Comparison study on SVD-based face classification

Yong Xu
Shenzhen graduate school, Harbin Institute of Technology, Shenzhen, China
Laterall2@yahoo.com.cn

Yingnan Zhao
College of physics science & information engineering, Jishou University, Jishou, China
ann_zhao_99@163.com

Abstract

In this paper, a novel SVD-based method is developed to classify face images. The experimental results show that the novel method outperforms the existing SVD-based face recognition methods. In addition, we investigate the performances on face recognition of singular values, left and right singular vectors generated from SVD and assess the existing SVD-based face recognition methods. It appears that by compared with singular values and singular vectors, SVD-based reconstruction images are more useful in classifying face images. In practice, through the SVD-based image reconstruction process, we may weaken the side effects on face classification of varying imaging conditions and facial expressions.

1. Introduction

Automatic face recognition plays an important role in our society [1-14]. Two basic and important processes should be included in a face recognition system. That is, feature extraction and face classification. Over the last two decades, a number of methods, such as principal component analysis (PCA)[1,2], linear discriminant analysis (LDA)[3,4,5] and kernel discriminant analysis [6,7,16,17] have been developed or used to extract features of face images. Among them, singular value decomposition (SVD) [10-14] is an important technique. Singular values of face images were first exploited to classify face images by Hong et al. It was reported that several theoretical properties were held for singular values of images [8,9]. Wang et al. combined singular values and other ones such as radial basis function neural network to classify faces [10,11]. On the other hand, Tian et al.[14]. concluded that the left and right singular vectors of face images included more information than the singular values and they could produce a higher recognition rate. In addition, some face recognition methods derived from SVD were also reported [12, 13,18].

In this work, we firstly investigate several SVD-based face identification methods. Furthermore, we develop a face recognition method, which exploits the SVD-based image reconstruction technique, the PCA transformation and the LDA technique in succession. The method has the following bases: some extent of noise usually exists in images and the SVD-based image reconstruction is helpful to weaken the noise and the disadvantageous information for recognition. The disadvantageous information may be caused by varying facial expressions and lighting conditions. In practice, combined with the PCA and LDA techniques, this SVD-based image reconstruction method outperforms the existing SVD-based methods such as the ones in [8,14]. The rest of this paper is organized as follows. In Section 2 we introduce the SVD technique. In Section 3 we propose the SVD-based face recognition scheme. In Section 4 we present the experimental result. Finally we offer our conclusion in Section 5.

2. SVD and image reconstruction

Theorem 1[15]. Suppose that A is an $m \times n$ real matrix. Then, there are matrices $U \in R^{m \times p}, V \in R^{n \times p}$, which satisfy

$$U^T AV = \Lambda = diag(\sigma_1, \sigma_2, ..., \sigma_p), \qquad (1)$$

where $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_p > 0$, $U = [u_1 \quad u_2 \ . \ . \ . \ u_p]$ $V = [v_1 \quad v_2 \ . \ . \ . \ v_p]$. $\sigma_1, \sigma_2, ..., \sigma_p$ are called singular values of A. u_i, v_i are called the $i-th$ left and right singular vector, respectively. The column vectors of both U and V are unitary and orthogonal vectors.

Corollary 1. The matrix A can be reconstructed accurately using $A = \sum_{i=1}^{p} \sigma_i u_i v_i^T$.

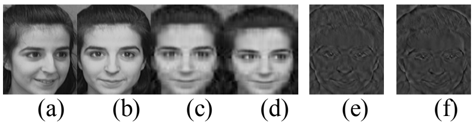
Corollary 2.[15] Suppose that the singular value decomposition of A is in the form of (1). If $A_1 = \sum_{i=1}^{k} \sigma_i u_i v_i^T, k \leq p$, then $\| A - A_1 \|_F^2 = \sigma_{k+1}^2 + \sigma_{k+2}^2 + ... + \sigma_p^2$.

Corollary 2 shows that if the value of $\sum_{i=k+1}^{p} \sigma_i^2$ is small, the reconstruction image will approximate to the original one well.

