Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®
Pattern Recognition Letters 25 (2004) 429–436
Pattern Recognition Letters
www.elsevier.com/locate/patrec

An improved face recognition technique based on modular PCA approach

Rajkiran Gottumukkal, Vijayan K. Asari *

VLSI Systems Laboratory, Department of Electrical and Computer Engineering, Old Dominion University, 231 Kaufman Hall, Norfolk, VA 23529-0246, USA

Received 13 September 2002; received in revised form 3 November 2003

Abstract

A face recognition algorithm based on modular PCA approach is presented in this paper. The proposed algorithm when compared with conventional PCA algorithm has an improved recognition rate for face images with large variations in lighting direction and facial expression. In the proposed technique, the face images are divided into smaller sub-images and the PCA approach is applied to each of these sub-images. Since some of the local facial features of an individual do not vary even when the pose, lighting direction and facial expression vary, we expect the proposed method to be able to cope with these variations. The accuracy of the conventional PCA method and modular PCA method are evaluated under the conditions of varying expression, illumination and pose using standard face databases.

Keywords: PCA; Face recognition; Modular PCA; Pose invariance; Illumination invariance

1. Introduction

Face recognition is a difficult problem because of the generally similar shape of faces combined with the numerous variations between images of the same face. The image of a face changes with facial expression, age, viewpoint, illumination conditions, noise etc. The task of a face recognition system is to recognize a face in a manner that is as independent as possible of these image variations.

Automatic recognition of faces is considered as one of the fundamental problems in computer vision and pattern analysis, and many scientists from different areas have addressed it. Chellappa et al. (1995) presented a survey on several statistical-based, neural network-based and feature-based methods for face recognition. Currently, one of the methods that yields promising results on frontal face recognition is the principal component analysis (PCA), which is a statistical approach where face images are expressed as a subset of their eigenvectors, and hence called eigenfaces (Sirovich and Kirby, 1987; Turk and Pentland, 1991;

* Corresponding author. Tel.: +1-757-683-3752; fax: +1-757-683-3220.
E-mail address: vasari@odu.edu (V.K. Asari).

0167-8655/$ - see front matter
doi:10.1016/j.patrec.2003.11.005

Moghaddam and Pentland, 1997; Martinez, 2000; Graham and Allinson, 1998). PCA has also been used for handprint recognition (Murase et al., 1981), human-made object recognition (Murase and Nayar, 1995), industrial robotics (Nayar et al., 1996), and mobile robotics (Weng, 1996). But results show that the recognition rate is not satisfactory for pose variations exceeding 30° and extreme changes in illumination.

The main objective of this research is to improve the accuracy of face recognition subjected to varying facial expression, illumination and head pose. As stated before, PCA method has been a popular technique in facial image recognition. But this technique is not highly accurate when the illumination and pose of the facial images vary considerably. In this research work an attempt is made to improve the accuracy of this technique under the conditions of varying facial expression, illumination and pose. We propose the modular PCA method, which is an extension of the conventional PCA method. In the modular PCA method the face images are divided into smaller images and the PCA method is applied on each of them. Whereas in the traditional PCA method the entire face image is considered, hence large variation in pose or illumination will affect the recognition rate profoundly. Since in the case of modular PCA method the original face image is divided into sub-images the variations in pose or illumination in the image will affect only some of the sub-images, hence we expect this method to have better recognition rate than the conventional PCA method. A similar method called modular eigenspaces was proposed by Pentland et al. (1994). In this method PCA is performed on the eyes and nose of the face image.

This paper is organized as follows: Section 2 describes the conventional PCA method. Section 3 explains the modular PCA method. Section 4 describes the face databases used for testing the face recognition methods. Section 5 presents simulation results obtained by applying the PCA method and the proposed modular PCA method to the face image sets with large light and pose variations. Finally, a conclusion is drawn in Section 6.

2. Review of the PCA method

The PCA method has been extensively applied for the task of face recognition. Approximate reconstruction of faces in the ensemble was performed using a weighted combination of eigenvectors (eigenpictures), obtained from that ensemble (Sirovich and Kirby, 1987). The weights that characterize the expansion of the given image in terms of eigenpictures are seen as global facial features. In an extension of that work, Kirby and Sirovich (1990) included the inherent symmetry of faces in the eigenpictures.

