P^2CA: A NEW FACE RECOGNITION SCHEME COMBINING 2D AND 3D INFORMATION[1]

Antonio Rama, Francesc Tarres

Technical University of Catalonia, Barcelona, Spain
{alrama, tarres}@gps.tsc.upc.edu

***Abstract*- This paper presents a novel face recognition approach which uses only partial information in the recognition stage. The algorithm is based on an extension of the classical PCA and is called Partial PCA (P2CA). The P^2CA is a combined 2D-3D scheme which requires 3D face data in the training process but can process 2D pictures in the recognition stage. The strategy has been proven to be very robust in pose variation scenarios showing that the 3D training process retains all the spatial information of the face while the 2D picture effectively recovers the face information from the available data. Simulation results with a multi-view face database have shown recognition rates above 92% when using 180º texture face images in the training stage and 2D face pictures taken from different angles (from -90º to +90º) in the recognition stage.**

I. INTRODUCTION

In the last three decades almost all face recognition techniques have evolved in order to overcome two main challenges: illumination and pose variation [1]. Any of these problems can cause serious performance degradation in a face recognition system. Pose variations drastically change the appearance of a face, and in most cases these differences are larger than differences between individuals, which make the recognition task very difficult. The same statement holds for light variation.

Some of the new face recognition strategies tend to overcome both challenges from a 3D perspective. Most of these methods try to reconstruct 3D face models either from multiple images of the same person acquired from a multi camera system or directly obtained from 3D devices like 3D lasers and scanners. The advantage of using 3D data is that apart from texture, depth information is also available for face recognition making the whole system more robust toward changes in illumination and pose.

The main drawback of almost all of this kind of 3D face recognition methods is the acquisition of the 3D data in the recognition stage. The accuracy of 3D reconstruction algorithms is proportionally related to the acquisition parameters, thus, a controlled scenario with all its components well calibrated and synchronized is required, as well as the cooperation of the person who has to be recognized. These two requirements can be available in the training stage when the database is built up or extended, but are not during the recognition. Most of the surveillance and control access applications present uncontrolled scenarios where only one normal 2D picture of the subject is available for the recognition. Our research is focused on the development of a new flexible algorithm which uses 3D data for the description of the images on the database but may combine the possibility of processing faces using either 3D or 2D in the recognition stage. This concept of comparing 2D with 3D data can be also foreseen as using only partial information which has to be projected in the whole space. Since this technique is based on this concept and on Principal Component Analysis [2], it has been named Partial Principal Component Analysis (P2CA).

The rest of the paper is organized as follows. In section 2 an overview of the P^2CA algorithm is given and the mathematical procedure is formulated in detail. Section 3 describes the UPC face database and the morphing process that has been used for estimating a 180º texture representation of the faces in cylindrical coordinates. Section 4 summarizes the main results obtained in our experiments. Finally, section 5 contains the conclusions together with the current and future research which is being developed.

II. FACE RECOGNITION USING P^2CA

A. *Overview of the algorithm*

The objective of P^2CA is to implement a mixed 2D-3D method, where either 2D (normal pictures) or 3D data (180º texture images in cylindrical coordinates). can be used in the recognition stage. However, the method requires a cylindrical representation of the 3D face data for training and including new individuals in the database.

The 3D face information required in the training stage could be obtained by means of 3D laser scanners, by binocular/trinocular camera sets using stereoscopic techniques or simply by a cylindrical scanning of the face with a single camera mounted on a moving structure. In this paper, however, the main objective was to show the validity

[1] ACKNOWLEDGMENT: The work presented was developed within VISNET, a European Network of Excellence (http://www.visnet-noe.org), funded under the European Commission IST FP6 programme.

of the method and a simpler approach based on manual morphing of 2D pictures taken from different views has been followed.

Let us present the main idea of P^2CA. As in PCA the dimensionality of the face images is reduced through the projection into a set of M optimal vectors. The vectors representing the i^{th} individual are obtained as

$$\mathbf{r}_k^{\,i} = \mathbf{A}_i^T \cdot \mathbf{v}_k \qquad k = 1,..,M \tag{1}$$

where A_i^T is the transpose of the texture image representing individual i and $\mathbf{v}_k$ are the M optimal projection vectors that maximize the energy of the projected vectors $\mathbf{r}_k$ averaged through the whole database. The main difference with conventional PCA is that the whole image is represented as a 2D matrix instead of a 1D vector arrangement representing the image. Clearly, the optimization problem for finding the projection vectors has to be reformulated as will be discussed in section 2.2.

