Lecture Notes in Artificial Intelligence, vol. 1793, pp. 197-206, April 2000,
Springer-Verlag press http://www.springer.de/comp/lncs/index.html
(MICAI-2000, Acapulco)

Eigenfaces versus Eigeneyes: First Steps Toward Performance Assessment of Representations for Face Recognition

Teófilo Emídio de Campos
Rogério Schmidt Feris
Roberto Marcondes Cesar Junior

{teo, rferis, cesar}@ime.usp.br
CreatiVision - Creativision Research Group
http://www.ime.usp.br/~cesar/creativision
Department of Computer Science
DCC-IME-USP, University of São Paulo
Rua do Matão, 1010, São Paulo, SP, 05508-900, Brazil
Phone: +55 (0) 11 818 6235, Fax: +55 (0) 11 818 6134

Abstract. The Principal Components Analysis (PCA) is one of the most successfull techniques that have been used to recognize faces in images. This technique consists of extracting the eigenvectors and eigenvalues of an image from a covariance matrix, which is constructed from an image database. These eigenvectors and eigenvalues are used for image classification, obtaining nice results as far as face recognition is concerned. However, the high computational cost is a major problem of this technique, mainly when real-time applications are involved. There are some evidences that the performance of a PCA-based system that uses only the region around the eyes as input is very close to a system that uses the whole face. In this case, it is possible to implement faster PCA-based face recognition systems, because only a small region of the image is considered. This paper reports some results that corroborate this thesis, which have been obtained within the context of an ongoing project for the development of a performance assessment framework for face recognition systems. The results of two PCA-based recognition experiments are reported: the first one considers a more complete face region (from the eyebrows to the chin), while the second is a sub-region of the first, containing only the eyes. The main contributions of the present paper are the description of the performance assessment framework (which is still under development), the results of the two experiments and a discussion of some possible reasons for them.

1 Introduction

Research on automatic recognition of faces is relatively recent, but it has been addressed by a many scientists from several different areas. According to Chellapa [1], several methods have been proposed, such as statistical-based, neural networks and feature-based. Currently, one of the methods that yields one of the most 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, hence called *eigenfaces*. This representation is used together with some classification technique for face recognition, e.g. a neural network. Next section discusses with more detail this technique.

Despite the nice results that can be obtained, this technique has the disadvantage of being computationally expensive because all pixels in the image are necessary to obtain the representation used to match the input image with all others in the database. This paper presents the experiments and the results of an approach that aims at reducing the computational effort of this approach. This technique is discussed below.

Some researchers have used eigenfaces and other eigenfeatures in order to perform recognition. The term *eigenfeature* has been used by Baback in [2], referring to the application of PCA in restricted areas of the image in order to obtain the main components of feature points of the face, such as the mouth (eigenmouth), the nose (eigennose), and the eyes (eigeneyes). In this sense, Brunelli's work [3] presents some interesting results. The results reported in that work obtained using a template covering only the eyes region are surprisingly better than the results obtained using a template that covered the whole face. Baback [2] has also obtained better results with eigenfeatures that included the eyes, the nose and the mouth than with eigenfaces.

The experiments reported here belong to a broader project aiming at the establishment of a performance assessment framework for face recognition problems. In this context, the results presented here have been obtained to either confirm or reject the results achieved by Brunelli in a PCA-based system, but using a different image database and specific preprocessing. The difference between this work and Baback's is that the present experiments only consider the eyes. As it will be seen, the present results corroborate those works, thus paving the way for the implementation of PCA-based face recognition systems that are faster and more efficient, due to the fact that they consider a smaller window.

The next section describes a PCA-based recognition system. We then describe the face image database used and the generation of the data for the training the classifier and testing. Section 4 shows the obtained results. We then move on to discuss the results and future work. The last section presents this work's conclusions.

2 Methodology

The experiments require that the sub-images (face and eyes region) be extracted from the original images. In order to improve PCA classification, the segmented images are normalized so that the face and the eyes images are of the same size. The obtained images are then used to train the PCA system, and to perform the classification tests. This section describes these procedures, namely the Principal Components Analysis system, and the image database formation.

