Research on Improved PCA Face Recognition Algorithm

Wei-yi YANG, Hui ZHANG*, Meng-xia LI and Fa-nian LAI
Faculty of Information and Control Engineering, Shenyang Jianzhu University, China
*Corresponding author

Keywords: Principal component analysis (PCA), Face recognition, Feature extraction, Linear discriminant analysis (LDA).

Abstract. To solve the problems of that principal component analysis (PCA) is sensitive to light and affected its recognition rate by external interference factors, this paper puts forward an improved algorithm which based on the fusion of PCA and LDA. Firstly, the input face image is preprocessed with eliminating noise, normalization and gray level distribution equalization. Then using PCA algorithm to project the face training image to obtain low-rank feature subspace, and then the subspace face feature is derived from LDA, thus the fusion feature space is acquired. Finally the training and test samples are projected to the fusion feature space, and identify test samples based on the nearest neighbor rule. The experiment shows that this algorithm can fuse the advantage of PCA and LDA effectively and improve the robustness and efficiency of the system.

Introduction

With the development of society and network technology, information security is becoming more and more important. We need accurate personal identification on financial, judicial, security procedure, network transmission and so on. How to identify the identity of people quickly and effectively seems more and more important, too. Face Recognition is a typical biometric recognition technology. Face image is the most common image in our daily life. Face recognition is more safe and simple than other identification methods.

Face image recognition is a practical technology, which uses computer technology to extract the low-dimensional essential characteristic of face and identify the identity of people. It is a fusion product of multi-subject knowledge, such as pattern recognition, computer vision technology, image data processing and etc. As a result, it widely applies to network system, public security crime identity recognition system, monitoring system, computer recognition system and so on.

In recent years, famous research results on face image recognition have as follows: Gudivada[1] established a face recognition system which based on facial feature; Eickeler[2] used pseudo-2D hidden markov model to search face from face database. In China, to combine the urgent needs of public security criminal investigation, Yuena Ma made a preliminary study of the retrieval of facial simulated portrait and put forward a adaptive edge-detection algorithm of face image; Zhiguang Yang put forward a retrieval algorithm of face image which based on cluster Analysis, etc.

At Present, the methods of face image feature extraction and linear dimensionality reduction mainly includes principal component analysis (PCA)[3] and linear discriminant analysis (LDA)[4], etc. PCA is a unsupervised feature extraction method which based on K-L transformation. Under the constraint of the maximum variance of data in each dimension, the original data is projected to the low-dimensional linear subspace, so that the loss of information is minimal on the basis of retaining the distribution characteristics of each dimension. PCA algorithm focuses on extracting the feature information of the original data and has a high recognition rate and recognition speed, but PCA algorithm is sensitive to light and shooting angles and has not fully considered of the category label information available in the image[5], as a result, the characteristics we obtained is not the most effective class resolution feature. LDA is a method of supervision dimension reduction, it can fully considered the feature information between similar and inhomogeneity so that it can extract low-dimensional feature which has category differential diagnosis ability from high-dimensional...
data space, improving the recognition rate of image under non-uniform illumination. However, its performance is bad when the number of training samples is smaller than sample dimension.

This paper fuses both advantage and disadvantage of PCA and LDA algorithm, using PCA to project high dimension space sample to lower dimensional space to guarantee the within-class scatter matrix nonsingular, and then use LDA to extract featuring the fusion of the two characteristic subspace. Finally, this algorithm project face test set to this fusional characteristic subspace, making a classification basis for the nearest neighbor rule to improve its robustness and recognition rate.

**Face Recognition Based on PCA and LDA**

**PCA Algorithm**
Turk and Pentlad first proposed PCA algorithm for face recognition[6]. This algorithm is based on K-L transformation. Its essence is to extract feature from data. The objective of feature extraction is to obtain the transformation matrix from the measuring space to the feature space. The feature extraction must meet two principle. The one is feature space should preserve the information of measuring space as possible. Another is the dimension of feature space far less than measuring space. The principle of PCA space construction is introduced as following.