Face images are usually acquired under

Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP'06)
0-7695-2745-0/06 $20.00 © 2006 IEEE

varying conditions. For instance, the facial expression and pose are usually not fixed, which makes face recognition difficult. $\sigma_i u_i v_i^T (1 \le i \le p)$ may be regarded as sub-images and the reconstruction image is the sum of these sub-images. If some sub-images primarily reflect the essential characteristic of a face and they are only slightly affected by the varying conditions, we may expect that the reconstruction images based on these sub-images are approximately independent of the variant conditions. As a result, face recognition using these reconstruction images has higher accuracy than face recognition using original face images. In order to measure the linear separability between different categories, we define the inter-class distance and inner-class distance of images as follows: $D_b = \frac{1}{M}\sum_{i=1}^{M}\sqrt{(y_i - y)^T (y_i - y)}$, $D_w = \frac{1}{M}\sum_{i=1}^{M}\frac{1}{n_i}\sum_{j=1}^{n_i}\sqrt{(y_{ij} - y_i)^T (y_{ij} - y_i)}$. Here each image has been transformed into a vector by stacking its columns in advance. y_i denotes the mean of the images of the $i-th$ class, while y means the mean of all the images. y_{ij} denotes the $j-th$ image in the $i-th$ class. M , n_i are respectively the numbers of classes and the training samples in the $i-th$ class. Fig. 1 illustrates some SVD-based reconstruction images and the corresponding original images of the same subject. The inter-class distances and the inner-class distances of the original images and the reconstruction images show that compared with the original images, the reconstruction images associated with large singular values of the same subject are more similar to each other, whereas the reconstruction images of different subjects appear to be more distinctive from each other. Fig. 1 also shows that in the reconstruction images associated with small singular values and the corresponding singular vectors, the fine detail of faces comes into prominence. The corresponding ratio of the inter-class distance to the inner-class distance is quite small.



(a) (b) (c) (d) (e) (f)

Fig. 1. Two original face images (a,b) and the corresponding reconstruction images of the same subject in ORL. (c) and (d) respectively mean the construction images that are reconstructed using the first 8 left and right singular vectors and the corresponding singular values, while (e) and (f) respectively mean the construction images that are reconstructed using the last 84 left and right singular vectors and the corresponding singular values.

3. Method description

Our method takes advantages of SVD, PCA and LDA to classify faces. PCA, also known as K-L transform, is a classical dimensionality reduction technique. It has been applied to numerous image processing and feature extraction tasks. PCA transforms data into new ones whose components are uncorrelated to each other. LDA transforms samples into a novel space, in which the inter-class distance of samples is maximized and the corresponding inner-class distance is minimized. If the samples of each class are generated from a normal distribution, the LDA transform will be optimal for the above goal. Our method integrates these techniques to perform face recognition as follows.

Step 1. SVD is implemented for each face image. Then every image is reconstructed using the first k left and right singular vectors and the corresponding singular values.

Step 2. Principal component analysis (PCA) is performed based on the reconstructed training images. The first l eigenvectors of the generative matrix of PCA, corresponding to the first l largest eigenvalues are used as projecting axes. All the samples are transformed into new data by projecting them onto the projecting axes, respectively.

Step 3. Linear discriminant analysis is carried out for the data obtained by step 2, based on the following inter-class scatter matrix and the inner-class scatter matrix:

$$S_w = \sum_{i=1}^{M} p(\omega_i) E[(X - m_i)(X - m_i)^T \mid \omega_i], \quad (2)$$

$$S_b = \sum_{i=1}^{M} p(\omega_i)(m_i - m_0)(m_i - m_0)^T \quad (3)$$

where $p(\omega_i)$ and m_i are the prior probability and the mean of the $i-th$ class , respectively. m_0 is the mean of the total training samples . In practice, $E[(X - m_i)(X - m_i)^T \mid \omega_i]$ can be computed using $\frac{1}{n_i}\sum_{j=1}^{n_i}(x_{ij} - m_i)(x_{ij} - m_i)^T$, where n_i is still defined as the number of the training samples in the $i-th$ category, and x_{ij} is the $j-th$ sample of the $i-th$ class . If the number of training samples of each class is greater than 1, we can solve the following equation to obtain optimal transforming axes:

Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP'06)
0-7695-2745-0/06 $20.00 © 2006 IEEE

$S_b \varphi = \lambda S_w \varphi$. (4)

Replacing S_w with $S_w + \mu I$ allows us to get numerically stable solutions to Eq.(4), where I is the identity matrix and μ is a positive constant. The first $M-1$ eigenvectors corresponding to the first $M-1$ largest eigenvalues of Eq.(4) are used as optimal transforming axes, where M is the number of classes of face images. We project each sample datum from step 2 onto these transforming axes and take the corresponding projection coefficients as features of this sample. For the case of one training sample per class, we solve another equation $S_b \varphi = \lambda \varphi$ .Then we project every sample from step 2 onto each of the first $M-1$ eigenvectors of this equation and take the coefficients as features of this sample. The above procedure may avoid the problem that S_w it is a zero matrix for the case of one training sample per class. Finally we classify test samples using the nearest neighbor classier.

