Contextual Constraints Based Kernel Discriminant Analysis for Face Recognition

Xian Wu #1, Xiao-Qi Sun 2, Xiao-Jun Wu 3, Zhen-Hua Feng &4
# School of Humanities, Jiangnan University
Wuxi 214122, China
1xian_wu@126.com

* School of IoT Engineering, Jiangnan University
Wuxi 214122, China
2sxqghl126.com
3xiaojun_wu@163.com

& Centre for Vision, Speech and Signal Processing, University of Surrey
Guildford GU2 7XH, United Kingdom
4z.feng@surrey.ac.uk

Abstract—In this paper, an improved subspace learning method using contextual constraints based linear discriminant analysis (CCLDA) is proposed for face recognition. The linear CCLDA approach does not consider the higher order non-linear information in facial images. However, the wide face variations posed by some factors, such as viewpoint, illumination and expression, existing in non-linear subspaces may lead many difficulties in face recognition and classification. To counteract the above problem, we incorporate the contextual information into kernel discriminant analysis by using kernel PCA, which provides more useful information for face recognition and classification. Experimental results on ORL, Yale and XM2VTS databases invalidate the effectiveness of the proposed method.

I. INTRODUCTION

Face recognition is one of the most popular topics in pattern recognition and computer vision, which involves problems of face detection, feature extraction and classification. To a certain extent, feature extraction is the most fundamental process in face recognition, which directly affects the performance of a face recognition system. Numerous algorithms have been reported to solve the face feature extraction problem in the past half century. Generally, these algorithms could be divided into two categories: linear [1] [2] [3] and non-linear approaches [4] [5] [6]. Linear approaches, such as PCA [1] and LDA [2], which find a projection matrix by defining different criterion function to map the observations from the original high dimensional image space to a low dimensional feature space and subsequently attempt to recognition or classification. However, linear methods are not always suitable for representing the internal structure of a given data set. Therefore, kernel technique was proposed to solve this problem, such as KPCA [7] [8], KDA [9] [10] [11]and so on. Kernel-based methods extend linear subspace methods to non-linear representations and have achieved great success in pattern recognition. One of the most important improvements on kernel-based approaches is the KPCA+LDA method, proposed by Yang et al. [12], which is a two-phase kernel Fisher Discriminant analysis (KFD) framework based on rigorous theoretical derivation in Hilbert space. It shows the essence of kernel discriminant analysis and gives simple steps for feature extraction.

Unfortunately, a fundamental problem of the above face feature extraction methods (both linear and non-linear) is that they do not consider the spatial relationship among pixels in training images. Human face images distribute sparsely in a high dimensional feature space and the dependence among pixels is very useful and important for applications to the problems of recognition and classification. Thus, using contextual information in images has been to be one of the most important techniques in image understanding and provides useful knowledge for recognition and classification [13] [14] [15]. Contextual constraints based linear discriminant analysis (CCLDA) proposed by Lei et al. [14]incorporates the contextual information into LDA and has been successfully applied the proposed CCLDA to face recognition. The key of CCLDA is to add a contextual constraint to the classical LDA algorithm and subsequently apply this to facial feature extraction. However, the CCLDA is a linear method and hard to describe the non-linear variations in human face images. To solve this problem, we proposed an contextual constraints based kernel discriminant analysis (CCKDA) by combing the KPCA with CCLDA. The effectiveness of CCKDA algorithm on face image recognition can be found obviously from our experimental results on several well-known face databases, ORL, XM2VTS and Yale face databases. According to the framework of KPCA+LDA [12], our proposed two-phase framework first map the training observation vectors to a new space using kernel technique and then perform the CCLDA algorithm at mapped space to achieve a further face feature extraction. Finally, the nearest neighbour classifier is adopted in the face recognition phase.

The rest of this paper is organized as follows. Section 2 details the basic theory of the proposed CCKDA algorithm. Experimental results on ORL, Yale and XM2VTS databases
are presented in Section 3. Finally, conclusions are drawn in the last section.

