Face Recognition Using Principal Component Analysis

Javier de Alfonso Miñambres ‡, A.M. Tomé

***Abstract* – This paper discusses the application of Principal Component Analysis (PCA) in face recognition systems. PCA subspace models are commonly used to perform image dimension reduction before the input of the classifier. More recently, PCA subspace models are estimated for one face and the comparison of models via a subspace distance allows face identification. Both strategies of applying PCA were compared for a repository of faces of famous people in uncontrolled poses.**

***Resumo* – Este trabalho apresenta e discute a aplicação da decomposição em componentes principais (PCA) em sistemas de reconhecimento. PCA começou por ser utilizado para reduzir a dimensão das imagens, como um bloco de pre-processamento à entrada do classificador. Mais recentemente, o modelo PCA é utilizado como modelo de uma face e a comparação entre os modelos, via uma medida de subespaço, é a base do processo de decisão. As duas estratégias na utilização do PCA são estudadas utilizando um conjunto de imagens de pessoas conhecidas em poses não-controladas.**

***Keywords* – Principal Component Analysis (PCA), eigenfaces, nearest-neighbor, subspace distance**

***Palavras chave* – Decomposição em componentes principais (PCA), eigenfaces, vizinho-mais-próximo.**

I. INTRODUCTION

Face recognition is one of the most studied problems in the computer vision field [1], [2], so many different solutions to perform face recognition are proposed [3], [4]. Principal Component Analysis (PCA) is a very simple and effective way to perform dimension reduction of raw data in face recognition systems. In the early face recognition systems that use PCA, each image is projected onto a subspace model and this new representation constitutes the input for the decision block. More recently, PCA re-gain interest for video applications where the subspace models can be computed to represent one person. Therefore, the decision block has different subspace models to be compared as input.

In this work these two strategies of using PCA in face recognition systems are presented. The first strategy is based on a single PCA block for the database and the second one uses a PCA block for each person in the database. The decision blocks in both strategies are nearest neighbor classifiers but using different metrics: the first uses euclidian distance to calculate distances between projections while the second uses principal angle based distance to compare individual PCA subspace models. Preliminary results concerning the recognition rate in a public available repository of faces of well-known people are shown.

‡-ERASMUS in University of Aveiro during 2009/2010 originally from University Carlos III of Madrid

II. FACE RECOGNITION AND PCA MODELS

In face recognition systems, the subspace model is usually computed during the training phase and stored to be used during the test phase. In this section we present, first, how to compute this model and its main properties, and, afterwards, a description of the integration of the model in face recognition systems.

A. PCA Subspace Model

The roots of subspace PCA models can be found in singular value decomposition (SVD). Forming an $M \times N$ matrix $\mathbf{X}$, with one image per column, obtained by the row or column concatenation, SVD allows to explain the data set $\mathbf{X}$ as a product of matrices

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{1}$$

and, assuming that $M > N$ (number of pixels larger than size of the data set), these matrices are

- an $M \times N$ matrix $\mathbf{U}$ with orthogonal columns, $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. Matrix $\mathbf{U}$ constitutes the matrix of basis vectors, e.g., the subspace model (usually called eigenfaces), in the space of dimension M.
- an $N \times N$ diagonal matrix with the singular values ordered by decreasing order $\Sigma = diag(\sigma_1, \sigma_2, \dots, \sigma_N)$.
- an $N \times N$ orthogonal matrix $\mathbf{V}$, e.g, $\mathbf{V}^T\mathbf{V} = \mathbf{V}\mathbf{V}^T = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix.

Usually PCA subspace model $\mathbf{U}$ can be achieved by computing the eigendecomposition of the matrix $\mathbf{S} = \mathbf{X}\mathbf{X}^T$. However, in face recognition applications the dimension $M \times M$ of $\mathbf{S}$ is too large to turn viable the application of an eigendecomposition algorithm. Therefore, the eigendecomposition of matrix

$$\mathbf{K} = \mathbf{X}^T\mathbf{X} = \mathbf{V}\mathbf{\Sigma}^T\mathbf{\Sigma}\mathbf{V}^T = \mathbf{V}\mathbf{D}\mathbf{V}^T \tag{2}$$

is performed, where the diagonal matrix $\mathbf{D} = diag(d_1, d_2 \dots d_N)$ is the eigenvalue matrix. The eigenvalues are the square of the singular values. By multiplying both sides of eqn. 1 by $\mathbf{V}\mathbf{D}^{-1/2}$, the so called dual form of the subspace model is obtained

$$\mathbf{U} = \mathbf{X}\mathbf{V}\mathbf{D}^{-1/2} \tag{3}$$

Thus, the maximum number of columns of matrix $\mathbf{U}$ is N, which corresponds to the maximum possible number of non-zero singular values of the data set $\mathbf{X}$.