2.1 Image Database

The image database adopted is composed of sixteen images of adult people; lots of them wear glasses, moustache and/or a beard. There are also major variations in their hair lengths. Most men are white, but there are also other ethnic groups present. Moreover, the face images may vary with respect to illumination, face expressions and pose. Nevertheless, in all the considered images the two eyes are visible, i.e. There is no image with self-occlusion problems, as it would be the case for profile images.

The database has six different images for each person. Two tests have been carried out, the first one had been made using three images to train the system, and one test image; and another one using five images in the set of training for each person and one for testing. Tests with images belonging to the training set have also been performed.

The training set has been chosen randomly from the available images of each person. If we choose only “nice” images for the training set, the performance of the recognition system would decrease [4], because if that was the case, then the classification algorithm would have more difficulties in classifying faces with different pose, facial expression or different illumination conditions.

In the case of the tests with the region of the eyes, the images are generated from the original database. Such images are hand-cropped from the original face images so that only the region from the eyebrows down to a little below of the eyes are taken into account. On the other hand, in the full-face images, only the region that encloses from the forehead until the chin is actually used and, therefore, the hair is not considered in these tests.

As it has been commented above, both images sets (eyes and faces) have been hand-cropped. An alternative to this approach would be to adopt automatic methods for detecting feature points in faces (such as [7, 8 and 9]). Nevertheless, there is no well-established, general and really reliable algorithm for performing this task, which has motivated the use of hand-cropped images.

Fig. 1. Example of images used in the training phase.

Fig. 2. Example of test image.

Therefore, the image database is composed of 192 images of 16 people, being 96 images of faces and 96 of eyes and each person being represented by four different images of eyes and face. The systems are trained independently with respect to the eyes and the faces.

The original image resolution is 512 x 342 pixels with 256 gray levels. After the segmentation and normalization (described in section 2.3), both the images of eyes and of faces have been represented with resolution of 64 x 64, with 256 gray levels.

The figure 1 shows some examples of images from the database. Figure 1 shows the images used in the training set while figure 2 presents the test image.

2.2 Segmentation and Normalization

As it has already been commented, each image has been hand-cropped in order to generate two sub-images: *img1*, corresponding to the eyes region, and *img2*, encompassing the eyes, nose, mouth and part of the chin. An interactive program using specially developed GUI has been implemented. In this program, the human operator firstly clicks on the center of the left and of the right iris. From the obtained iris coordinates, the image is rotated so that the line between the two points become horizontally oriented. Next, the sub-images *img1* and *img2* are automatically cropped based on the distance between the eyes (i.e. the distance between the marked iris centered points). Therefore, let d be the distance between the clicked points. *img1* (the eyes image) is of size $0.65d$ x $1.8d$ pixels, with the clicked points being located at line $0.4d$, which implies that *img1*encloses a larger region above the in-between iris line, including the eyebrows, as desired. *img2* is obtained in an analogous way, except that it has $2.15d$ rows. These proportions have been found empirically after some experiments with the original face database.

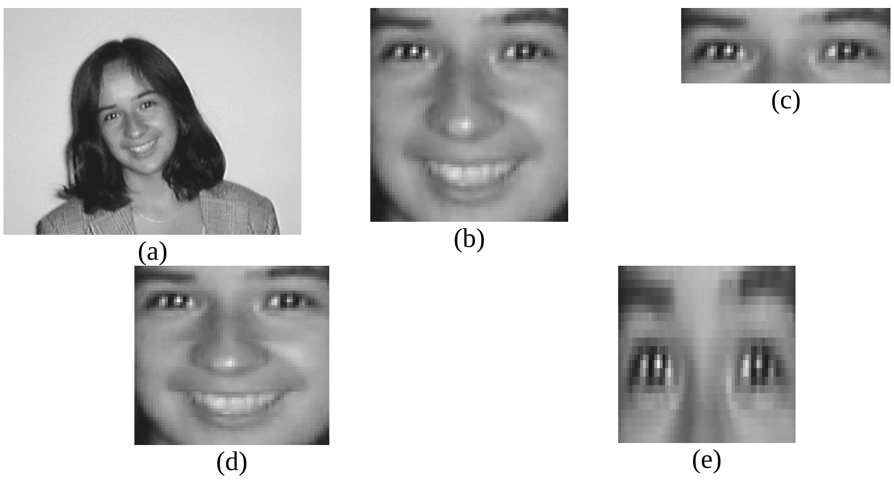

Fig. 3. Example of the pre-processing steps (from top to bottom and left to right): (a) original image (512x342); (b) result of the rotation and the segmentation of the face region; (c) resulting eyes region; (d) resizing of the face image; (e) resizing of the eyes region (both with 64x64 pixels).

