Research on 2D Face Representation and Recognition

Hengliang Tang, Jie Zhu, Tao Liu, Mingru Zhao
Beijing Wuzi University, Beijing, China

ABSTRACT: Considering the effect of the face pose, expression and illumination, a novel feature selection scheme and a face recognition method based on sparse representation were proposed in this paper, which focused on face feature representation and classification. First, the principal component analysis (PCA) was applied on the original samples to remove redundancies and promote the calculate efficiency. After that, a feature selection scheme based on the linear discriminant analysis (LDA) was designed to extract the face features which were more suitable for classification. Finally, the sparse representation framework (RS) was adopted to collect all the face features, and address the recognition task. The experiments tested on FERET, YALE and ORL face database, demonstrate the proposed method is effective and robust to facial pose, expression and illumination conditions to some extent.

1 INTRODUCTION

Face recognition research has made a lot of progress, but there are still many problems worthy to further study. At present, most of work is devoted to improving the face representation method to improve the robustness of face recognition.

Liu, X.Z. & Ye, H.W. (2015) proposed a method of dual-kernel based two dimensional linear discriminant analysis (D-K2DLDA), by integrating multiple kernel discriminant analysis with the existing K2DFDA method. Mahbubur Rahman, S.M. et al. (2016) presented the effectiveness of selecting the discriminatory set of KCMs as the global and local face features as opposed to traditional features obtained from heuristic choice of fixed-order moments or projection of the moments for recognizing an identity. Huang, Z.H. et al. (2015) proposed a method for face recognition by using the two-dimensional discrete wavelet transform (2D-DWT) and a non-uniform patch strategy for the top-level's low-frequency sub-band. Lee, S.J. et al. (2014) proposed a newly developed 2D-LDA method based on weighted covariance scatter for face recognition. They applied 2D-LDA training matrix based on WPS-LDA to calculate the reciprocal of distance between classes and applied this weight to between class scatter matrix. Sun, W.H. (2014) researched a face recognition scheme with single training sample using 2D Gabor filter and 2D(PCM2A under varying light conditions, and the matrix-based similarity nearest neighbor classifier was used to address the recognition task. Cament, L.A. et al. (2015) proposed a new Gabor-based method which modified the grid from which the Gabor features were extracted using a mesh to model face deformations produced by varying pose. And a statistical model of the scores computed by using the Gabor features is used to improve recognition performance across pose. Ho, H.T. & Gopalan, R. (2014) accounted for the domain shift by deriving a latent subspace or domain, which characterized the multifactor variations using appropriate image formation models for each factor, and performed recognition across domain shift using statistics consistent with the tensor geometry. Al-Shiha, A.A.M. et al. (2014) proposed a method of supervised and unsupervised multi-linear neighborhood preserving projection (MNPP) for face recognition. It operated directly on tensorial data, and solved problems of tensorial representation for multi-dimensional feature extraction, classification and recognition. Wang, W. et al. (2015) presented a method of discriminant analysis on Riemannian manifold of Gaussian distributions (DARG) to solve the problem of face recognition with image sets. Lu, J. et al. (2015) proposed a simultaneous local binary feature learning and encoding (SLBFLE) method for face recognition, which was an unsupervised feature learning approach and automatically learned from raw pixels. Hafez, S.F. et al. (2015) presented a face recognition approach based on 2D face image features using subset of non-correlated and Orthogonal Gabor Filters. And the multi-stage image processing technique was designed to compensate the illumination variation. Darwish, S.M. & Mohammed, A.H. (2014) dealt with the design of intelligent 2D face recognition system using interval type-2 fuzzy logic for diminishing the effects of uncertainty formed by variations in light direction, face pose and facial expression.
In recent years, sparse representation (SR) method has been successfully applied to the study of face recognition. Gao, S. et al. (2013) extended sparse representation to kernel sparse representation (KSR). The kernel method can extract nonlinear features of samples, and then apply the sparse representation method on these nonlinear features. Elhamifar, E. & Vidal, R. (2013) studied the clustering method based on sparse subspace (SSC), which can be used to cluster the samples into a joint low dimensional subspace. He, R. et al. (2013) proposed two-stage sparse representation (TSSR), in order to solve the problem of face recognition in large scale data sets. Cheng, G.T. & Song, Z.G. (2014) presented an efficient face recognition algorithm via sparse representation in 2D Fisherface space. They transformed the 2D image into 2D Fisherface in preprocessing, and classified the testing image via sparse representation in the 2D Fisherface space.

In this paper, a new method for face representation and recognition based on face feature selection and sparse representation is proposed (Figure 1). The more efficient features are extracted for face representation, and the sparse representation framework is used for classification. The experiments demonstrate that our method can greatly improve the robustness of face representation and recognition.