A.1 Model Order Selection

By selecting $L < N$ eigenvalues and their corresponding eigenvectors, the subspace model $\mathbf{U}_L$ will have L basis vectors. The most widely used criteria to choose L is to take the eigenvalues in decreasing order and compute the normalized cumulative sum of eigenvalues, which reads

$$\frac{d_1+d_2+d_3+\ldots+d_L}{d_1+d_2+d_3+\ldots+d_N} \geq \theta, \quad d_1 \geq d_2 \geq \ldots \geq d_N \qquad (4)$$

Then, the order (L) of the model is chosen to match an user defined threshold (θ). The subspace model $\mathbf{U}_L$ will then have L columns corresponding to the L selected eigenvalues.

A.2 Pre-processing

Usually the data set $\mathbf{X}$ is assumed to be centered in the space of dimension M. Therefore the mean image (see fig. 1) can be subtracted to all the images of the data set before computing de subspace model. However, in digital image processing applications the number of gray levels can be an alternative to this traditional centering of the data. In this case, in images with 8 bits per pixel, the value x of each pixel is updated by

$$x = \frac{x - 128}{128} \qquad (5)$$

This way, pixel values are normalized in the range $[-1, 1]$.

Fig. 1 - Average face of the data set

The pre-processing step has influence in the profile of the eigenvalues, as it can be verified with the normalized cumulative sum of eigenvalues (see eqn. 4). Fig 2 shows that the profile after centering the data, normalizing pixel values or simply using the raw data. The first two approaches have a similar profile while with the raw data the first eigenvalue accumulates more than 95% of the information of the data set. The advantage of using eqn. 5 to transform the value of the pixels is that the operation can be applied without any side information. If centering is applied the average image of the training set (figure 1) should be stored to do the pre-processing of every new image projected onto the subspace model. However, all the numerical simulations presented here will be discussed using those three strategies of performing the pre-processing.

B. Face Recognition

In this work, PCA model was used as pre-processing block to perform face recognition using two strategies. The first is called *single PCA* model [1] and the decision block is a

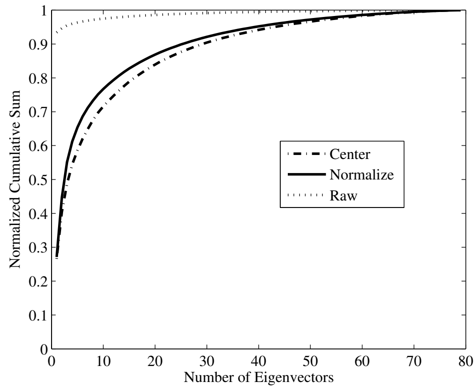

Fig. 2 - Normalized cumulative sum of eigenvalues depending on the pre-processing step.

nearest neighbor classifier . The other alternative is called *multiple PCA* model and face recognition is achieved by using a subspace distance based on principal angles [5].

B.1 Single PCA model

The subspace model $\mathbf{U}_L$ is computed for the labeled training set $\mathbf{X}_{train}$ which should include images of all persons. Fig. 3 shows the $L = 22$ first eigenfaces of the data set. This PCA model is spanned using all the images available except for the image set which will be used as test group. This generates a facespace where all images are represented by a combination of eigenfaces.

This new representation of the training set is formed

Fig. 3 - Eigenfaces of the database

by projecting the training set onto the basis vectors $\mathbf{Z} = \mathbf{U}_L^T \mathbf{X}_{train}$. Any new image $\mathbf{y}$ also can be projected onto the subspace model $\mathbf{z}_y = \mathbf{U}_L^T \mathbf{y}$ and the following euclidian distance [6] is computed

$$d_k = \|\mathbf{z}_y - \mathbf{z}_k\| \quad k = 1, 2, \ldots$$

for all the elements k of the training set. The label of the closest element in the training set, e.g., $min(d_k)$ is used to identify $\mathbf{y}$.

