IEEE ISIE 2005, June 20-23, 2005, Dubrovnik, Croatia

Detecting Faces and Recognizing Facial Features Using Color Segmentation and 2DPCA in the Normalized RG Space

M. S. Escudero, L. M. Bergasa, E. López, R. Barea*, A. Delicado** , J. Nuevo*
* Electronics Department. University of Alcala. Campus Universitario. Ctra. Madrid-Barcelona, Km. 33,600, Spain
** Grupo Tecnológico IT-Deusto. Avda. Llano Castellano 13-5. Madrid, Spain
marisol@depeca.uah.es, bergasa@depeca.uah.es, elena@depeca.uah.es, barea@depeca.uah.es,
adelicado@itdeusto.com, nuevo@depeca.uah.es

***Abstract*— Face detection and recognition is very challenging due to the diverse variation of face appearance, facial expressions, variable recording conditions (changes in illumination, scale differences, varying face position…) and the complexity of image background. In this paper, we propose a new system which integrate color segmentation and Two Dimensional Principal Component Analysis in the normalized RG space (2DPCA, to compress the red information as well as the green information) in order to detect faces and recognize facial features in color images that have not been preprocessed.**

We show some experimental results, using our own face database and the AR and PICS face databases. Then, we have compared results obtained with 2DPCA technique in the normalized RG space and other typical methods (2DPCA in the gray level space, PCA, Fisherfaces, Kernel PCA and Kernel Fisherfaces). Conclusions and future works have finally presented.

I. INTRODUCTION

Face images have received considerable attention, which has increased exponentially in the last decade.

This interest is motivated by the broad range of potential applications for systems able to code, interpret and recognize faces images. Examples include:

- personal identification and access control [1-7]
- systems for locating faces [8-10]
- systems for locating facial features [11-13]
- systems for recognizing facial expressions [14-21]

Nowdays, it is possible to work with a lot of databases and videos (or sequences) of faces, where everybody can probe and compare the implement algorithms with other algorithms of other authors or colleagues.

In this way, we can mark the following databases of gray level face images:

- Cohn-Kanade database
- FERET database
- Oliveti (ORL) database
- MIT database
- JAFFE database
- Yale database
- UMIST database

And the following databases of color level face images:

- OULU database
- M2VTS database
- CMU PIE database
- AR database [22]
- PICS database [23]

In this article, to build a robust face detection system, we work with the AR and PICS databases and with other images obtained by the authors (with face images which have turns in the image plane), in order to make a set of complex and highly variable face images. Sources of variability include pose, facial expressions, individual appearance, and lighting.

Respecting the recognition method used, PCA [24] has been widely investigated and has become one of the most successful approaches in face recognition. Recently, a new method called 2DPCA has been developed [25]. This last method has two important advantages over PCA. Firstly, it is easier to evaluate the covariance matrix accurately. Secondly, less time is required to determine the corresponding eigenvectors. In other hands, effectiveness and robustness of 2DPCA is higher to other techniques as *Fisherfaces*, *ICA* (Independent Component Analysis) and their corresponding non-linear techniques, *Kernel PCA and Kernel Fisherfaces*, as it has been shown in [25] using three face images database: ORL, AR and Yale. This is the reason because we have applied the 2DPCA methodology in our recognition system.

In this paper we evaluate performance of 2DPCA method in the normalized RG space and we compare its results with the obtained with other traditional methods, in order to show its advantages over the reference methods.

The organization of the paper is as follows:

- Section two describes the proposed face detection system
- Section three explains the abilities that the system provides for personal identification and recognizing appearance (wearing sun glasses, wearing scarf, the head is turn around…).
- Section four shows results.
- Section five draws some conclusions.