The size of every face image is $M \times N$ (M is length, N is width), we can obtain a one-dimensional column vector $D = M \times N$. This column vector $D$ is the dimension of face image, meanwhile it is the dimension of face space. Assuming that the number of face image is $N$, which is used to training, then $x_i$ represents the face feature vector of picture i, the vector $u$ represents the average value of the training image.

$$u = \frac{1}{N}\sum_{i=1}^{N} x_i$$

(1)

The difference of image vector and average vector compose covariance matrix $C$:

$$C = \frac{1}{n}\sum_{i=1}^{n}(x_i - u)(x_i - u)^T$$

(2)

We can calculate the eigenvalue and eigenvector of the face image covariance matrix by singular value decomposition theorem to determine eigen-face. Because the dimension of covariance matrix is very high and the computational complexity is large and complex, we can first use singular value decomposition theorem to calculate the eigenvalue and eigenvector of matrix $AA^T$. From the singular value decomposition theorem we can know that:

$$C\phi_i = \lambda_i\phi_i$$

(3)

At the Eq. 3, $\lambda_i$ is the ith eigenvalue of the covariance matrix $C$, $\phi_i$ is the corresponding eigenvector of the ith eigenvalue. Sorting $\lambda_1 , \lambda_2 , \ldots, \lambda_n$ from big to small and adjusting the sequence of corresponding eigenvector at the same time. Then choose a part of eigenvector to construct characteristic subspace. Assuming that the number of nonzero eigenvalue of covariance matrix as k, its value often $< n_1$ and $k << m$, but the value of k is still very high under normal case, so we can reduce some eigenvectors which have little information according to actual. The most important is that We use the method to ensure that the information ratio of the remaining eigenvectors to the total information is greater than a certain threshold ($e$), the calculation formula as equation 4:

$$e = \frac{\sum_{i=1}^{k}\lambda_i}{\sum_{j=1}^{n}\lambda_j}$$

(4)

When $e > 90\%$, It can be considered that the selected k eigenvalues have retained most of the information, preserving the previous i eigenvector. Delete the eigenvector which eigenvalue close to zero, the image corresponding to the eigenvector we preserved is highly similar with face, so we often call it as "eigen-face". Projecting training image and test image to feature space, every face image which is projected to subspace can correspond a point in the subspace. Also, each point in the
subspace correspond an image. Each image which is waited recognition can be projected to this sample space, as the basis for face recognition.

**LDA Algorithm**

LDA algorithm extracts some important feature from feature space, these feature have judging ability. At the same time, it can aggregate similar samples and separate different samples. Sleeting the feature which has the maximum ratio between between-class scatter($S_B$) and within-class scatter($S_W$). The bigger $S_B$, the better it is. $S_W$ is opposite to $S_B$. The calculation equation the following.

\[
S_B = \sum_{i=1}^{C} p_i (\mu_i - \mu)(\mu_i - \mu)^T
\]  
\[
S_W = \sum_{i=1}^{M} \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T
\]

Here, $C$ is the total of category, $P_i$ is prior probability, $\mu_i$ is the sample mean of the $C_i$-th category, $X_i$ belongs to the $i$th category sample.

Calculate the maximum value of the following Eq. 7:

\[
W = \arg \max |W^T S_B W|
\]

According to linear algebra theory, we can know $W$ meets the following equality, corresponding eigenvector $\lambda_i$(The maximum value of matrix $S_B^{-1}S_B$).

\[
S_B W_i = \lambda_i S_W W_i \quad (i = 1, 2, \ldots m)
\]

The small sample size problem which LDA often meets can be solved by the fusion of PCA-LDA. First, we can use PCA to make a dimension reduction of face image and project face image to the characteristic subspace, ensuring the within-class scatter matrix is nonsingular and using LDA to make a optimal transformation.

**Experiment and Result Analysis**

We use the ORL face database from Olivetti laboratory, Cambridge University to complete this experiment, this database has 400 face images which can be divided into 40 classes from different age, gender and race and each class of face includes 10 image from different direction, light, facial expression and facial occlusion.

**Eigenvector Experiment**

After we construct the vector matrix by the eigenvector which we obtained by the formula, projecting the face in database to eigenvectors matrix to get feature face and reconstructing the face by feature face, the effect of the reconstruction is very different if we choose different eigenvectors. As shown in Figure 1 and Figure 2.

![Figure 1. Face reconstruction of database personnel.](image1)

![Figure 2. Face reconstruction of non-database personnel.](image2)

The effect of reconstruction is quite different by Fig. 1 and Fig. 2, Fig 1 is fed from the face in the database and it participated in the feature extraction, so its feature is all in the characteristic subspace.
With the increase of the number of corresponding feature vectors, the reconstructed face is more distinct. Because the original Fig. 2 is not in the database, it can’t partake the former treatment of the face image. Therefore, the reconstructed image is always unclear. So face reconstruction out of the database is very poor or can't be reconstructed. It shows that PCA-LDA algorithm has a very obvious feature extraction effect.