B.2 Multiple PCA model

In this case, PCA models are generated for each person of the data set using only their images. Several subspace models are obtained $\mathbf{U}_{L_1}, \mathbf{U}_{L_2}, \ldots \mathbf{U}_{L_k}$. Each model has order L_k and is related with person k of the database. Figure 4 illustrates the subspace model of one person of the database.

To find out the identity of a person u, several test images should be available and its subspace model $\mathbf{U}_{L_u}$ computed. Following, a decision is taken using a distance based on

Fig. 4 - Eigenfaces of one person of the database.

principal angles [7], [8], [9]

$$dist_{(k,u)} = \|\mathbf{U}_{L_u}\mathbf{U}_{L_u}^T - \mathbf{U}_{L_k}\mathbf{U}_{L_k}^T\|_F = \sqrt{L_u + L_k + 2\,trace(\mathbf{U}_{L_u}^T\mathbf{U}_{L_k})} \quad (6)$$

where $\|.\|_F$ is the Frobenius norm. Then, the person u will have the label of the person k which has $min(dist_{(k,u)})$.

III. NUMERICAL SIMULATIONS

Both strategies of using PCA model will be presented and discussed with a data set of famous people photographed under uncontrolled lighting, position and facial expression circumstances [10], [11]. Furthermore, both strategies are compared using the recognition rate achieved at the output of the decision blocks following cross-validation techniques [12].

A. Data Set

The data set used is an extract of the LFWcrop Face Data set (Labeled Faces in the Wild cropped) [10]. It contains grey-scale images of size 64×64 pixels in pgm format. The origin of the images makes them very uneven in terms of facial expression of the subject, which results in a database with pictures of a same person smiling and showing the teeth, with and without a moustache, or occasionally wearing sunglasses. Some subjects have their face partially covered by hair, and others have the head pretty turned around. From the database, 80 images of 10 persons (8 images for each) were selected. Each person has been assigned a label $k = 1, 2, \ldots, 10$ and all the images are tagged with each person's labels.

To overcome a problem of a small sample set, synthetic images can be generated from the original ones by shifting them [7]. The procedure is fairly simple and doesn't add computational or storage costs since the images generated are simply shifted versions of the original one. Images $\mathbf{\Phi}$ are generated from an image represented by matrix $\mathbf{\Psi}$ of dimension $P \times Q$, in our case 64×64, getting pq images of dimension $S \times R$ as follows:

$$\mathbf{\Phi}_{i,j} = \mathbf{\Psi}(i : (S + i - 1), j : (R + j - 1)) \quad (7)$$

With $1 \le i \le p$ and $1 \le j \le q$, where p and q are parameters that choose the amount of images to synthesize (pq), $S = P - p + 1$ and $R = N - q + 1$. Figure 5 illustrates the result of creating four images for each image of one person.

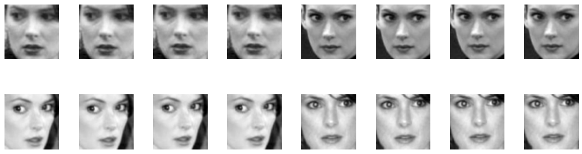

Fig. 5 - Extended database: 1 original image gives 4 new images ($p = q = 2$)

A.1 Single PCA Model

The tests were conducted with the raw data set, using $8-$fold (with an image of each person in the test set) and leave-one-out cross-validation strategies to measure the performance rate at the output of the nearest neighbor classifier. And as expected the results weren't what literature usually shows for data set with faces under controlled circumstances. The recognition rate was close to 40% if the subspace model was computed with $\theta = 1$. However, computing the subspace model for the available data set, and testing the recognition rate with the images included in the training set the performance rate is higher than 70%. Fig. 6 shows the results varying the threshold $\theta \times 100$ to compute the subspace model after the different strategies to pre-process the training data set. Performance is not significantly dependent on the model order, meaning that the retrieval of the data base is possible with a small number of eigenvectors. This result also suggests that if the test image is close to the ones of the training set, the performance would naturally increase.

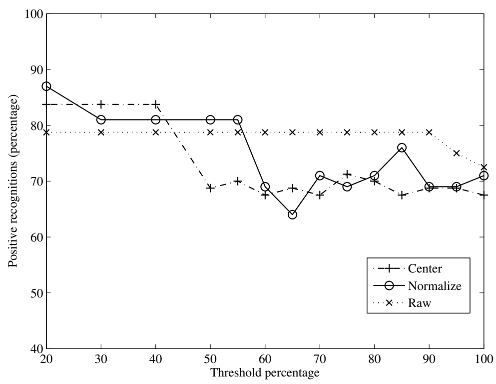

Fig. 6 - Performance (Leave-One-Out test) with the test image in the training set.