The projection described in Eq. (1) is depicted in Fig 1. The image of the subject i is represented by the M vectors $\boldsymbol{r}_k^i$. Each vector $\boldsymbol{r}_k^i$ has n components where n is the dimension of the matrix A_i in the horizontal direction (vertical when transposed).

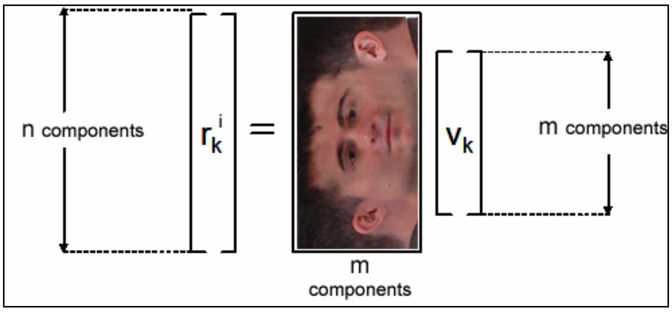

Fig 1 Description of a cylindrical coordinate image by means of projection vectors (training stage)

If complete 3D data of the individual is available the recognition stage is quite obvious. In fact, it is only necessary to convert the 3D data to cylindrical coordinates and compute the resulting M vectors $\boldsymbol{r}_k$. The best match is found for the individual i that minimizes the Euclidean distance:

$$\min_i \left\{ \xi_k = \sum_{k=1}^{M} \sum_{l=1}^{n} \left(r_k(l) - r_k^{\,i}(l) \right)^2 \right\} \qquad i = 1,..,L \tag{2}$$

where L represents the number of individuals in the database.

The main advantage of this representation scheme is that it can also be used when only partial information of the individual is available. Consider, for instance, the situation depicted in Fig.2, where it is supposed that only one 2D picture of the individual is available. Each of the 2D pictures of the subject (frontal and lateral) show a high correlation with the corresponding area of the cylindrical representation of the 3D image.

The procedure for extracting the information of a 2D picture is illustrated in Fig.3.

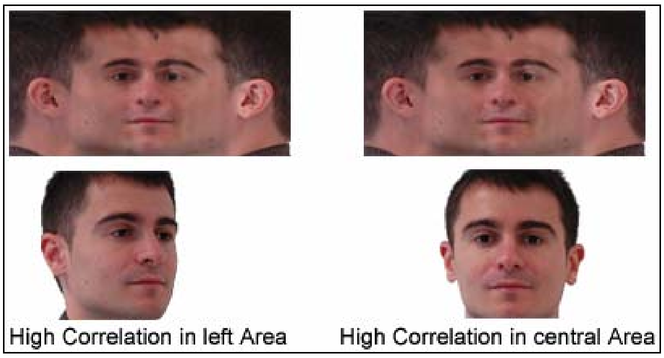

Fig 2 Comparing 2D pictures with a cylindrical representation of the subject

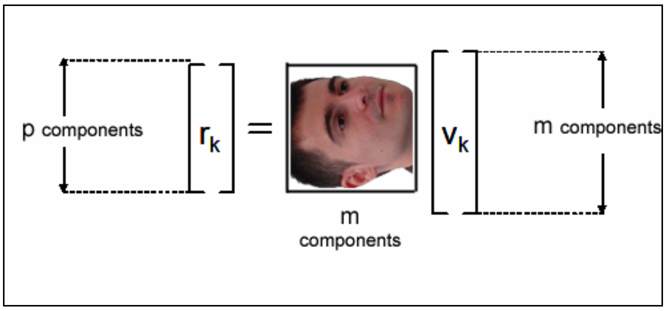

Fig 3 Projection of a partial 2D picture through the vector set v_k (recognition stage)

In this case, the M vectors $\boldsymbol{r}_k$ representing the 2D picture, have a reduced dimension p. However, it is expected that these p components will be highly correlated with a section of p components in the complete vectors $\boldsymbol{r}_k^i$ computed during the training stage. Therefore, the measure proposed below can be used to identify the partial available information (p components) through the vectors r_k^i:

$$\min_{(i,j)} \left\{ \sum_{k=1}^{M} \sum_{l=1}^{p} \left(r_k(l) - r_k^{\,i}(l+j) \right)^2 \right\}$$
$$i = 1,..,L; \quad j = 0,..,n-p \tag{3}$$