The expected effects on face classification of our method can be presented simply as follows. By step 1 we aim to make the reconstruction face images of the same subject more similar to each other. For example, for the face images with variant facial expressions, there are usually obvious differences among the images of the identical individual. Through the process of constructing images, we may discard some information corresponding to small singular values to weaken the effects on imaging of varying facial expressions. Through the PCA technique adopted in step 2, we can obtain data with uncorrelated components and effectively reduce the dimension of data. Via step 3, the data are further transformed to achieve the maximum ratio of the inter-class distance to the inner-class distance.

4. Experimental results

The following experiment is performed on the ORL face database with the nearest neighbor classifier. The first 1,2,3,4 images of every subject are regarded as training samples and the others are taken as test samples , respectively.

This experiment aims to illustrate the performance of our method, the singular-values-based method [8] and the singular-vectors-based-method [14]. In Table 1, method 1 means the singular-values-based method, which uses singular values of images to classify faces. Method 2 and method 3 use left singular vectors and right singular vectors to categorize faces, respectively. Method 4 classifies faces based on the combination of left and right singular vectors, while in method 5 face recognition is implemented using the concatenation of three kinds of data from SVD. In method 6 classification is directly performed based on the SVD-reconstruction images. Method 7 means the singular-vectors-based-method [14]. Method 7 works as follows. It firstly perform the SVD decomposition for the mean image of each class. Then method 7 projects every training sample onto the left singular vectors and the corresponding right singular vectors to obtain its features [14]. So does every test sample.

Table 1 shows that our method achieves higher accuracy than the other methods, including PCA and LDA. Moreover, the experiment shows that method 6 outperforms method 1, method 2, method 3, method 4 and method 5. Face classification based on the left and right singular vectors also does not result in high accuracy. Though a face image can be reconstructed by using the left singular vectors, the right ones and the singular values generated from SVD, classification based on the concatenation of the three kinds of data does not bring a delightful recognition accuracy, either. The following fact may be one of the reasons. The simple integration of left singular vectors, the right ones and the singular values neglects the hint of how the information should be assembled. Therefore, it is difficult for method 5 to perform well by virtue of this simple integration. On the other hand, method 6 classifies faces well by using the SVD-based reconstruction images.

The fact that our method is better than method 6 implies that besides SVD-based reconstruction procedure, the PCA technique and the LDA transforms are of great benefit to face recognition. Moreover, Table 2 indicates that the reconstruction images have a much larger ratio of the inter-class distance to the inner-class distance than the original images, whose ratio value is 1.7. This means that the SVD-based reconstruction images have greater linear separability than original face images.

Table 1. The highest recognition accuracies (%) obtained using different methods on ORL

The number of training sample	1	2	3	4
Our method	74.4	85.3	88.9	92.1
Method 1	41.1	50.3	53.6	60.8
Method 2	67.8	73.4	75.7	77.9
Method 3	33.9	41.9	46.8	52.1
Method 4	66.4	75.0	77.5	82.1
Method 5	41.1	50.3	53.6	60.8

Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP'06)
0-7695-2745-0/06 $20.00 © 2006 IEEE

IEEE COMPUTER SOCIETY

Method 6	74.4	83.8	86.8	90.8
Method 7	49.7	61.9	66.4	77.1

Table 2. The ratio of the inter-class distance to the inner-class distance of the reconstruction images.

The number Of left(right) singular vectors	10	15	20	25	30	35
The ratio for the reconstruction images	5.1	4.9	4.8	4.8	4.7	4.7

5. Conclusion

It is revealed that the best way of using the SVD-technique for face recognition is to classify faces by virtue of the SVD-based reconstruction face images rather than the corresponding singular values or singular vectors. The SVD technique is capable of filtering the fine detail of face images. For a face image, the reconstruction image that is obtained using the first several singular values and the corresponding singular vectors, primary reflects the holistic characteristic of faces. As a result, the linear separability of the reconstruction images can be greater than the original images and face recognition using these reconstruction images can obtain a higher accuracy

The proposed novel SVD-based face recognition method that aims at combing the SVD technique, PCA and LDA to achieve better classification performance does outperform other SVD-based methods and bring very promising classification results.

