PARTIAL LDA VS PARTIAL PCA[1]

Antonio Rama, Francesc Tarres

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

ABSTRACT

Recently, 3D face recognition algorithms have outperformed 2D conventional approaches by adding depth data to the problem. However, independently of the nature (2D or 3D) of the approach, the majority of them required the same data format in the test stage than the data used for training the system. This issue represents the main drawback of 3D face research since 3D data should be acquired under highly controlled conditions and in most cases require the collaboration of the subject to be recognized. Thus, in real world applications (control access points or surveillance) this kind of 3D data may not be available during the recognition process. This leads to a new paradigm using some mixed 2D-3D face recognition systems where 3D data is used in the training but either 2D or 3D information can be used in the recognition depending on the scenario. Following this new concept, *Partial Linear Discriminant Analysis* (PLDA) is presented in this paper. Preliminary results have shown an improvement with respect to the Partial PCA approach [1].

1. PARTIAL INFORMATION CONCEPT

The performance of face recognition systems that use 2D intensity images depends highly on the conditions during the acquisition of the image, e.g. pose of the face, illumination, or facial expression. Since a face is a 3D object, new face recognition techniques have tried to add shape or depth information to make the system more robust towards pose and lighting variations. Additionally, 3D data acquisition is becoming faster and cheaper by means of special 3D scanner devices or multi-camera systems [2]. Therefore, 3D face recognition research is getting more and more important [3, 4, 5, 6, 7]. These 3D algorithms can be roughly divided in two categories: The approaches of the first group basically compute a depth and intensity map separately and then they perform a conventional 2D method to each modality and combine them as two different expert opinions [4,7]. The second category encloses model-based approaches that use complete 3D models of a face to perform the recognition [3,5,6]. The advantage of the first category is that it adds depth information to conventional approaches without increasing too much the computational cost; but, on the other hand, most of them are not true 3D approaches and they should be called 2.5D techniques since they may not have multi-view information. Furthermore, the input of the recognition stage of these approaches should maintain the same data format as the training images, i.e. if frontal views have been used during the training stage then a depth and/or intensity frontal image may be required in the recognition stage [4]. On the contrary, the majority of the model-based 3D face approaches intend to fit texture images on some 3D models. After this adjustment, they extract some relevant features, in most of the cases geometrical parameters, that will be used in the recognition stage. In this case, the input images for the recognition phase could be common 2D intensity images which can be available in any kind of application either under controlled or uncontrolled acquisition conditions. However, the process of fitting an intensity image on a generic model is very computationally demanding and not a very precise task. Thus, if the algorithm uses only one generic model it may not provide sufficient discrimination information, and if it uses one model for each person of the database the computation time will be excessive. Moreover, in real world scenarios, either for 2D capture camera systems or 3D sensors, only partial views of the face will be available. Therefore, this second category methods are more appropriate since they address a scenario where different data formats are used in the training and in the test stages, i.e. they intend to combine 2D and 3D information for the face recognition problem. Recently, a novel approach called Partial Principal Component Analysis (P^2CA) has been presented [1]. This approach still presents some problems to cope with illumination changes. For this reason, this paper presents a possible extension of this approach which is less sensitive to illumination changes by linking the concepts of partial information and Linear Discriminant Analysis.

The rest of the paper is organized as follows. In section 2, the fundamentals of the P^2CA technique presented in [1] are reviewed and extended to the LDA space. Some

preliminary results are shown in section 3, whereas conclusions and future work are presented in section 4.

2. EXTENSION OF LDA TO PLDA

2.1. P²CA (Partial PCA) fundamentals

The objective of P²CA is to implement a mixed 2D-3D method, where either 2D (pictures or video frames) 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 the training stage. In this paper, it is supposed that 180º cylindrical texture images, as the ones shown in Fig 1, are available for the training stage.

Each identity of the database is characterized through the projection into a set of M optimal vectors v_k (face space) which are the eigenvectors of the Covariance Matrix of the training ensemble:

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

where A_i^T is the transpose of the 3D training texture image which represents individual i and v_k are the M optimal projection vectors that maximize the energy of the projected vectors r_k^i averaged through the whole database (weights or coefficients used for the recognition). Each vector r_k^i has W components where W is the dimension of the 180º texture training image (A_i) in the horizontal direction (vertical when transposed). These vectors are the extracted features (weights in Fig 1) that will be stored in the system during the training stage and used later in the recognition stage. Summarizing, each identity of the database will be represented by a WxM matrix of features. 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. The complete process is illustrated in Fig 1, and a more detailed explanation of the mathematics related to P²CA can be found in [1,8].