All the face images in the face database are represented as very long vectors, instead of the usual matrix representation. This makes up the entire image space where each image is a point. Since the faces have a similar structure (eye, nose and mouth, position, etc.), the vectors representing them will be correlated. We will see that faces of the same class will group at a certain location in the image space. Hence the face images are represented by a set of eigenvectors developed from a covariance matrix formed by the training of face images. The idea behind eigenimages (in our case eigenfaces) is to find a lower dimensional space in which shorter vectors will describe face images. Fig. 1 illustrates this idea graphically.

2.1. Computing eigenfaces

Consider the face images in the face database to be of size L by L. These images can be represented as a vector of dimension L^2, or a point in L^2-dimensional space. A set of images therefore corresponds to a set of points in this high dimensional

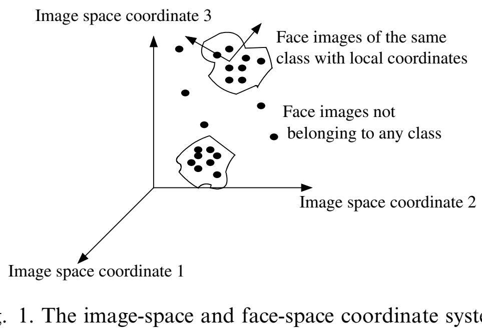

Fig. 1. The image-space and face-space coordinate system.

space. Since facial images are similar in structure, these points will not be randomly distributed, and therefore can be described by a lower dimensional subspace. PCA gives the basis vectors for this subspace (which is called the "face space"). Each basis vector is of length L^2, and is the eigenvector of the covariance matrix corresponding to the original face images.

Let $I_1, I_2, \ldots, I_M$ be the training set of face images. The average face is defined by

$$A = \frac{1}{M}\sum_{i=1}^{M} I_i \tag{1}$$

Each face differs from the average face by the vector $Y_i = I_i - A$. The covariance matrix C is obtained as

$$C = \frac{1}{M}\sum_{i=1}^{M} Y_i \cdot Y_i^{\mathrm{T}} \tag{2}$$

The eigenvectors of the covariance matrix are computed and the M' significant eigenvectors are chosen as those with the largest corresponding eigenvalues. From these eigenvectors, the weights for each image in the training set are computed as

$$W_{iK} = E_K^{\mathrm{T}} \cdot (I_i - A) \quad \forall i, K \tag{3}$$

where E_K's are the eigenvectors corresponding to the M' largest eigenvalues of C and K varies from 1 to M'.

2.2. Classification

A test image I_{test} is projected into face space by the following operation:

$$W_{\text{test}\,K} = E_K^{\mathrm{T}} \cdot (I_{\text{test}} - A) \quad \forall K \tag{4}$$

The weights W_{iK} form a vector $T_p^{\mathrm{T}} = [w_1, w_2, \ldots, w_{M'}]$, which describes the contribution of each eigenface in representing the input face image. This vector can then be used to fit the test image to a predefined face class. A simple technique is to compute distance of $W_{\text{test}\,K}$ from T_p, where T_p is the mean weight vector of the pth class. The test image can be classified to be in class p when $\min(D_p) < \theta_i$, where $D_p = \|W_{\text{test}} - T_p\|$ and θ_i is the threshold.

3. Modular PCA method

The PCA based face recognition method is not very effective under the conditions of varying pose and illumination, since it considers the global information of each face image and represents them with a set of weights. Under these conditions the weight vectors will vary considerably from the weight vectors of the images with normal pose and illumination, hence it is difficult to identify them correctly. On the other hand if the face images were divided into smaller regions and the weight vectors are computed for each of these regions, then the weights will be more representative of the local information of the face. When there is a variation in the pose or illumination, only some of the face regions will vary and rest of the regions will remain the same as the face regions of a normal image. Hence weights of the face regions not affected by varying pose and illumination will closely match with the weights of the same individual's face regions under normal conditions. Therefore it is expected that improved recognition rates can be obtained by following the modular PCA approach. We expect that if the face images are divided into very small regions the global information of the face may be lost and the accuracy of this method may deteriorate.

In this method, each image in the training set is divided into N smaller images. Hence the size of each sub-image will be L^2/N. These sub-images can be represented mathematically as

$$I_{ij}(m,n) = I_i\left(\frac{L}{\sqrt{N}}(j-1)+m, \frac{L}{\sqrt{N}}(j-1)+n\right) \quad \forall i,j \tag{5}$$

where i varies from 1 to M, M being the number of images in the training set, j varies from 1 to N, N being the number of sub-images and m and n vary from 1 to $L/\sqrt{N}$. Fig. 2 shows the result of dividing a face image into four smaller images using Eq. (5) for $N = 4$.