Finally, because the Principal Components Analysis involves some multiplication of arrays, it is important that normalize the size of all images. This is done by resizing all images to 64 x 64 pixels. This final image size has been chosen because of the trade-off between computational cost and minimum resolution in order to guarantee that information about eyes, nose and mouth is not lost in too small image versions. Figure 3 shows the results of the segmentation and normalization processes for an image from the database.

2.3 Obtaining the Feature Spaces for Eyes and for Faces

In the first step of PCA, each 2D image is transformed in a 1D vector by appending the second image line after the first, followed by the third line, the fourth, and so on. The length of the obtained vector is $w\ h$, where w and h are the number of columns and rows of the input image, respectively (recall that, in the case of the experiments of this paper, $w = h = 64$). Therefore, each image is represented as a vector of a ($w\ h$)-dimensional space.

In the present case, when a set of face images are expressed in the above vector-like representation, the corresponding vectors tend to form clusters in the vector space, because of the common visual features of these images (eyes, nose, mouth,...). PCA allows the representation of each face (a point in the aforementioned vector space) by a few components that are necessary to distinguish between faces of different people. A database is formed from the training set of each person, and recognition is achieved by assigning the input face (to be recognized) to a specific class (person) by proximity of these representation components. This idea is applied both for the eyes-based and the face-based recognition.

2.4 The Principal Components Analysis System

To compute the principal components of faces or eyes we must do the following steps. First, we have to get the L first principal components of the image database. The principal components considered are the eigenvectors of the covariance matrix W. This matrix are taken from this product: $W = X \times X^t$, where X is the array constructed from the database, in with each column of X is a vector of an image described on the previous sub-section, and X^t is the transposition of the matrix X.

In this context, the first eigenvector of W is oriented in the direction of the largest variance among the faces, and it corresponds to an "average face" or an "average eye", because this vector have features shared to all the pictures. Therefore, it seems like a blurred face image. This blurring effect occurs because of the variations between the pictures. The second eigenvector of W characterizes the face features that are different from the first eigenvector, the third characterizes the points that are different from the two others eigenvectors, and so on. The eigenvectors are orthogonal to each other. It is important to note that the eigenvectors are sorted according to their respective eigenvalues, that is, the first eigenvector has the largest eigenvalue. The number of eigenvectors used by the classifier is fundamental the recognition rate and execution time of the sorting. The recognition rate was tested varying the number of eigenvectors and the results are described later.

A PCA-based system can be implemented by using a self-associative memory or a statistical classifier. In the case of the face recognition system by a neural network, each element of the described vector in the previous sub-section is input for the classifier and each neuron has an output that contains a reconstructed pattern. The weights of this network are gotten from the matrix *W*.

In the statistical classifier approach, *W* it is treated as a covariance matrix that is used to create a spatial basis where the covariance among its elements is minimal. In this sense, it is possible to obtain reduction of the faces or eyes basis. Therefore, we can recover an image through an input image by doing some matrix computations. The computational cost of this technique is less than the neural network approach, but it has a larger error rate. More details about PCA can be found in Romdhani [4] and Valentin [5]. The figure below shows some (four) of the eigenvectors of the image database, that has 150 images from 15 persons (10 images per person).

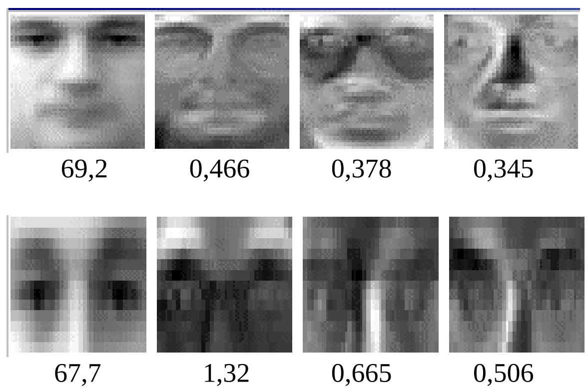

Fig. 4. Images of the first four eigenvectors with their respective eigenvalues of the faces database (above) and eyes database (below), which have been created using 150 images from 15 people (10 images per person: 5 face images and 5 eye images).