In the recognition stage (upper box in Fig 1), two different cases can occur depending on the nature of the test face. 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 r_k. The best match is found for the individual i that minimizes the Euclidean distance. The main advantage of this representation scheme is that it can also be used when only partial information of the individual is available. Consider the second situation depicted in Fig 1, where it is supposed that only one 2D picture of the individual is available. In this case, the M vectors 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 r_k^i computed during the
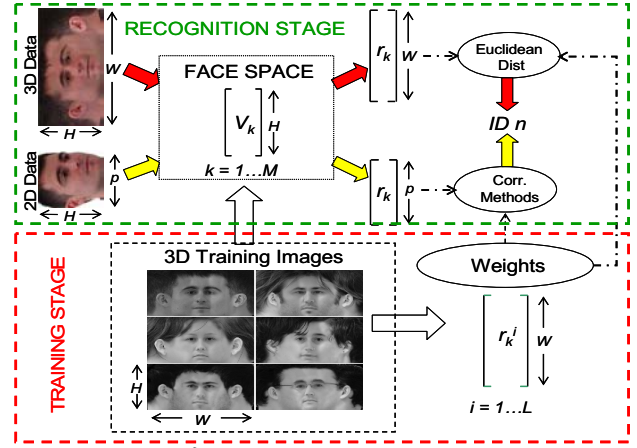

Fig.1 P²CA: General Block Diagram

training stage. Therefore, the measure proposed below can be used to identify the partial available information (p components) through the vectors r_k^i:

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

2.2. Partial Linear Discriminant Analysis (PLDA)

After analyzing the performance of P²CA [1] two major drawbacks can be extracted: The first one refers to the dimension of the feature space used for the recognition. Since P²CA is mathematically based on 2DPCA [8], one disadvantage is that more coefficients are needed to represent each image (in fact WxM coefficients). Thus, it is necessary to reduce as much as possible the number of vectors (M) of the face space where the images are projected. The second drawback is that the results presented in [1] show a certain weakness of the technique towards big illumination variations. Both disadvantages may be solved by generalizing the partial information concept with *Linear Discriminant Analysis*.

LDA or more precisely Fisherfaces [9] have demonstrated to be more robust against lighting changes than PCA (or Eigenfaces). The Linear Discriminant Analysis [9] uses the class membership information to develop a set of feature vectors where the variations of different faces are emphasized while the changes due to illumination conditions, facial expressions and orientations are de-emphasized. Formally the idea is to construct two scatter matrices that represent the between-class (S_B) and the within-class (S_W) dispersion of the training data. The between-class matrix S_B is collecting the dispersion of the mean vectors in each class with respect to the overall mean. On the other hand, the within-class matrix S_W represents the dispersion of the elements of a given class with respect to the mean of this class. This matrix is a probability weighted sum of the covariance matrix within each class. Therefore,

S_W is in some way measuring the noise of each sample with respect to the mean of its class. The Linear Discriminant Analysis method proposes to project the samples on a set of k orthogonal vectors that maximize the following function:

$$V_{opt} = \arg\max_V \left\{ \frac{\left| V^T S_B V \right|}{\left| V^T S_W V \right|} \right\} \quad (3)$$

where V_{opt} represents a matrix with k orthogonal column vectors with n components each. Clearly, the above optimization criteria implies that V_{opt} will maximize the projected distance between vectors belonging to different classes but also will try to collect together the projected samples belonging to the same class. The method explicitly uses the information of the within-class scattering matrix to optimize the clustering of data after the projection. On the contrary, while constructing the covariance matrix in the PCA approach no prior knowledge of the existence of different classes is introduced. The solution of the above optimization problem reduces to a generalized eigenvector equation provided that the within-class scattering matrix S_W is non-singular.