The average image of all the training sub-images is computed as

$$A = \frac{1}{M \cdot N}\sum_{i=1}^{M}\sum_{j=1}^{N} I_{ij} \tag{6}$$

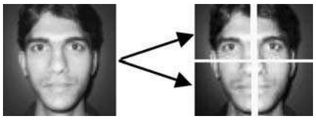

Fig. 2. A face image divided into N smaller images, where $N = 4$.

The next step is to normalize each training sub-image by subtracting it from the mean as

$$Y_{ij} = I_{ij} - A \quad \forall i, j \tag{7}$$

From the normalized sub-images the covariance matrix is computed as

$$C = \frac{1}{M \cdot N} \sum_{i=1}^{M} \sum_{j=1}^{N} Y_{ij} \cdot Y_{ij}^{\mathrm{T}} \tag{8}$$

Next we find the eigenvectors of C that are associated with the M' largest eigenvalues. We represent the eigenvectors as $E_1, E_2, \ldots, E_{M'}$. The weights are computed from the eigenvectors as shown below:

$$W_{pnjK} = E_K^{\mathrm{T}} \cdot (I_{pnj} - A) \quad \forall p, n, j, K \tag{9}$$

where K takes the values $1, 2, \ldots, M'$, n varies from 1 to Γ, Γ being the number of images per individual, and p varies from 1 to P, P being the number of individuals in the training set. Weights are also computed for the test sub-images using the eigenvectors as shown in the next equation:

$$W_{\mathrm{test}\,jK} = E_K^{\mathrm{T}} \cdot \left(I_{\mathrm{test}\,j} - A\right) \quad \forall j, K \tag{10}$$

Mean weight set of each class in the training set is computed from the weight sets of the class as shown below:

$$T_{pjK} = \frac{1}{\Gamma} \sum_{K=1}^{M'} \sum_{n=1}^{\Gamma} W_{pnjK} \quad \forall p, j \tag{11}$$

Next the minimum distance is computed as shown below:

$$D_{pj} = \frac{1}{M'} \sum_{K=1}^{M'} \left| W_{\mathrm{test}\,jK} - T_{pjK} \right| \tag{12}$$

$$D_p = \frac{1}{N} \sum_{j=1}^{N} D_{pj} \tag{13}$$

$\min(D_p) < \theta_i$ for a particular value of p, the corresponding face class in the training set is the closest one to the test image. Hence the test image is recognized as belonging to the pth face class.

4. Image databases

The performance of the conventional PCA based algorithm and the modular PCA based algorithm were evaluated with two image databases, UMIST and Yale. The UMIST database consists of images with varying pose and the Yale database consists of images with varying illumination and expressions. All the images in both the databases were normalized and cropped to a size of 64×64 pixels.

4.1. UMIST database-pose variant

For our tests we took a partial set of face images consisting of 10 images each of 20 different individuals from the UMIST face database. Each image of a person is taken at a different pose, with a normal expression. Out of the ten images of a person, only eight were used for training and the remaining two were used to test the recognition rate. Fig. 3a and b show the set of images of a person used for training and testing respectively. The choice of the training and testing images was

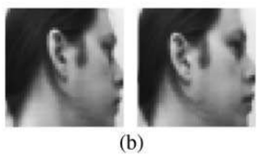

Fig. 3. Images of an individual used for (a) training and (b) testing.

made to test both the algorithms with head pose angles that lie outside the head pose angles they were trained with. The PCA and modular PCA methods may perform poorly with this selection of training and testing images, but our aim is to compare their performance for test images whose head pose angles lie outside the head pose angles of the training images.

4.2. Yale database-expression and illumination variant

The Yale database has 165 images of 15 adults, 11 images per person. The face images vary with respect to facial expression and illumination. The images have normal, sad, happy, sleepy, surprised, and winking expressions. There are also images where the position of the light source is at the center, left and right. In addition to these there are images with and without glasses. Out of the 11 images of a person, only eight were used for training and the remaining three were used to test the recognition rates. Fig. 4a and b show the set of images of a person used for training and testing respectively. The choice of the training and test images was made to facilitate comparison of performance of both the methods for test images with uneven illumination and partial occlusion.