Figure 5 (in the next page) shows some examples of faces reconstructed thru the matrix *W*, that was obtained from the training image base described above.

2.5 Tests

As mentioned in section 1, the aim of the tests is to obtain a comparative study between person recognition using images that only contain the region of the eyes and images with the whole face.

Both tests have been done using the same recognition system based on PCA, the same image database, and the training and test sets corresponding to the same original images. The recognition rate of the system is adopted as the comparative criterion to analyze the results. Next section presents a discussion of our preliminary results.

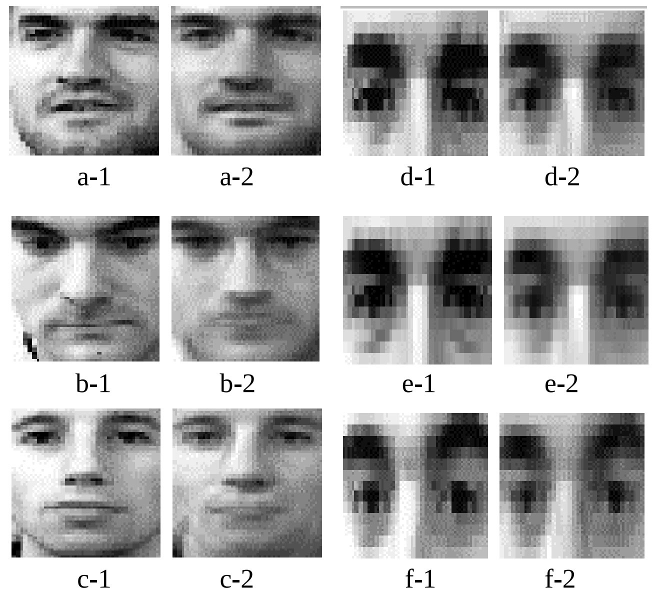

Fig. 5. a, b and c: face images; d, e and f: respective eye images. a-1, b-1, c-1, d-1, e-1 and f-1: original images; a-2, b-2, c-2, d-2, e-2 and f-2; reconstructed images. The reconstruction was done trhu the covariance matrix of a PCA system trained with 15 persons, 6 images per person.

3 Preliminary Results

As it has been noted before, the present work presents the current (preliminary) results of an ongoing performance assessment project. Section 4 presents a description of the future developmetns of the project.

Table 1 contains the results obtained by the experiments done with 16 people images. In this test set, we have used 3 images per person in the training set and 1 image per person in the test set. The results shown below was taken with training and test sets without intersection.

Table 1. Recognition rate (%) of the PCA system using eyes and face. The training was done with 3 images per person.

Number of Eigenvectors	**Eyes**	**Faces**
3	25,00	31,25
4	25,00	37,50
5	50,00	37,50
10	56,25	37,50
13	62,50	43,75
15	62,50	43,75
24	62,50	43,75
48	62,50	43,75

The above results show a poor performance because of the limited number of samples of the training set for each person. In order to improve the results, some experiments using a larger training set have been carried out, and Table 2 shows some results for this case. Here we used images from 15 people, 5 images per person in the training set and 1 image per person in the test sets.

Table 2. Recognition rate (%) of the PCA system using eyes and face. The training was carried out with 5 images per person.

Number of Eigenvectors	Eyes	Faces
3	40,00	46,67
15	73,33	66,66

Note that, if the classifier uses more than four eigenvectors, the performance of the eye recognition system is superior to the face recognition system. Moreover, the recognition rate increases significantly for both classifiers when the training set size is increased. The better performance for the eyes based classifier can be explained by two main reasons. Firstly, the inclusion of the nose and mouth region can reduce the recognition performance because face expressions implies strong distortions in this region. Furthermore, the complexity of a recognition system increases with the number of used features. This fact implies that the number of objects required to train the classifier and measure its performance increases exponentially with the number of characteristics. Therefore, adding characteristics that are either noisy or highly correlated to each other may decrease the classifier's performance if a large enough training set is not available. This fact is well known in the pattern recognition research area and can be found in further detail in [6,10,11 and 12]. Despite the lack of a larger set of classes (people) used as input and of tests with different training set sizes, the results obtained so far corroborates this theory.