II. Face Segmentation

The general process of face segmentation, that will be described in detail immediately after, is based on an initial clustering, in the normalized RG space [26], that classifies the image depending on their primary colors. After a more accurate adjustment of this initial clustering, it is chosen which cluster (or combination of clusters) belongs to the face category, comparing the Euclidean distances of the clusters (and combination of clusters) with cluster face patterns obtained previously. The closest category to the pattern is considered as a face.

As difference with almost all the existing face and facial features recognition system who works in gray level images, in our case, the recognition is carried out after projecting the clusters candidates for being faces and the pattern prototypes to a space of less dimensions using 2DPCA in order to compress the red information and the green information. This new method which works in color images is called 2DPCA in the normalized RG space.

A. Clustering in color by competitive learning

Firstly, a clustering algorithm has been used to detect the face [27]. This algorithm, starting from the normalize RG color, calculates the number of different colors in an image by using a VQ (Vector Quantization) neuronal network that trains itself with competitive learning [28]. A histogram of the low-resolution image is used to initiate the network. The right number of categories of the clustering is calculated automatically evaluating the adaptation of the categories to the topology of the histogram in order to increase higher the variance among categories than the inner variance of each category.

Some experimental results of the general process of clustering can be observed in Fig. 1.

Fig. 1 Clustering based on color information in scenes without deep changes in illumination

B. Formation of the candidate clusters for being faces

Once the image is divided in clusters (from the color information), we have to do a more accurate adjustment of the clusters before trying to select which cluster can be considered as a face blob. In order to do it:

1.- A harder research of every generated cluster is done to see if it is convenient to divide the cluster. This process is necessary to improve the clustering because, even though the image is classified in primary colors in the beginning, sometimes there are mistakes with images in dim light. However, it is important to highlight that this embarrassing process is not very common; it is just done where the clustering makes an excessive grouping of colors, dividing the image in just two colors (see Fig. 2).

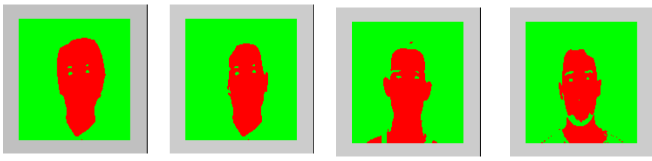

Fig. 2 Clustering in two categories

2.- The algorithm takes from every cluster (or combination of two clusters) the biggest blob, taking away the holes of it (so the eyes or mouth would not be eliminated from the head). Afterwards, the blob is turned around (using PCA with the x-y coordinates of every pixel of the blob) so the system can work properly with turns in the image plane (see Fig. 3).

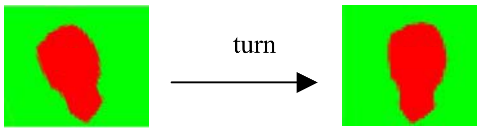

Fig. 3 Turning the cluster

To conclude, that blob is introduced in “the smallest possible box”, and it has to be normalized the size of that cluster_image_box (created by pixels which belong to the “cluster in color”, and the rest in black) with a certain height and length, previously calculated, of 128x128.

The creation of the cluster_image_box from every color cluster is shown in figure 4 for a user with sunglasses as well as for a user without sunglasses.

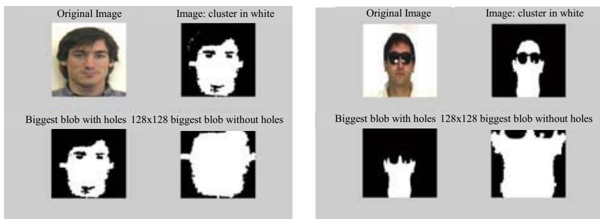

Fig. 4 Formation of a cluster_image_box from a cluster

That way, it is obtained as many cluster_image_box as clusters (and combinations of two clusters) have been obtained from the last step.

Letting combinations of clusters is due to face images which have got a very illuminated area and another dark area (because of, for instance, a spotlight which light up the left side of the face), so it would only be obtained a half of the face if combinations of clusters were not used (see Fig. 5).