$$S_B \cdot v_k^{LDA} = \lambda_i S_W v_k^{LDA}, \quad k = 1,2,...,M \quad (4)$$

where v_k^{LDA} are the eigenvectors (columns of V_{opt}) associated to the M largest eigenvalues. Following the Fisherfaces [9] framework, LDA can be extrapolated to a two dimensional approach like the one presented in [8]. First, the Between-class and Within-class scatter matrixes will be defined as:

$$S_B = \sum_{i=1}^{L} P_i \cdot \left(\overline{\mathbf{A}}_i - \overline{\mathbf{A}} \right) \cdot \left(\overline{\mathbf{A}}_i - \overline{\mathbf{A}} \right)^T \quad (5)$$

$$S_W = \sum_{i=1}^{L} P_i \sum_{j=1}^{N_i} \left(\mathbf{A}_j^{(i)} - \overline{\mathbf{A}}_i \right) \cdot \left(\mathbf{A}_j^{(i)} - \overline{\mathbf{A}}_i \right)^T \quad (6)$$

where it is assumed that L classes (identities) are given with an a-priori probability of P_i. The number of samples per each class is N_i. The images in the training set are represented as HxW matrices where $A_j^{(i)}$ denotes the jth sample belonging to class i. $\overline{\mathbf{A}}_i$ is the mean image of each class and finally $\overline{\mathbf{A}}$ is the total mean image of the training ensemble, both treated also like matrices.

The Fisherfaces technique is based on reducing the dimensionality of the data vectors through the PCA before the LDA is applied. Following the same criteria, in this paper the optimal face space will be constructed from the computation of the P^{2}CA space and the two dimensional extension of LDA using the following expression:

$$V_{PLDA_k} = V_{2DFLD_k}^T \cdot V_{P^2CA_k}, \quad k = 1,2,....M \quad (7)$$

where V_{P^2CA} is the face space presented in Fig 1,

$$V_{2DFLD} = \arg\max_V \left\{ \frac{\left| V^T V_{P^2CA}^T S_B V_{P^2CA} V \right|}{\left| V^T V_{P^2CA}^T S_W V_{P^2CA} V \right|} \right\}$$

and $\quad (8)$

In Partial LDA V_{PLDA_k} will substitute the v_k vectors presented in equation 1 and Fig 1 for feature extraction. The recognition stage is the same as the one presented for P^{2}CA in Fig.1 and the measure of equation 2 will be used to compare 2D and 3D data.

3. EXPERIMENTAL RESULTS

3.1. Description of the database

The UPC face database [10] contains a total of 756 images corresponding to 28 persons with 27 pictures per person acquired under different pose views (0º, ±30º, ±45º, ±60º and ±90º) and three different illuminations (environment or natural light, strong light source from an angle of 45º, and finally an almost frontal mid-strong light source). The images have been normalized to an output resolution of 122x100 pixels. The 180º cylindrical training images have been created by manually morphing five images (0º, ±45ºand ±90º) that have been acquired in a different session than the rest of the pictures under environmental light conditions.

In order to model different illuminations, two images like the ones shown in Fig 2 have been synthetically created using an image application like Adobe Photoshop®. These two synthetic images try to simulate the 45º strong light source and the medium light source of the acquisition session without using the test material. It is preferred to perform this experiment with two illumination conditions that simulated the ones used for the test material because in real applications it is almost impossible to acquire an image under exactly the same illumination. For the feature extraction of each identity, only image (a) of Fig 2 has been used.
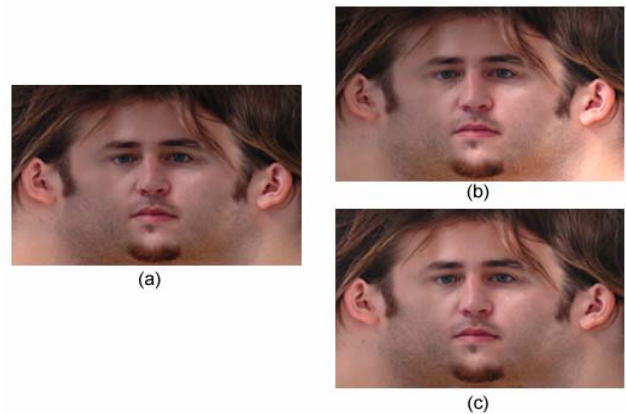

Fig.2 (a) 180º Texture Image using 5 views under environmental light conditions. (b) and (c) synthetic illumination conditions

3.2. Face Recognition

P^2CA and PLDA have been tested using the material described above. Fig 3 represents the recognition rate as a function of the face space dimension. From the results, it can be concluded that PLDA presents better results than the P^2CA technique when using less than 60 eigenvector for the face space. In this case PLDA reveals a maximal recognition accuracy of 79.49% in front of the 71.82% of P^2CA.

It should be remarked that the maximum recognition rate for P^2CA corresponds to a dimension of the face space of 21 eigenvectors, whereas the maximum for PLDA is reached using 14 eigenvectors. This represents not only an improvement in terms of recognition accuracy but also in terms of computational cost due to the feature space reduction.

