4×8)	80.69	98.28	100	100	100

TABLE III. THE RUNNING TIME ON XM2VTS DATABASE

Method	G2		G4		G6	
	Time of Feature Extraction (S)	Running Time (S)	Time of Feature Extraction (S)	Running Time (S)	Time of Feature Extraction (S)	Running Time (S)
CCLDA	8.30	45.72	28.02	48.39	65.44	73.67
M2DPCA+CCLDA (2×4)	8.70	45.83	21.36	42.61	41.72	49.97
M2DPCA+CCLDA (4×4)	18.92	56	37.48	58.22	66.14	74.36
M2DPCA+CCLDA (4×8)	18.83	56.84	37.66	58.94	67.73	75.84

M2DPCA +CCKDA (2×4)	10.15	50.74	25.47	48.79	47.29	53.31
M2DPCA +CCKDA (4×4)	21.73	59.51	39.46	62.18	68.42	76.37
M2DPCA +CCKDA (4×8)	22.17	60.26	39.75	62.83	68.79	77.62

Table I and Table II respectively show the recognition performance of M2DPCA+CCKDA with different sub-images on ORL and XM2VTS databases. To evaluate the performance of the proposed method, we also conduct some typical methods on the two databases. Table III shows the running time of different methods on XM2VTS database. From the experimental results, we can get some conclusions as follow:

(1)Firstly, we can see that M2DPCA+CCLDA and M2DPCA+CCKDA can get much higher face recognition rate than CCLDA in both databases. Moreover, the performance of the proposed method has increased a lot than the other two methods. It proves that the new method is effective in solving nonlinear problems by utilizing kernel tricks.

(2)Table I shows that the performance of different methods on ORL database. We can see that the M2DPCA+CCLDA and M2DPCA+CCKDA can get desired results even in less training samples. Especially, when we have more than 5 training samples, the recognition rate can get 100%. But the performance of the new method is better than M2DPCA+CCLDA. Finally, we can see that the number of sub-images has some influence on the performance of the modular methods. Overall, the performance is desired. From Table II, we can get the same story as Table I.

(3)Table III shows the running time of different methods on XM2VTS. As the number of training samples is gradually increasing, both feature extraction time and running time of different algorithms are increasing.

Moreover, the computational time is affected by the number of sub-images in M2DPCA, and the time is much less when we partition the original images into 2×4 sub-images.

IV. COUCLUSIONS

In this paper, we propose an improve face recognition method. The new method can fully make use of the advantages of M2DPCA and CCKDA. By utilizing kernel tricks, we can fully use the information in non-linear high-dimensional feature space, which can provide more effective information contained in the higher order relationships among the image pixels of a face pattern. Moreover, by adopting M2DPCA to reduce the dimension, we can reduce the computation complexity and preserve the local feature of the images. Extensive experiments show the efficiency of the proposed method.

REFERENCES

- [1] M. Turk, A. Pentland, Eigenfaces for Recognition [J], Cognitive Neuroscience, 1991,3(1), 71-86.
- [2] Belhumeur, P., Hespanha, J., Kriegman, D., 1997. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE Trans. Pattern Anal Machine Intell. 19 (7), 711–720.
- [3] Jian Yang, David Zhang, etal. Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(1),131-137.
- [4] Xiong, H., Swamy, M.N.S., Ahmad, M.O., 2005. Two-dimensional FLD for face recognition. Pattern Recognition 38, 1121–1124.
- [5] Dass S. C., Jain A. K., Markov face models [C]. IEEE International Conference on Computer Vision. 2001,vol.2,680-687.
- [6] Dass S. C., Jain A. K., Lu X., Face detection and synthesis using Markov random field models [C].IEEE International Conference on Pattern Recognition, 2002, vol.4, 201-204.
- [7] Wang L., Zhang Y., Feng J., On the Euclidean distance of images[J]. IEEE Trans. PAMI. 2005,27 (8), 1334–1339.
- [8] Zhen Lei, Stan Z. Li. Contextual Constraints Based Linear Discriminant Analysis [J]. Pattern Recognition Letters. 2011, 32, 626-632.
- [9] Jian Yang, Alejandro F. Frangi, Jing-yu Yang, David Zhang, Senior Member, IEEE, and Zhong Jin, KPCA Plus LDA: A Complete Kernel Fisher Discriminant Framework for Feature Extraction and Recognition [J], IEEE Trans. PAMI. 2005,27(2), 230-244.