2 FACE FEATURE SELECTION AND REPRESENTATION

Due to the low-dimensional face features are more redundant to some extent, if all the low-dimensional features are used for recognition, the algorithm will be more complex and time-consuming. The huge low-dimensional features may cause the optimization problem based on SR more difficult or even unsolved. Furthermore, not all the low-dimensional features are in favor of face recognition task. Studies have shown that these features are more affected by face expressions and other factors, which will make the face representation ability decreased, result the sparsity of training data descent, and influence the recognition performance. Therefore, in order to improve the facial representation ability and the computational efficiency, it is needed to effectively extract the face features which own more discriminative ability.

In this paper, the face representation method is deeply studied, and a feature selection and representation method based on LDA is proposed. Since the normalized face samples are still high-dimensional, the principal components of samples must be extracted firstly to solve the dimension crisis.

These features ensure the best approximative representation for original samples, and also promote the algorithm efficiency. While the principal components analysis is an unsupervised subspace method, the target is to extract a linear subspace for the original data, and ensure the minimal reconstruction error for all the raw data. But the projection features of samples in the linear subspace may not be necessarily propitious to classification, and the experiments in this paper also verify this viewpoint. Therefore, how to make the face representation more effective is an important problem to be solved in this paper.

An important research shows that the similarity between different samples from the same class is higher than those from different categories. Through the analysis and evaluation of a variety of image feature extraction algorithms, the linear discriminant analysis is adopted to extract and select the discriminant face features. The LDA is a supervised subspace method, and is also designed to obtain a linear subspace. While all the samples are projected to this subspace, the samples from the same class are gathered more closely, and those from different classes are gathered more dispersive. This method ensures that the original samples have good separability after projected in the subspace, so it is more suitable for face recognition.

According to the discriminant criterion, the better discriminant feature components should make the similarity between samples belonging to the same class much smaller, and that between different-class samples much larger. It also guarantees that the ratio between the within-class covariance mean value $S_W$ and between-class covariance mean value $S_B$ is smaller and the objective function can be defined as $J=S_W/S_B$. Under that, the facial representative features can be selectively chosen to improve the recognition efficiency.

3 FACE RECOGNITION

After the processing of Sec.2, the face samples are represented as feature vectors, and are also reduced into a low-dimensional subspace. In the low-dimensional subspace, the nearest neighbor classifier (NN), a linear classification method, is usually adopted for classification. Although these feature vectors is with certain separability under the low-dimensional subspace, studies show that it is still unsatisfying by simply applying the NN method on
classification stage. Therefore, this section intends to use the sparse representation method to integrate the face features, and to achieve the final classification task. Compared with the traditional method, the objective function of SR is more suitable for classification.

After the stage of feature selection and representation, each face sample is abstracted into a set of feature vectors which also can be regarded as a set of face individual descriptors. Under the premise of the higher similarity between the same-class samples, each test sample can be approximately represented by the training samples from the same class through a linear combination, according to the SR principle.

For a certain face sample \( i \), \( S_i \) denotes its low-dimensional feature vector, and all the feature vectors belonging to the training samples can be arrayed to form a training matrix \( S \) as

\[
S = [s_1, s_2, \cdots, s_n], \quad S \in \mathbb{R}^{m \times n}
\]  

(1)

Where \( n \) represents the number of training samples, and \( m \) represents the dimensionality of a feature vector.

For a test sample \( p \), the recognition task can be transformed into a \( L_0 \)-norm minimization problem under the SR framework.

\[
\hat{x} = \arg \min \| x \|_0, \quad s.t. \, Sx \approx s_p
\]  

(2)

Where \( x \) is the sparse coefficient vector with the dimension of \( m \).

In the ideal case, \( x \) contains \( k \) nonzero elements at most, and so it is usually considered that the feature vector \( s_p \) is with \( k \) sparsity respected to the training matrix \( S \). The smaller the value of \( k \) is, the sparser \( s_p \) is, and also the stronger the representation ability of \( S \) relative to \( s_p \). However, the optimization problem described by formula 2 is a famous NP-hard problem, which cannot be solved directly. The study of Wright, J. et al. (2009) demonstrated that if the optimal solution of formula 2 is sufficiently sparse, it can be approximately equivalent to solving the minimization of \( L_1 \)-norm, and then can be solved by linear programming method.

\[
\hat{x} = \arg \min \| x \|_1, \quad s.t. \, Sx \approx s_p
\]  

(3)

In the recognition stage, the optimal solution of formula 3 can be used to calculate the residuals between the test sample and all the training samples, and the training sample with the minimum residual error would be considered to be the best match. The residual error between the test sample \( s_p \) and the specific training sample \( s_i \) can be defined as

\[
r_i = \| s_p - \sum \hat{x}_i s_i \|_2
\]  

(4)

4 EXPERIMENTS AND DISCUSSIONS

4.1 Experiments

The presented method is tested on the FERET, ORL, and YALE public face databases. To further test the face recognition performance for different feature representations and classifier algorithms, and verify the effectiveness and superiority of the method in this paper, five kinds of experimental schemes are designed and carried out from the feature selection method and classification method. The experimental results are shown in Table 1 below.