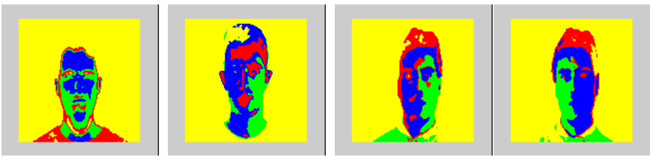

Fig. 5 Clustering in scenes with deep changes in illumination

To solve this typical problem of changes in illumination, cluster_image_box formed by images from the combination of two clusters were also created. Therefore, after step B.1 the image is divided in K clusters (with 2<K<10), there will be a number of combinations of:

$$n = \frac{[1+(K-1)]}{2}(K-1) = \frac{K}{2}(K-1) \tag{1}$$

3.- The whole cluster_image_box candidates for being faces are created after eliminating those clusters or combination of clusters in which the shape of the blob does not resemble a face. Therefore, it does not take into consideration the outlines that are not approximately square, rectangular, round or elliptic figures (typical outlines of faces).

C. Selection of the face cluster

To select the face cluster_image_box, several face pattern blobs of 128x128 pixels are previously obtained, in order to form the training set, following the previous steps, and choosing manually the face pattern cluster between the candidate clusters for being faces.

To reduce the memory requirements, every image pattern and every candidate cluster_image_box for being face are compress to a vector of *2d* components ($1 \le d \le 40$), the first *d* of these components are the result of using 2DPCA to the green component of the image of 128x128 pixels; the other *d* components are the result of using 2DPCA to the red component of the image of 128x128 pixels.

The 2DPCA algorithm [25], applied first to the red component and then to the green component of the normalized RG image, is, with more detail, as follows:

Let **X** denote an n-dimensional unitary column vector. The idea is to project image **A**, an m x n random matrix, onto **X** by the following linear transformation:

$$Y = A\,X \tag{2}$$

Thus, an m-dimensional projected vector **Y** is obtained. This is called the projected feature vector of image **A**. How is determined a good projection vector **X**? Let $\mathbf{A_j}$ (j=1,2,...,M) denote the set of sample images (*m x n*) of a group of users. The total scatter of the projected sample images is a good measure of the discriminatory power of the projected vectors and can be evaluated by the trace of the covariance matrix ($\mathbf{S_x}$) of the feature vectors. The projection vectors are optimal when the total scatter of the projected samples is maximized.

$$J(X) = tr(S_x) = X^T G_t X \tag{3}$$

The optimal projection vectors are the set of eigenvectors of $\mathbf{G_t}$, the *image covariance (scatter) matrix* of the sample images. $\mathbf{G_t}$ is an *n x n* nonnegative definite matrix which can be evaluated by:

$$G_t = \frac{1}{M}\sum_{j=1}^{M}\left(A_j - \overline{A}\right)^T\left(A_j - \overline{A}\right) \tag{4}$$

where $\overline{A}$ is the average image of the training samples:

$$\overline{A} = \frac{1}{M}\sum_{j=1}^{M} A_j \tag{5}$$

In general, it is enough to select the first *d* orthonormal eigenvectors corresponding to the first d largest eigenvalues $\{\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_d}\}$, in order to characterize the whole image.

The optimal projection vectors of 2DPCA, $\{\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_d}\}$, are use for feature extraction. For a given image sample **A**, let

$$Y_k = A\,X_k,\quad k = 1,2,...,d \tag{6}$$

Then, a family of projected feature vectors, $\{\mathbf{Y_1}, \mathbf{Y_2}, ..., \mathbf{Y_d}\}$ are obtained, which are called the principal components (vectors) of the sample image **A**. The principal component vectors obtained are used to form an *m x d* matrix $\mathbf{B} = [\mathbf{Y_1}, \mathbf{Y_2}, ..., \mathbf{Y_d}]$, which is called the *feature matrix or feature image* of the image sample **A**.