II. CCKDA-BASED FEATURE EXTRACTION

A. KPCA

Suppose we have a set of training face images \( X = \{X_1, X_2, \cdots, X_N\} \), we can obtain the corresponding training set by convert 2D images to 1D vectors:

\[
Y = \{y_1, y_2, \cdots, y_N\},
\]

where \( N \) is the number of training samples, \( y_n \in \mathbb{R}^M \) is the \( n \)th high dimensional training vector. We can map the training vectors to a low dimensional feature space \( F \) from the original space \( \mathbb{R}^M \) by using a non-linear mapping \( \Phi \):

\[
\Phi : y \in \mathbb{R}^M \mapsto \Phi(y) \in F.
\]

Then we can obtain a feature level description of the original training set \( Q = \{\Phi(y_1), \Phi(y_2), \cdots, \Phi(y_N)\} \) and the corresponding covariance matrix:

\[
\Sigma_f = \frac{1}{N} \sum_{j=1}^{N} (\Phi(y_j) - \Phi(y)) (\Phi(y_j) - \Phi(y))^T,
\]

where \( \Phi(y) = \frac{1}{N} \sum_{j=1}^{N} \Phi(y_j) \). The orthonormal transformation matrix \( P = [p_1, p_2, \cdots, p_m] \) can be obtained by applying the method in [7] and [12] to the covariance matrix \( \Sigma_f \).

Thus, for an given test observation vector \( y' \), the reduced dimensionality description \( z = [z_1, z_2, \cdots, z_m] \) at the space \( F \) can be computed by:

\[
z = P^T \Phi(y'),
\]

where \( \mathbb{R}^m \). Up to now, the test observation vector \( y' \) has been mapped from the space \( \mathbb{R}^M \) to \( \mathbb{R}^m \).

In the KFD framework by using KPCA+LDA [12], the classical LDA algorithm is performed to do a further feature extraction at the space \( \mathbb{R}^m \). However, the classical LDA algorithm does not consider the correlations among pixels in images. To improve the KPCA+LDA, we use the CCLDA to replace the classical LDA to propose our CCKDA approach.

B. CCKDA

In classical LDA approach, the aim is to learn a projection matrix to maximize the following projective function:

\[
J = \frac{W^T S_b W}{W^T S_w W},
\]

where \( S_b \) is the between class scatter matrix and \( S_w \) is the within class scatter matrix. In the proposed framework, \( S_b \) and \( S_w \) are defined by:

\[
S_b = \frac{1}{N} \sum_{i=1}^{L} N_i (m_i - m)(m_i - m)^T,
\]

\[
S_w = \frac{1}{N} \sum_{i=1}^{L} \sum_{j=1}^{N_i} (z_i - m_i)(z_i - m_i)^T,
\]

where \( L \) is the number of classes, \( N_i \) is the number of training samples in the \( i \)th class, \( m_i \) is the mean value of all the training samples in the \( i \)th class and \( m \) is the mean value of all the training samples.

To make the full use of contextual information among pixels in face images, Lei and Li proposed a new constraint function to the modified the objective function (5) of LDA in CCLDA [14]:

\[
\tilde{J} = \frac{W^T S_b W}{W^T S_w W + \eta J^*(W)},
\]

where \( \eta \) is the weight to balance the new projective function, and the new constraint \( J^* \) can be formulated and simplified by:

\[
J^* = \frac{1}{2} \sum_{i,j} (W_i - W_j)^2 S_{ij}
\]

\[
= W^T (D - S) W,
\]

where \( D \) is obtained by

\[
D_{ij} = \sum_j S_{ij}.
\]

The matrix \( S \) stands for the relationships among different pixels:

\[
S_{i,j} = \begin{cases} \frac{e^{||f_i - f_j||^2/\sigma^2}}{||f_i - f_j||^2/\sigma^2}, & \text{when pixels } i \text{ and } j \text{ are neighbours} \vspace{-0.2cm} \\ 0, & \text{otherwise} \end{cases}
\]

where \( f_i \) and \( f_j \) are feature vectors at \( i \) and \( j \). At the original image space, this feature vector \( f \) can be obtained by summing the gray-level values of the form all the training images and the parameter \( \sigma \) could be set to be the average distance among these feature vectors. However, when we perform this to the space transformed from KPCA, we must map this contextual constraint to the corresponding space. Here, we assume that the contextual information could be preserved from the original space to the new space.

Then, for a given observation \( y' \), the feature level description obtained by CCKDA can be computed by:

\[
v = W^T z = W^T P^T \Phi(y').
\]

At last, the nearest neighbour method can be performed for the face recognition phase.
III. EXPERIMENTAL RESULTS

Three face databases, ORL, XM2VTS and Yale, are used to evaluate the performance of the proposed CCDKA algorithm. In experiments, the database was randomly divided into training and testing set. Then the nearest neighbour classifier was used for the classification. The value of $\eta$ in Eq. (8) is set to be 0.0005 to achieve better face recognition rate. Moreover, we use the typical polynomial kernel function $k(x,y) = (\langle x, y \rangle + a)^b$, where $a = 1$ and $b = 2$ in our proposed framework.