Acknowledgments

This article is partly supported by National Natural Science Foundation of China (No. 60602038, 60472060, 60473039 and 60472061) and Natural Science Foundation of Guangdong Province, China (No. 06300862).

References

[1] L. Sirovich, M. Kirby. Low-dimensional procedure for characteristicization of human faces[J]. Journal of Optical Society of America, 1987, 4(3): 519-524.

[2] M. Kirby and L. Sirovich. Application of the KL Procedure for the Characteristicization of Human Faces[J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 1990,12(1): 103-108.

[3] J. Yang and J.Y. Yang. Optimal FLD Algorithm for Facial Feature Extraction[A]. SPIE Intelligent Robots and Computer Vision XX: Algorithms, Techniques, and Active Vision[C], pp. 438-444, Oct. 2001.

[4]Y. Xu, J.-Y. Yang, Z. Jin. Theory analysis on FSLDA and ULDA[J]. Pattern Recognition,2003,36(12): 3031-3033.

[5] Y. Xu, J.-Y. Yang, Z. Jin. A novel method for Fisher discriminant Analysis[J]. Pattern Recognition, 2004, 37 (2): 381-384.

[6] Y. Xu, J.-Y. Yang, J. Yang. A reformative kernel Fisher discriminant analysis[J], Pattern Recognition, 2004, 37 (6): 1299-1302.

[7] Y. Xu, J.-Y. Yang, J. Lu, D. Yu. An efficient renovation on kernel Fisher discriminant analysis and face recognition experiments[J]. Pattern Recognition, 2004,37 (10): 2091-2094.

[8] Z. Hong. Algebraic feature extraction of image for recognition[J]. Pattern Recognition, 1991, 24: 211-219.

[9] Y.-Q. Cheng, et al. Human face recognition method based on the statistical model of small sample size[A]. SPIE Proc: Intell. Robots and Comput. Vision[C], Vol.1607: 85-95, 1991.

[10] Y.-H. Wang, T.-N.Tan, Y. Zhu. Face Identification Based on Singular Values Decomposition and Data Fusion[J]. Chinese Journal of Computer (in Chinese), 23(6), June 2000.

[11] Y. H. Wang, T. N. Tan and Y. Zhu. Face verification based on singular value decomposition and radial basis function neural network[A]. Proc. 4th. Asian Conference on Computer Vision (ACCV)[C], pp.432-436, Taipei, Taiwan, 2000.

[12] H. Wang and Na. Ahuja. Facial expression decomposition[A]. IEEE International Conference on Computer Vision (ICCV) [C], pp. 958-965, , Nice, France, 14-17 October, 2003.

[13] R. Chellappa, C. Wilson, S. Sirohey. Human and machine recognition of faces: A survey[A]. Technical Report CAR-TR-731, CS-TR3339, University of Maryland, August 1994.

[14] Y. Tian, T. Tan, Y. Wang, and Y. Fang. Do singular values contain adequate information for face recognition?[J]. Pattern recognition, 36(3),2003: 649-655.

[15] G. H. Golub, C. V. Loan. Matrix Computations, 2nd Edition[M]. Baltimore: The Johns Hopkins University Press, 1989.

[16] Y. Xu, D. Zhang, Z. Jin, M. Li, J.-Y. Yang. A fast kernel-based nonlinear discriminant analysis for multi-class classification[J]. Pattern Recognition, ,2006, 39(6) : 1026-1033..

[17] Y. Xu, J.-Y. Yang, J.-F. Lu. An efficient kernel-based nonlinear regression method for two-class classification [A]. Proceedings of 2005 International Conference on Machine Learning and Cybernetics[C], pp.4442-4445, Guangzhou, China, August, 2005.

Proceedings of the 2006 International Conference on Intelligent
Information Hiding and Multimedia Signal Processing (IIH-MSP'06)
0-7695-2745-0/06 $20.00 © 2006 IEEE

IEEE COMPUTER SOCIETY