After a transformation, a feature matrix is obtained for each image. Then, two arbitrary feature images can be compared using a nearest neighbor classifier.

The distance between the matrices:

$B_i = \left[Y_1^{(i)}, Y_2^{(i)}, ..., Y_d^{(i)}\right]$ and $B_j = \left[Y_1^{(j)}, Y_2^{(j)}, ..., Y_d^{(j)}\right]$, is defined by:

$$d\left(B_i, B_j\right) = \sum_{k=1}^{d}\left\|Y_k^{(i)} - Y_k^{(j)}\right\| \tag{7}$$

where $\left\|Y_k^{(i)} - Y_k^{(j)}\right\|$ represent the Euclidean distance between the two principal component vectors $Y_k^{(i)}$ and $Y_k^{(j)}$.

The feature matrix of the training images are grouped by users, then, a user, $\mathbf{U^z}$, will have associate the features matrix corresponding to his training images, $U^z = \left[B_1^z, B_2^z, ..., B_M^z\right]$. Given a test feature matrix (of the probe set), called **B?**, the face cluster_image_box will be that whose feature image B_j^z has the closest distance d to **B?**

$$if\ d\left(B?,B_l^w\right)=\min_{\forall j,\forall z} d\left(B?,B_j^z\right) \quad then \quad B?\in U^w \qquad (8)$$

In short, the cluster_image_box whose Euclidean distance to any pattern is a minimum is considered as a face (the distance algorithm gets better results with k=1 [29]).

III. RECOGNISING USERS AND FACIAL FEATURES

The patterns that belong to a face are classified taking into account the following facial features:

- Frontal face
  - standard
  - with sunglasses
  - with a scarf or other element in the lower part of the face
  - with a hat or other element in the upper part of the face
- Face turn 90º to the left side
  - standard
  - with sunglasses
  - with a scarf or other element in the lower part of the face
  - with a hat or other element in the upper part of the face
- Face turn 90º to the right side
  - standard
  - with sunglasses
  - with a scarf or other element in the lower part of the face
  - with a hat or other element in the upper part of the face

Consequently, the distance to the closest pattern decides the facial features of the user (e.g. frontal face with sunglasses). Furthermore, the identified user can be forced to go with the user-pattern (Fig. 6).

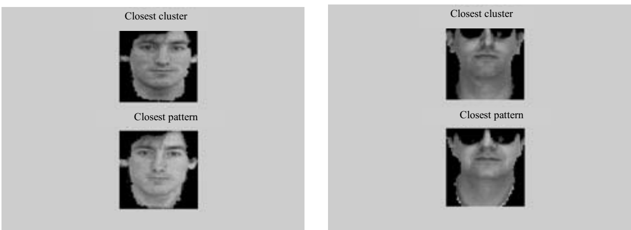

Fig. 6 Recognition taking into account the distance to the closest neighbour

IV. EXPERIMENTAL RESULTS

The goal of this paper is evaluate performance of the 2DPCA technique in the normalized RG space in a facial features recognition system and compare its results with the obtained with the traditional PCA technique.

The results have been obtained varying the number of principal components vectors from 1 to 40 for both methods (PCA, 2DPCA in the normalized RG space), and the number of training samples from 5 to 15 as can be seen in figure 7. The top recognition accuracy (98,66%) is achieved by 2DPCA in the normalized RG space with 5 double eigenvectors (5 eigenvectors in the red space and 5 eigenvectors in the green space) and 15 samples. For the PCA method the maximum recognition accuracy is 97,77% for 20 principal components and 15 samples. As can be observed for the 2DPCA in the normalized RG space, the best accuracy results are obtained for a low number of components vectors while for the PCA technique is necessary a higher number of principal components in order to obtain similar accuracy values. For the both methods, the more samples training used, the better performance is achieved.