B. Multiple PCA Models

The subspace models were computed for each person of the data base. To apply the decision based on subspace angles (see 6), half of the images of each person are randomly chosen for the training phase and the remaining are left for the test phase. The first test is to train the model as it is, with four images for test and four for training. The results were very poor (as expected, see fig. 7), so the database should be extended somehow. Extended versions of each image were considered, generated by shifting the original image to create new ones [13]. The training and testing are performed in two manners: firstly, we will generate the sub-

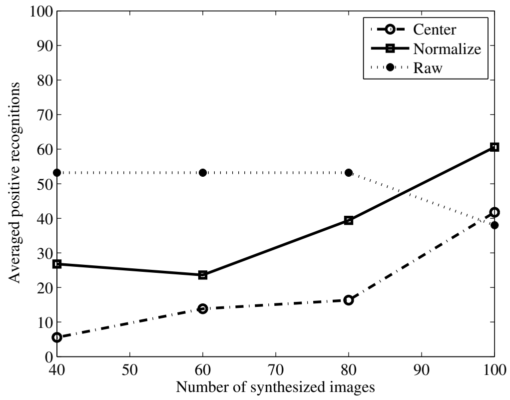

Fig. 7 - Evolution graph for Single Person Subspace averaged after 50 tests

space model by shifting just one image of that person. We will call this *Single Image Extended Database*, or SIED. Secondly, we will shift the available four images of each person to generate the subspace models, obtaining what we call a *Multiple Image Extended Database* or MIED.
The SIED strategy gives recognition rates close to 60% with a model order $L_k = 4$ having the normalized database as input. This could be due to only having one image to compare. Since the database is very heterogenic, images of a person can be very different since they were obtained under uncontrolled circumstances, and that may hinder recognition. It has to be noticed that the order of the subspace models is very low in this strategy, in some cases $L_k = 1$ almost contain 85% of the information (see fig. 8). With the Multiple Image Extended Database (MIED) strategy, the same level of information is achieved with a higher order model (fig. 5). And this is naturally related with the number of original images used to compute the models. The recognition rates are in this case closer to the ones that are usually discussed in literature for face recognition systems (fig. 9). The figure also shows that the normalization step can substitute the centering step and that way we can avoid storing mean image. The figure shows the performance obtained by creating 16 images given the 4 original images of the data set. However enlarging the database with more images does not change the success rate as it can be verified with the results of table I. Generating more images out of the original doesn't mean that the profile of the cumulative sum of eigenvalues changes substantially. Therefore the information held by the synthetic data set do not improve increasing the number of images.

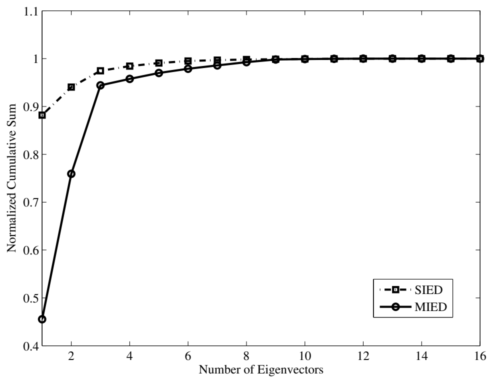

Fig. 8 - Order of the subspace model using SIED and MIED strategies to compute subspace models.

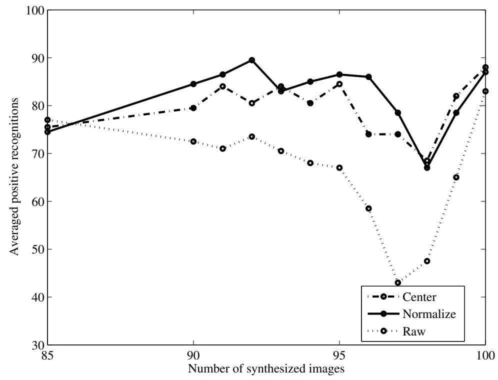

Fig. 9 - Recognition rate using multiple PCA models with four original images in the training and test sets to generate a synthetic database.

TABLE I
PERFORMANCE AVERAGED THROUGHOUT DIFFERENT MODEL ORDERS (θ) FOR DIFFERENT SIZES OF THE SYNTHETIC DATABASE CREATED WITH FOUR ORIGINAL IMAGES.