B. Obtention of the projection vectors

The set of vectors which maximize the projection of Eq. (1) may be obtained as the solution to the following optimization problem: Find v_k for k=1,..,M such that $\xi = \sum_k \sum_n (r_k^n)^T \cdot r_k^n$ is maximum, where r_k^n is defined as the projection of image n through the vector v_k and n accounts for the number of images in the training set. The function to be maximized may be expressed as:

$$\xi = \sum_k \sum_n \left(A_n^T v_k \right)^T \cdot \left(A_n^T v_k \right) = \sum_k v_k^T \left(\sum_n A_n \cdot A_n^T \right) \cdot v_k$$

which states that the optimum projection vectors may be obtained as the eigenvectors associated to the M largest eigenvalues of the mxm positive definite matrix C_s

$$C_s = \sum_n A_n \cdot A_n^T$$

This vector set will be used for feature extraction and recognition from partial information:

$$\{\mathbf{v}_1, \ldots, \underline{\mathbf{v}}_M\}$$

The procedure for feature extraction from an intensity image A consists in projecting the transposed image through every eigenvector:

$$\mathbf{r}_k = A^T \cdot \mathbf{v}_k \quad k = 1, \ldots, M$$

Therefore, a total of M feature vectors are available, with n (width) components each, for the image. The image has been compressed to a total of nxM scalars with M always being smaller than m.

When a complete image sample A (mxn) is available, the recognition stage is straightforward. First, the projected vectors of the sample image are computed using the previous equation and then, the best match is found as the individual i whose representing vectors minimize the Euclidean distance using equation (2).

The procedure is quite different from conventional PCA. Certainly, in PCA a scalar number is obtained when the vector image is projected to one eigenvector, whereas in P^2CA, an n-dimensional vector ($\mathbf{r_k}$) is obtained, when the image (in matrix form) is projected to an eigenvector. It can seem that the P^2CA approach demands more computational cost because it uses vectors instead of numbers to represent the projections. However, the number of eigenvectors $\{\mathbf{v_k}\}$ needed in P^2CA for an accurate representation is much lower than in PCA.

It should be mentioned that the mathematical theory behind this approach is similar to one recent method which has extended the conventional PCA method [2] from 1D to 2D; this technique was called 2DPCA [3].

C. Recognition of an individual from partial information

The most outstanding point of this procedure is that the image projected in the n-dimensional space does not need to have dimension mxn during the recognition stage so that partial information can be used. It is possible to use a reduced pxn ($p<m$) image which is projected to a smaller subspace.

If only partial information is used, a classification method is needed to compare the partial projection with the data in the whole space. In this case, it is not possible to use nearest neighbour classifier like in conventional PCA and correlation of partial difference methods like the criteria defined in (3) have to be applied.

III. DATABASE DESCRIPTION

A. UPC Face Database Description

The majority of available face databases are focused on frontal views; therefore, the proposed P^2CA method has been tested on an own database which has been created at the *Universitat Politècnica de Catalunya* [5] with the main purpose of testing the robustness of face recognition algorithms against strong pose and illumination variations. This database includes a total of 18 persons with 27 pictures per person which correspond to different pose views (0º, ±30º, ±45º, ±60º and ±90º) under three different illuminations (environment or natural light, strong light source from an angle of 45º, and finally an almost frontal mid-strong light source).

B. 180º Texture Training Images

The work presented in this paper is focused on the recognition algorithm; thus, a manual method has been used for the generation of the cylindrical 180º 3D texture images which will be used as 3D information for the training stage. From the 9 different views under a natural illumination only five of them are selected (frontal, ±45º and ±90º views) and morphed together as illustrated in Fig. 4. The eyes of the frontal image are used to fuse that view with the ±45º views, whereas the ears are the junction point for the combination of the ±45º and ±90º face images. 20 different 180º 3D texture images (one for each person) and their mirror images have been used for the training phase.

C. Normalization of the Test Images

A total of 486 normal 2D images have been used for the recognition stage (9 views x 3 illuminations x 18 persons). A scale factor has been computed using the frontal view picture of each person and each illumination and considering a distance of 40 pixels between the two eyes and an output resolution of 122x100. This scale factor has been applied to all non-frontal pictures. Fig. 5 shows some images used for testing the performance of P^2CA. The face images where the person is principally looking to the left side (left-view images illustrated on the first line of Fig. 5) will be considered positive angles, whereas the views where the person is looking to the right side (right-view images depicted on the bottom line of Fig.5) are defined as negative angles.