A. Data preparation

ORL: The ORL face database has 400 images (112 * 92) captured from 40 identities, each identity contains 10 images. The images are all at frontal pose with slight tilt of the head, including illumination, expression and pose variations taken at different times. Fig. 1 shows some typical faces in ORL face database.

![Fig. 1. ORL face database](image)

Yale: The Yale face database consists of 165 images from 15 individuals under different poses, illumination conditions and expressions. The sizes of them are 160 * 121. Everyone has 11 images, the characteristics of these ones are ordered as the following rule, front lighting, wearing glasses, happy expression, left lighting, not wearing glasses, normal expression, right lighting, sad expression, tired expression, astonished expression and blink. Fig. 2 shows some of the images on Yale database.

![Fig. 2. Yale face database](image)

XM2VTS: The XM2VTS face database [16] contains 295 subjects captures from 4 sessions with pose and illumination variations. In our experiments, each identity has 2 images selected from every session. The face region have been extracted by a bounding box and resized to 55 * 51. Fig. 3 shows some typical faces in XM2VTS face database.

![Fig. 3. XM2VTS face database](image)

B. Experimental results and analysis

Fig. 4 - Fig. 6 show the comparison of the performance of face recognition rates on ORL, Yale and XM2VTS face databases by using LDA, KDA, CCLDA and CCKDA respectively.

Generally, in most of the experimental results, the face recognition performance of the CCKDA is better than all the other algorithms including LDA, KDA and CCLDA. It proves the effectiveness of the proposed CCKDA approach for solving face recognition problem. In addition, the kernel-based approaches show their advantages on ORL, Yale and XM2VTS databases compared with the linear LDA approach. The reason is that the kernel-based non-linear approach can represent complex non-rigid objects, such as human face, better than that of the classical approaches, which lead to the wide real-world applications of the kernel-bases algorithms.

![Fig. 4. Comparison on ORL face database](image)

Fig. 4 shows that the performance of CCKDA on ORL database is quite ideal even when we have a very small number of the training samples. The face recognition rates of CCKDA are all over 85% when we have more than 3 training samples. This suggests that CCKDA has a good stability on ORL database.

From Fig. 5, the performance of methods using contextual constraints, CCLDA and CCKDA, is much better than the others. Since the expression variations are rich enough in Yale face database, both the CCLDA and CCKDA algorithms makes full use of the contextual information in images and subsequently achieve a better performance than that of the LDA and KDA methods.
Face Recognition on Yale

![Comparison on Yale face database](image)

Face Recognition on XM2VTS

![Comparison on XM2VTS face database](image)

Fig. 6 shows the face recognition results on the XM2VTS face databases, which validate the advantages of kernel-based algorithms, KDA and CCKDA, i.e. the recognition rates of KDA and CCKDA are much higher than that of the LDA and CCLDA methods. The difference between the results on XM2VTS and Yale reveals that kernel-based methods are sensitive to illumination variation. According to the performance of KDA and CCKDA on XM2VTS database, the results of the proposed CCKDA algorithm is slightly better than that of the KDA method.

However, the advantages of CCKDA can be observed from the experimental results on all three face databases due to the use of contextual information and the kernel based non-linear representation together.

IV. CONCLUSION

In this paper, we propose an improved two-step face feature extraction method, CCKDA, by combing the contextual constraints based LDA to the kernel method and then apply the proposed CCKDA approach to face recognition. The motivation is that the virtue of kernel tricks, which can make full use of the non-linear information in a high-dimensional space. Furthermore, the use of contextual constraints in LDA can provide relationships among pixels in training images. Therefore, both of them provide important information face recognition problem. Extensive experimental results on three well-known face databases show the validity of our proposed method.

ACKNOWLEDGMENT

This work was supported in part by the following projects: 111 Project of Chinese Ministry of Education (Grant No.: B12018), Key Grant Project of Chinese Ministry of Education (Grant No.: 311024) and National Natural Science Foundation of P. R. China (Grant No.: 60973094, 61103128).

REFERENCES