Number Synthetic	Center	Normalize	Raw
16 ($p = 2, q = 2$)	79.6	82.2	66.4
36 ($p = 3, q = 3$)	78.9	84.2	70.6
64 ($p = 4, q = 4$)	79.3	82.4	69.3

IV. CONCLUDING REMARKS

In this work two different approaches to include PCA models in face recognition systems were presented. Having seen that the database used, LFWcrop Image database, is a set of images taken in *the wild*, the results obtained are according to our expectations.
Aside from the kind of experiment performed, we have seen that the preprocessing methods help improve PCA's performance. The preferred option regarding the results obtained and considering its computational benefits would be normalizing the images. In all of the tests carried out, using the normalized database showed that the performance is pretty much the same as with the centered database.This fact, added to not needing to calculate the mean face and subtract it to all of the images (with the inconvenience that if a new image is added to the database, the centering has to be done again), and store it and use it as an input parameter for PCA, we can conclude that normalizing is the best option indeed.
Studies and papers published often use databases which have been created under controlled circumstances. The preliminary results discussed in this work show that PCA subspace models are reliable for face recognition even in non-controlled environments, although the main goal for its application has to be different from the dimension reduction goal of the pioneer systems.

The results also show that it is far more efficient to extend the databases by shifting a few images from all the available (MIED) than generating that same amount of images from just one (SIED). However, it was also concluded that increasing the number of synthetic images do not mean that the performance is always improving. That means that it is preferable to generate synthetic images from different than to generate synthetic images from a single one. Nonetheless, even with the MIED strategy, only increasing the number of images does not grant an increasing performance. This is in an important aspect to be taken into account if the PCA models are generated from video sequences.

REFERENCES

[1] Matthew Turk and Alex Pentland, "Face Recognition Using Eigenfaces", in *Proceedings of the IEEE Computer Vision and Pattern Recognition*, Maui, Hawaii, USA, 3-6 June 1991, pp. 586–591.

[2] Matthew Turk and Alex Pentland, "Eigenfaces for Recognition", *Journal of Cognitive Neurosicence*, vol. Volume 3, no. No.1, pp. pages 71–86, 1991.

[3] Benjamin Weyrauch, J. Huang, B. Heisele, and V. Blanz, "Component-based face recognition with 3d morphable models", in *IEEE Workshop on Face Processing in Video*, 2003, pp. 1–5.

[4] L. Wiskott, J.-M. Fellous, N. Krueuger, and C. von der Malsburg, "Face Recognition by Elastic Bunch Graph Matching, Chapter 11 in Intelligent Biometric Techniques in Fingerprint and Face Recognition", 1999.

[5] Xichen Sun, Liwei Wang, and Jufu Feng, "Further results on the Subspace Distance", *Pattern Recognition*, vol. 40, pp. 328–329, 2007.

[6] T Cover and P Hart, "Nearest Neighbor Pattern Classification", *IEEE Transactions on Information Theory*, vol. Volume 13, no. Issue 1, pp. Pages 21 – 27, January 1967.

[7] Jun Liu, Songcan Chen, Zhi-Hua Zhou, and Xiaoyang Tan, "Single image subspace for face recognition", in *Proceedings of the 3rd international conference on Analysis and modeling of faces and gestures*, AMFG'07, pp. 205–219. Springer-Verlag, Berlin, Heidelberg, 2007.

[8] Aurél Galántai, *Projections and Projection Methods*, 2004.

[9] J Ross Beveridge, Bruce A. Draper, Jen-Mei Chang, Michael Kirby, Holger Kley, and Chris Patterson, "Principal Angles Separate Subjects Illumination Spaces in YDB and CMU-PIE", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. volume 31, no. Issue 2, pp. 351–356, February 2009.

[10] "LFWcrop Face Database", `http://itee.uq.edu.au/~conrad/lfwcrop/`.

[11] Conrad Sanderson and Brian Lovell, "Multi-region probabilistic histograms for robust and scalable identity inference", in *Advances in Biometrics*, Massimo Tistarelli and Mark Nixon, Eds., vol. 5558 of *Lecture Notes in Computer Science*, pp. 199–208. Springer Berlin / Heidelberg, 2009.

[12] Ian H. Witten and Eibe Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, Elsevier Science Ltd, second edition, June 2005.

[13] Jun Liu, Songcan Chen, Zhi-Hua Zhou, and Xiaoyang Tan, "Single image subspace for face recognition".