We also conducted experiments by leaving out one image from each individual's set of 11 images during training and testing the recognition rate with the images left out. This was repeated 11 times by leaving out a different image each time. This kind of testing is referred to as leave out one testing in the remainder of the paper.

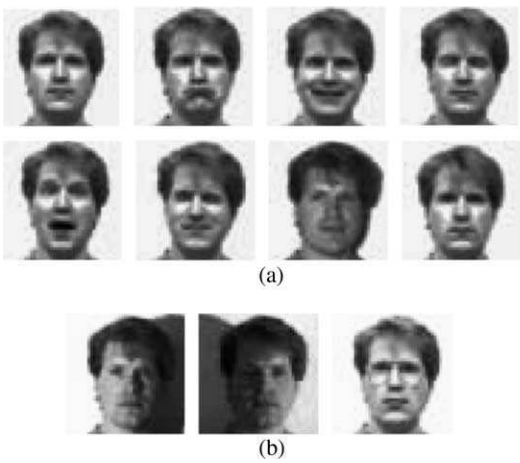

Fig. 4. Images of an individual used for (a) training and (b) testing.

5. Test results

We tested the performance of PCA and modular PCA algorithms for varying number of eigenvectors. Considering more eigenvectors results in increased recognition rates, however the increase in computational cost is linear with the number of eigenvectors. Fig. 5 shows the recognition rates of PCA and modular PCA for varying number of eigenvectors. The results shown in Fig. 5 were obtained using the Yale face database by leaving out one testing. Threshold was not used for this testing; hence there are no rejections, only correct recognition or false recognition. It can also be observed from Fig. 5 that the recognition rate is increasing in both PCA and modular PCA methods as we increase the value of M', and there is not much improvement for $M' > 30$. Similar results have been observed for values $N = 4$, 16, 64, 256 and 1024. The modular PCA results have also been compared with the results of modular eigenspaces described by Pentland et al. (1994). It has been observed that the modular PCA algorithm provides better recognition rate with the added advantage that it does not require the detection of specific features like eye, nose and mouth.

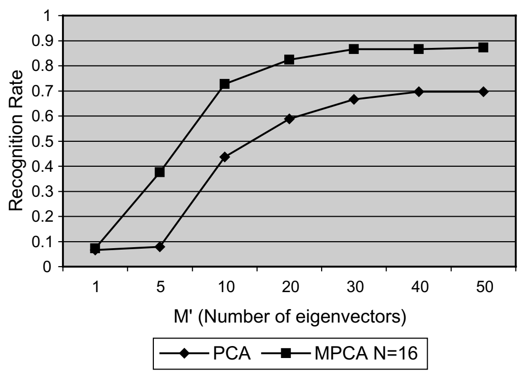

Fig. 5. Recognition rates of PCA and modular PCA with varying M'.

Tests in the remainder of this paper were conducted at $M' = 20$, i.e. eigenvectors corresponding to the 20 maximum eigenvalues of the covariance matrix. The aim of our tests was to compare the two algorithms with varying pose and illumination face images and varying M' would have the same effect on both the algorithms as shown in Fig. 5, hence only the first 20 eigenvectors were considered for the tests.

When the size of the sub-images is less than or equal to 4×4 ($N \geqslant 256$), the number of eigenvectors that can be obtained from the covariance matrix will be less than 20, since the size of the covariance matrix is less than or equal to 16×16. For the case where $L^2/N < 20$, L^2/N eigenvectors are considered. For example when $N = 256$, 16 eigenvectors are considered. For the case where $N = 4096$, the algorithm reduces to comparing the pixel values of test image and training images, pixel by pixel.

We applied the PCA method and the modular PCA method to reconstruct the test images. In the case of the PCA method the image is reconstructed as

$$I_{\text{test}} = A + E_K^{\text{T}} \cdot W_{\text{test}\,K} \tag{14}$$

The test image is reconstructed in a similar manner for the modular PCA method and is given as

$$I_{\text{test}\,j} = A + E_K^{\text{T}} \cdot W_{\text{test}\,jK} \tag{15}$$

Figs. 6 and 7 show the reconstructed images of a face image from the test set of the UMIST and Yale databases using both the methods. In the figures, the reconstructed images obtained for the modular PCA method are concatenated to facilitate visual comparison with the reconstructed image obtained for PCA method.