<table>
<thead>
<tr>
<th>Face Database</th>
<th>Experimental scheme</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>FERET</td>
<td>PCA+NN</td>
<td>43.3</td>
</tr>
<tr>
<td></td>
<td>LDA+NN</td>
<td>58.2</td>
</tr>
<tr>
<td></td>
<td>PCA+LDA+NN</td>
<td>80.8</td>
</tr>
<tr>
<td></td>
<td>PCA+SR</td>
<td>89.1</td>
</tr>
<tr>
<td></td>
<td>Our Method</td>
<td>94.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PCA+NN</td>
<td>40.7</td>
</tr>
<tr>
<td></td>
<td>LDA+NN</td>
<td>51.9</td>
</tr>
<tr>
<td>YALE</td>
<td>PCA+LDA+NN</td>
<td>79.3</td>
</tr>
<tr>
<td></td>
<td>PCA+SR</td>
<td>88.9</td>
</tr>
<tr>
<td></td>
<td>Our Method</td>
<td>93.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORL</td>
<td>PCA+NN</td>
<td>45.3</td>
</tr>
<tr>
<td></td>
<td>LDA+NN</td>
<td>57.8</td>
</tr>
<tr>
<td></td>
<td>PCA+LDA+NN</td>
<td>82.8</td>
</tr>
<tr>
<td></td>
<td>PCA+SR</td>
<td>90.1</td>
</tr>
<tr>
<td></td>
<td>Our Method</td>
<td>96.2</td>
</tr>
</tbody>
</table>

4.2 Discussions

Compared with the experimental results in Table 1, the proposed method is more superior to the other four. In terms of feature selection and representation, the principal component and linear discriminant methods are used in face representation, which not only reduces the spatial dimension of original samples, but also makes the selected features more suitable for classification. The principal component method focuses on data reconstruction, and ensures the minimum reconstruction error. It improves the efficiency of the algorithm, and guarantees the data after dimension reduction can be near lossless representation for the original sample. The linear discriminant method devotes to the data linear separability, which makes the samples projected in the discriminant subspace easy to be detached. In recognition, the RS framework is applied on the test, which can be considered as a kind of observation of the library samples under some kind of transformation. According to the SR principle, under the premise that the samples from the same class are more similar, the probe sample can be approximately represented by the linear combination of the samples from the same class. As long as the face pose and illumination is little, the face samples can meet the premise of sparse representation. The experimental
results verify the sparse representation framework for face recognition is feasible and effective, and the presented method test on three databases get the best efficiency of 94.1%, 93.3% and 96.2%.

In the experiments, the schemes of PCA+NN and LDA+NN make only 45% and 55% of the recognition efficiency. Those are all subspace-based methods, and need abundant and spread training samples. Under that, the extracted bases for the face subspace are more scientific, and with more representation ability. Because the PCA is to ensure the minimum reconstruction error, it is usually adopted for dimension reduction. The essence of LDA is to ensure the samples from the same class projected to the subspace can be gathered more closely, while those from different classes are more disperse. Therefore, the LDA is more suitable for classification. Each of the above methods owns special advantages, and they are usually combined for the recognition problem. Firstly the PCA is used for dimensionality reduction, and then the LDA is used for selecting the more beneficial features for classification. The fusion method takes the dimensionality reduction and classification into account, so gets better recognition efficiency than the single method. The experiments shown in Table 1 verify this point.

In addition, the ORL face database focuses on a slight gesture and expression changes, and the YALE library contains rich expression changes, while the FERET library is consist of greater changes in the attitude. Comparing the experimental results on the three face databases, the performances tested on ORL are slightly better than the other, which can draw the conclusion that the method promoted in this paper is robust to mild attitude and expression to some extent. Because the experimental samples are different with others, the recognition results are not compared with others, and only compared with different algorithms tested on the same samples.

5 CONCLUSIONS

In this paper, a method for face representation and recognition based on discriminant analysis and sparse representation is proposed. The principal component and linear discriminant methods are combined to reduce the dimensionality and represent the original face samples. Then, the sparse representation framework is designed for classification and recognition, and the minimum residual error is used to measure the best match. The experiments on FERET, YALE and ORL face database verify the feasibility and effectiveness of the proposed method. This paper mainly focuses on the feature representation and the sparse representation for face recognition. In the future, the feature selection method and recognition algorithm can be intensively studied to improve the overall performance.

6. ACKNOWLEDGMENTS

This paper is supported by the Beijing Key Laboratory of Intelligent Logistics System (No.BZ0211), Beijing Intelligent Logistics System Collaborative Innovation Center, Beijing Social Science Fund, General Program of Science and Technology Development Project of Beijing Municipal Education Commission of China (No.KM201610037001, KM201710037001), Open Project of State Key Laboratory of Virtual Reality Technology and System of China (No.BUAA-VR-16KF-21).

REFERENCES