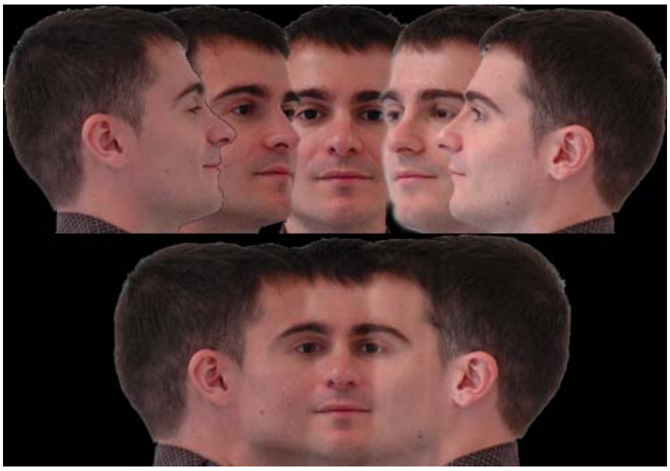

Fig 4 (a) Set of images used for the creation of the training data; (b) Example of a 180º texture training image

Fig 5 Examples of the set of test images

IV. Results

Two different experiments have been carried out. Firstly, the 162 test images under natural light have been tested separately to check the sensitiveness of the P^2CA approach towards pose variations; and secondly, all the test images under the three different illuminations (a total of 486) have been used for the computation of the recognition results in order to test the pose and illumination challenges at the same time. Table 1 and 2 present the results for both experiments depending on the angle view of the face image.

Table 1 Recognition Accuracy of P^2CA under natural light

Angle	Recog.	Angle	Recog.	Angle	Recog.
-90º	88.9%	-30º	94.4%	+45º	100%
-60º	88.9%	0º	100%	+60º	83.33%
-45º	100%	+30º	83.33%	+90º	88.9%

Table 2 Recognition under 3 different illumination conditions

Angle	Recog.	Angle	Recog.	Angle	Recog.
-90	79.63%	-30	77.77%	+45	72.22%
-60	77.77%	0	77.77%	+60	72.22%
-45	79.63%	+30	68.5%	+90	74.07%

Table 1 shows that the new technique is considerably robust towards changes in the pose of the face. On the other hand, when testing the images under the three different illuminations (especially the strong light source from an angle of 45º) the algorithm presents lower recognition accuracy as depicted in Table 2. Moreover, it can also be concluded that the positive pose variations present a lower recognition rate. The reason for this reduction in the recognition accuracy is probably due to the acquisition method of the training images. As illustrated in Fig.4a the left-view images (+45º and +90º) present a higher brightness in the part used for creating the 180º training images. This problem would be corrected in the acquisition of the extended version of the UPC multi-view face database.

Finally, in order to verify the robustness of P^2CA in front of other strategies the conventional PCA has been implemented using the same set of training and recognition images. In this case, the training set is composed of the five images of each subject (and the mirror versions) which have been used in the creation of the 180º texture images of the P^2CA method as depicted in Fig. 4a.

Table 3 Recognition Accuracy of the PCA algorithm

Method \ Exp.	**Only Pose**	**3 illuminations**
PCA	72.22%	60.37%
P^2CA	92.59%	75.51%

The results presented in Table 3 show that the new P^2CA technique is considerably robust towards changes in the pose of the face image improving the recognition accuracy of conventional PCA in a 20% when only pose is tested and in a 15% when pose and illumination are taken into account. Nevertheless, when testing the images under the three different illuminations the P^2CA algorithm presents a recognition accuracy of only 75.51% due to the strong light source from an angle of 45º that provokes several false recognition matches.

V. Conclusions and Future Work

This paper has presented a new face recognition technique which shows robustness against pose variation. Additionally, the P^2CA method permits to use either 2D or 3D data in the recognition stage, making the whole system more flexible.

After analyzing the results, it has been concluded that the main false recognition matches were due to problem of strong illumination variations; thus, an equivalent *Partial Linear Discriminant Analysis* technique is currently under development to make the whole approach less sensitive against strong illumination variations. Another way to reduce the effects of the illumination conditions is to acquire the 3D training data by more sophisticated and accurate methods like 3D laser scanners or binocular/trinocular camera sets using stereoscopic techniques. In fact, it is planned to obtain complete 3D face models using the approach presented in [4] with the cooperation of the *Politecnico de Milano* where not only texture but also depth information would be available.

References

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey" *ACM Computing Surveys*, Vol.35, Nº4, December 2003

[2] M. A. Turk, A. P. Pentland, "Face recognition using eigenfaces", *Proc. of the IEEE Computer Society Conf. on Computer. Vision and Pattern Recognition*, 1991.

[3] J. Yang, D. Zhang, A.F. Frangi, and J.Yang, "Two-Dimensional PCA: A New Approach to Appearance-based Face Representation and Recognition", in *IEEE Trans. on Pattern Analysis and Machine. Intel.*, Jan. 2004

[4] D. Onofrio, A. Sarti, and S. Tubaro, "Area Matching Based on Belief Propagation with Applications to Face Modeling", in *IEEE International Conference on Image Processing*, Singapore 2004

[5] "UPC Face Database" in http://gps-tsc.upc.es/GTAV