4 Conclusions

This paper describes some results of a PCA-based recognition's technique applied to people recognition within the context of performance assessment. Tests with eigenfaces and eigeneyes were performed, and it was found that in most cases eigeneyes provide a superior recognition performance than eigenfaces.

In spite of eigeneyes have less information than eigenfaces, these obtained results are understandable because an increasing number of features also increases the complexity of the system. Although only a limited number of tests have been performed, the results show that images which contain only eyes are sufficient to obtain good results in face recognition. In fact, eyes differ considerably from person to person.

Observing the results, we conclude that faster face recognition systems based on PCA can be implemented by using eigeneyes instead of eigenfaces. It is easy to realize that, in the pre-processing step, less computational effort can be required using only eye images, since the images can be smaller. Therefore, also the required training set is smaller when we use eigeneyes.

We are currently working to perform more tests considering variations in the number of eigenvalues and variations in the training set size. Furthermore, the number of people to be recognized will be increased, using face databases with variations in pose, facial expression and illumination. Therefore, more reliable results will be obtained. It is important to note that all these issues of representation for recognition are central to the ongoing research on face recognition by the vision community [12].

Acknowledgments

Roberto M. Cesar Jr. is grateful to FAPESP for the financial support (98/07722-0), as well as to CNPq (300722/98-2). Rogério Feris and Teófilo Campos are grateful to FAPESP (99/01487-1 and 99/01488-8).

We are grateful to Carlos Hitoshi Morimoto (IME-USP) for useful discussions.

References

[1] R. Chellappa, C. L. Wilson and S Sirohey, “Human and machine recognition of faces: a survey”, *Proceedings of the IEEE*, vol. 83, no. 5, may 1995, pp 705-640.

[2] B. Moghaddam and A. Pentland, “Face Recognition using View-Based and Modular Eigenspaces”, *In: Proc. of Automatic Systems for the Identification and Inspection of Humans*, SPIE vol. 2277, July 1994.

[3] R. Brunelli and T. Poggio, “Face recognition: features versus templates”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, no. 10, pp. 1042-1052. October 1993.

[4] S. Romdhani, “Face recognition using principal conponents analysis”, *MSc Thesis, Department of Electronics and Electrical Engineering, University of Glasgow*, March 13th 1996.

[5] D. Valentin, H. Abdi, A. J. O'Toole (in press)., “Principal component and neural network analyses of face images: Exploratinos into the nature of information available for classifying faces by Sex.”, In C. Dowling, F.S. Roberts, P. Theuns, *Progress in Mathematical Psychology*. Hillsdale: Erlbaum.

[6] K. R. Castleman, “Digital Image Processing”, *Prentice-Hall*, Englewood Cliffs, NJ, 1996.

[7] L. Silva, K. Aizawa, M. Hatori, “Detection and tracking of facial features.” *In Proc. of SPIE Visual Communications and Image Processing'95* (VCIP'95), pp.2501/1161-2501/1172, Taipei, Taiwan, May. 1995. Awarded the SPIE (The International Society for Optical Engineering) Best student paper.

[8] H. A. Rowley, S. Baluja, T. Kanade, “Neural Network based Face Detection.”, *In IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp:23-38, January 1998.

[9] M. La Cascia e S. Sclaroff, “Fast, Reliable Head Tracking under Varying Illumination”, *Technical Report*, *Boston University* BU CS TR98-018. To appear in Proc. IEEE CVPR99.

[10] A. Jain and D. Zongker, “Feature Selection: Evaluation, Application, and Small Sample Performance”, *in IEEE Transactions on Pattern Analysis and Machine Inteligence*, pp. 153-158 vol. 19, no. 2, February 1997.

[11] L. I. Perlovsky, “Conundrum of Combinatorial Complexity”, *IEEE Transactions on Pattern Analysis and Machine Inteligence*, pp. 666-670 vol. 20, no. 6, June 1998.

[12] N. Costen Craw, T. Kato, and S. Akamatsu , “How Should We Represent Faces for Automatic Recognition?”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8):725-736., 1999