The process time (recognition time) and the size of necessary stored data have been too analyzed in order to compare 2DPCA in the normalized RG space and PCA techniques. For PCA it is necessary to store the eigenvectors matrix, the average matrix and the principal components for every training sample. For 2DPCA in the normalized RG space, eigenvectors matrix and feature matrix of the red and green component per training sample are needed. Table 1 shows recognition time and size of stored data for PCA and “2DPCA in the normalized RG space” methods in the case of maximum recognition accuracy. This situation is reached in the case of 2DPCA in the normalized RG space for 15 training samples and 5 double eigenvectors ($2d$ = 2·5 = 10 eigenvectors) and in the case of PCA for 20 eigenvectors. As can be seen, the training (feature extraction) and recognition time are significantly better for 2DPCA in the normalized RG space than for PCA. On the other hand, the size of stored data is higher for 2DPCA in the normalized RG space than for PCA.

In order to obtain a more exhaustive study of the 2DPCA performance in the normalized RG space (RG 2DPCA) we have also compared it with other typical methods, very known in the recognition literature, as 2DPCA in the gray level space (gray 2DPCA) [25], Fisherfaces (FF) [30], Kernel PCA (KPCA) [31] and Kernel Fisherfaces (KFF) [32]. The experiment results are listed in Table 1. Respecting the maximum recognition accuracy, 2DPCA in the normalized RG space was better than other methods. Other methods needed a very higher number of eigenvectors in order to obtain a similar recognition accuracy than the 2DPCA in the normalized RG space. Then, the size of the stored data and the process time were higher for the other methods respecting the 2DPCA in the normalized RG space. In conclusion, general performance of the 2DPCA method in the normalized RG space is better than the obtained for the other typical methods which it was compared.

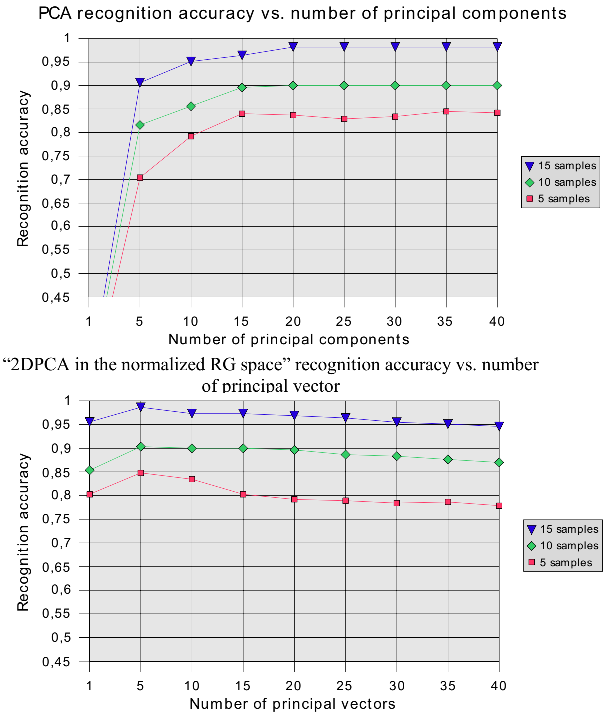

Fig. 7 .Recognition accuracy performance for 2DPCA in the normalized RG space and PCA under the number of principal components and number of samples

TABLE I
COMPARATION BETWEEN TRADITIONAL METHODS AND 2DPCA IN THE NORMALIZED RG SPACE

	Max. Recognition accuracy	**Training time (s)**	**Recogn. time (ms)**	**Size of stored data (KB)**	**Number of eigenvectors**
RG 2DPCA	98,66 % (222/225)	1.54	124.6	979.92	10 = 5x2
gray 2DPCA	98.22 % (221/225)	0.84	95.24	489.96	5
PCA	97.77 % (220/225)	2.33	112.18	120.70	20
FF	98.22 % (221/225)	11.22	125,63	1057.60	30
KPCA	97.77% (220/225)	16.94	583.68	2166.70	20
KFF	96.88%(218/ 225)	6.57	595.52	1995.70	35

V. CONCLUSIONS AND FUTURE WORKS

A new algorithm (2DPCA in the normalized RG space) has been created to detect and recognize face characteristics and users with very good results. Moreover, it can work properly with illumination, rotation in the image plane, background, expression and size changes (but the size of the face in the image has to be big enough to be one of the predominant colours in the image).