Training the first five images of each group of face database and testing the residual images. Chart 1 shows the connection between the recognition rate and the projection characteristics dimension of the algorithm which is fused by PCA and LDA. Table 1 shows the relationship between the recognition rate of PCA and PCA-LDA fusion algorithm and the projection feature dimension.

Table 1. The connection between recognition rate and feature dimension.

| Feature Dimension | PCA Recognition Rate% | PCA–LDA Recognition Rate% |
|-------------------|-----------------------|---------------------------|
| 45                | 92.3                  | 94.2                      |
| 55                | 92.3                  | 94.5                      |
| 65                | 93.5                  | 95.2                      |
| 75                | 92.8                  | 95.8                      |
| 85                | 91.5                  | 96.5                      |
| 95                | 90.0                  | 94.0                      |

As can be seen from table 1, when the feature space dimension is less than 65, with the increase of the feature dimension, the recognition accuracy of the two methods gradually increases and tends to be stable. When the feature space dimension continues to increase, there are too many dimensions in the feature space. All the feature vectors do not represent the effective projection direction, and the correct recognition rate gradually decreases with the increase of the dimension.

Every group of face database has 10 images. One part is used for training and the others are used for testing. Assuming that the accumulative contribution rate of PCA as 0.9, Chart 1 puts forward the connection between recognition rate and the number of training samples under different algorithm.

Table 2. Comparison of Recognition Rate under Different Algorithms.

| Sample Number | PCA Recognition Rate/% | PCA–LDA Recognition Rate/% |
|---------------|------------------------|---------------------------|
| 1             | 61.87                  | 75.02                     |
| 3             | 80.02                  | 89.12                     |
| 5             | 86.14                  | 95.10                     |
| 7             | 94.25                  | 98.62                     |
| 9             | 97.80                  | 100                       |

From table 2, we can see that if the value of training sample is 1, the recognition rate of LDA is only 61.87%. The recognition rate of the fusion algorithm is much higher than PCA under the effect of LDA; if the number of face image less than 3, the recognition rate of both PCA and the fusion algorithm is low and the rate of the fissional algorithm is less than 89%. However, the recognition rate of the two algorithms increase as adding the number of training sample and the growth rate of PCA-LDA is higher than PCA and LDA, its recognition rate reaches 100% when the system has 8 training samples. PCA-LDA can fuse the advantage of PCA and LDA, it has a better robustness and as a result we can get a higher correct rate.

Matlab GUI implementation of PCA-LDA algorithm

In order to make the operation more convenient and intuitive, the algorithm is verified by Matlab GUI tools. The GUI interface can display the results of image preprocessing; face recognition results and so on. The specific interface is shown in Figure 3.
Conclusions

In this paper, the preprocessed image is firstly projected to the feature space of PCA, realizing data dimension reduction, then using the category label information of LDA to obtain the within-class and between-class characteristics. The result of the experiment shows that PCA-LDA fusion algorithm combines the advantages of the two algorithms and effectively solves the small sample problem. It is robust to the uneven illumination, the change of expression and the direction of shooting. Its recognition rate is higher than that of the traditional PCA algorithm.

Acknowledgement

This research was financially supported by the National Natural Science Foundation of China (61602322).

References

[1] Gudivada VN, Raghavan VV. Modeling and retrieving images by content. Information Processing and Management, 1997, 33(4):427-452.

[2] Eickeler S. Face database retrieval using pseudo 2D hidden Markov models. In: proceedings of Conference on Automatic Face and Gesture Recognition, Washington DC, USA, 2002:65-70.

[3] Bahurpi S P, Chaudhair D S. Principal component analysis for face recognition [J]. International Journal of Engineering and Advanced Technology, 2012(5):91-94.

[4] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 19(7): 711-720.

[5] Yin Fei, Feng Da-zheng. Identification of Face Based on PCA Algorithm[J]. Computer Technology and Development, 2008, 18(18): 31-33.

[6] Ma Shuai Qi. Research on improved face recognition algorithm of PCA-LDA. Journal of Shaanxi University of Technology, 2019: 35 (2):62-66.