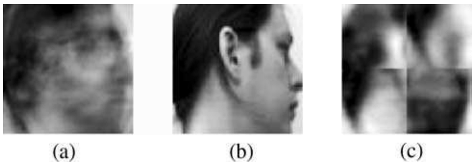

Fig. 6. (a) Reconstructed image using PCA method for a test image from the UMIST database, (b) original image from the UMIST database, and (c) reconstructed image using modular PCA method at $N = 4$ for a test image from the UMIST database.

Fig. 7. (a) Reconstructed image using PCA method for a test image from the Yale database, (b) original image from the Yale database, and (c) reconstructed image using modular PCA method at $N = 4$ for a test image from the Yale database.

5.1. Results for pose variation

In this experiment we compared the recognition rate, false recognition rate and false rejection rate of the two methods for large pose variations using the images in the UMIST database. The training and test images were chosen as described in Section 4.1. Furthermore we vary N from 4 to 4096 to observe the effect of N on face recognition. Since the size of all the images in the database is 64×64 pixels, the maximum value N can take is 4096, i.e. a sub-image is a single pixel. Fig. 8 shows the recognition rate, false recognition rate and false rejection rate for the modular PCA method with varying N. In the case of PCA the recognition rate was 0.3, false recognition rate was 0.625 and false rejection rate was 0.075.

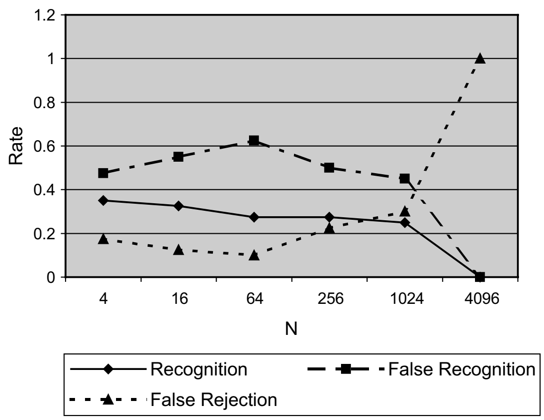

Fig. 8. Recognition, false recognition and false rejection rates of the modular PCA method for different values of N. For PCA the respective rates were 0.3, 0.625 and 0.075.

From the results we note that the modular PCA method has a slightly better recognition rate and false recognition rate at $N = 4$ and 16, but the conventional PCA method has a slightly lesser false rejection rate. Hence the proposed method does not have significant improvement over the PCA method under the condition of varying pose.

5.2. Results for expression and illumination variation

In this experiment we compared the recognition rate, false recognition rate and false rejection rate of the two methods for large expression and illumination variations using the images in the Yale database. The training and test images were chosen as described in Section 4.2. As before we vary the value of N from 4 to 4096 to observe the effect it has on face recognition. Fig. 9 shows the recognition rate, false recognition rate and false rejection rate for the modular PCA method with varying N. In the case of PCA the recognition rate was 0.44, false recognition rate was 0.31 and false rejection rate was 0.24.

A second set of experiments were performed by leaving out one testing. The results obtained for modular PCA are shown in Fig. 10. For PCA, recognition rate was 0.48, false recognition rate was 0.36 and false rejection rate was 0.16.

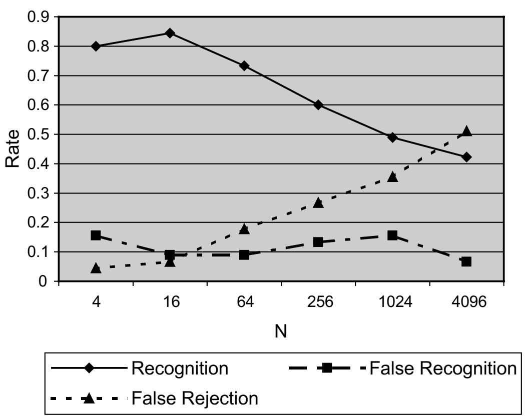

Fig. 9. Recognition, false recognition and false rejection rates of modular PCA method for different values of N. For PCA the respective rates were 0.44, 0.31 and 0.24.