If the results of Fig 3 are analyzed in more detail it could be concluded that PLDA is more robust towards illumination changes although only two additional synthetic illuminations have been added in the training stage. If the illumination is modeled with more significant samples, the results presented in Fig 3 will probably improve significantly.

Another remarkable aspect is that the recognition rate of P^2CA is more stable through the entire feature dimension axis, whereas PLDA's accuracy decreases as more eigenvectors are added. This is again related with the fact that more data is necessary to compute the within scatter matrix. Otherwise, the eigenvectors computed from the smallest eigenvalues of the face space can only model some noise which does not correspond to the illumination changes and that is useless for the recognition.
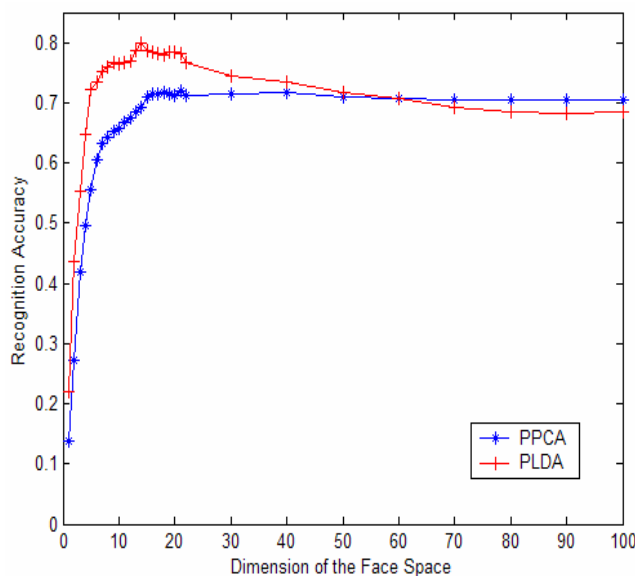

Fig.3 Recognition Rate Vs Number of Eigenvectors

4. CONCLUSIONS AND FUTURE WORK

In this paper, a two dimensional extension of the Fisherfaces presented in [9] have been proposed to develop the Partial LDA approach. PLDA has shown an improvement with respect to P^2CA [1] for illumination variations. However, PLDA cannot cope completely with illumination changes, since there are still several false recognition matches. Thus, more experiments should be performed using more training samples with more lighting variations. This will help to model a more consistent within scatter matrix. Additionally, P^2CA and PLDA face recognition techniques are being extended in order to integrate texture and depth data provided by a multicamera 3D reconstruction approach [2].

5. REFERENCES

[1] A.Rama, and F.Tarrés, "*P^2CA: A new face recognition scheme combining 2D and 3D information*", in IEEE ICIP, Genoa, Italy, September 11th-14th 2005

[2] D. Onofrio, S. Tubaro, A. Rama, F. Tarres, "3D Face Reconstruction with a four camera acquisition system". Proc. International Workshop on Very Low Bitrate Video Coding (VLBV05), Costa Rei, Sardinia, 15-16 September 2005

[3] X. Lu, and A.K. Jain, "Integrating Range and Texture Information for 3D Face Recognition", in Proc. IEEE WACV, Breckenridge, Colorado 2005

[4] K. I. Chang, K. W. Bowyer, and P. J. Flynn, "*An Evaluation of Multimodal 2D+3D Face Biometrics*", in IEEE Trans. on PAMI, Vol.27, pp 619-624, April 2005

[5] V. Blanz, and T. Vetter, "Face Recognition based on fitting 3D morphable model", in IEEE Trans. Pattern Analysis and Machine Intelligence, 25(9):1063-1074, 2003

[6] A. Ansari, M. Abdel-Mottaleb, "3D Face Modelling Using Two Ortogonal Views and a Generic Face Model", in Proc. of Int. Conf. and Multimedia and Expo, July 2003

[7] Y. Wang, C. Chua, and Y. Ho, "Facial Feature Detection and Face Recognition from 2D and 3D images", Pattern Recognition Letters, Vol 23, pp 1191-1202, 2002

[8] J. Yang, D. Zhang, A.F. Frangi, and J.Yang, "*Two-Dimensional PCA: A New Approach to Appearance-based Face Representation and Recognition*", PAMI, Jan. 2004

[9] P. N. Belhumeur, J.P. Hespanha, and D. J. Kriegman, „Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", in IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, Nº7, July 1987

[10] "UPC Face Database" in http://gps-tsc.upc.es\GTAV\ResearchAreas\GTAVDatabase.htm