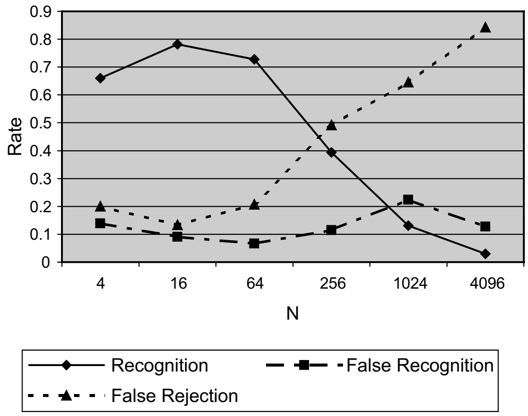

Fig. 10. Recognition, false recognition and false rejection rates of modular PCA method for different values of N. For PCA the respective rates were 0.48, 0.36 and 0.16.

We can observe from the results that the modular PCA method completely outperforms the PCA method in all aspects for N at 4, 16 and 64. However, best results were obtained for N at 16. Reconstruction of one of the test images was performed using PCA and modular PCA for N at 16. The results of the reconstruction are shown in Fig. 11, the first image is the reconstructed image obtained using PCA method, the second image is the original image and the third image is the concatenation of the reconstructed images obtained using the modular PCA method for $N = 16$.

The PCA based method was not very effective under the conditions of varying illumination, since it considers the global information of each face image and represents them with a set of weights. Under this condition the weight vectors of the test image will vary considerably from the weight vectors of the training images with normal illumination, hence it is difficult to identify them correctly. The huge improvement in the case of modular PCA was observed since the face images were divided into smaller regions and the weight

Fig. 11. Reconstructed images for a test image with varying illumination using the PCA and modular PCA method.

vectors were computed for each of these regions, hence weight vectors will be more representative of the local information of the face. Therefore for variations in illumination, the weights of the face regions not affected by varying illumination closely match with the weights of the same individual's face regions under normal conditions. This leads to better recognition results using modular PCA as observed in the experimental results.

6. Conclusion

A modular PCA method, which is an extension of the PCA method for face recognition has been proposed. The modular PCA method performs better than the PCA method under the conditions of large variations in expression and illumination. For large variations in pose there is no significant improvement in the performance of modular PCA. For face recognition, the modular PCA method can be used as an alternative to the PCA method. In particular, the modular PCA method will be useful for identification systems subjected to large variations in illumination and facial expression.

References

Chellappa, R., Wilson, C.L., Sirohey, S., 1995. Human and machine recognition of faces: A survey. Proc. IEEE 83 (5), 705–740.

Graham, D.B., Allinson, N.M., 1998. Characterizing virtual eigensignatures for general purpose face recognition. In: Face Recognition: From Theory to Applications, NATO ASI Series F, Computer and Systems Sciences, vol. 163, pp. 446–456.

Kirby, M., Sirovich, L., 1990. Application of the Karhunen–Loeve procedure for the characterization of human faces. IEEE Trans. Pattern Anal. Machine Intell. 12 (1), 103–108.

Martinez, A.M., 2000. Recognition of partially occluded and/or imprecisely localized faces using a probabilistic approach. In: Proc. of Computer Vision and Pattern Recognition, vol. 1, pp. 712–717.

Moghaddam, B., Pentland, A., 1997. Probabilistic visual learning for object representation. IEEE Trans. Pattern Anal. Machine Intell. PAMI-19 (7), 696–710.

Murase, H., Kimura, F., Yoshimura, M., Miyake, Y., 1981. An improvement of the auto-correlation matrix in pattern matching method and its applications to handprinted 'HIRAGANA'. Transactions on IECE J64-D(3).

Murase, H., Nayar, S., 1995. Visual learning and recognition of 3-D objects from appearance. Int. J. Computer Vision 14, 5–24.

Nayar, S.K., Nene, N.A., Murase, H., 1996. Subspace methods for Robot vision. IEEE Trans. Robot. Automat. RA-12 (5), 750–758.

Pentland, A., Moghaddam, B., Starner, T., 1994. View-based and modular eigenspaces for face recognition. IEEE Conf. on Computer Vision and Pattern Recognition.

Sirovich, L., Kirby, M., 1987. A low-dimensional procedure for the characterization of human faces. J. Opt. Soc. Amer. A 4 (3), 519–524.

Turk, M., Pentland, A., 1991. Eigenfaces for recognition. J. Cognitive Neurosci. 3 (1).

Weng, J.J., 1996. Crescepton and SHOSLIF: towards comprehensive visual learning. In: Nayar, S.K., Poggio, T. (Eds.), Early Visual Learning. Oxford University Press, pp. 183–214.