We should highlight about this new work:

- The use of color information to divide the image in clusters, letting an aggregation of clusters in order to get a robust system while detecting faces with great changes in illumination.
- The use of Principal Component Analysis to get a robust system in rotations (in the image plane).
- The use the 2DPCA to red and green component of the image in order to support the comparisons of the characteristics in color of the images.

We have analyzed performance of this technique over a set of complex and highly variable faces images which have been obtained from our own face database and the AR and PICS face databases, and we have obtained better results that obtained by others researchers over others well-known face image databases as ORL, AR and Yale. We have compared results obtained with 2DPCA technique in the normalized RG space with the traditional PCA one and we have concluded that the first one is simpler, computationally more efficient and more accurate, for a low number of components, than the second one. Then, we have compared 2DPCA in the normalized RG space and other typical methods (*2DPCA in the gray level space, Fisherfaces, Kernel Fisherfaces and kernel PCA*) performance and we have conclude that the first one is better than the second ones.

In the early future we have the intention of enhance the number of images for training of each user under extreme conditions and occluding structures in indoor environments for improving the performance of the system under extreme conditions.

ACKNOWLEDGMENT

This work has been supported by grants DPI2002-02193 (SIRAPEM Project) and TIC2002-10744-E (Special Action of the ADVOCATE II Project) from the Spanish Ministry of Science and Technology (MCyT).

REFERENCES

[1] F. S. Samaria, A. C. Harter. Parameterisation of a Stochastic Model for Human Face Identification. Proceedings of the 2nd IEEE Workshop on Applications of Computer Vision. Sarasota (Florida). December 1994

[2] A. Lanitis, C. J. Taylor, T.F. Cootes . An automatic face identification system using flexible appearance models. Proc. British Machine Vision Conference, BMVA Press, 1994, pp.66-75. 1994

[3] V. Espinosa y M. Faúndez. Face identification by means of a neural net classifier. 33º IEEE International Carnahan Conference on Security Technolgy. Madrid, 1999 pp.182-186. 1999

[4] T. Tan and H. Yan. Face recognition by fractal transformations. In Proc. IEEE ICASSP, pages 3537--3540, 1999

[5] Ara V. Nefian, Monson H. Hayes III. Maximum Likehood Training of the Embedded HMM for Face Detection and Recognition. Proceedings of the 2000 International Conference on Image Processing (ICIP 2000). Vancouver, BC, Canada, 2000.

[6] Hichem Sahbi, Nozha Boujemaa: Robust Face Recognition Using Dynamic Space Warping. Biometric Authentication, International ECCV 2002 Workshop Copenhagen, Denmark, Proceedings. Pp 121-132. 2002

[7] Stan Z. Li and Juwei Lu. Face Recognition Using the Nearest Feature Line Method. IEEE Transactions on Neural Networks. Vol 10, NO 2, pp 439 - 443. March 1999

[8] Y. Wang, B. Yuan. Fast method for face location and tracking by distributed behaviour-based agents.IEE Proc. Vis. Image Signal Process., Vol 149, No 3, pp 173-178. June 2002

[9] Linlin Huang Akinobu Shimizu Hidefumi Kobatake. Face Detection using a Modified Radial Basis Function Neural Network. Proceedings of the International Conference on Pattern Recognition (ICPR '2002). August, 2002

[10] P. Viola, M. Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. Proceedings IEEE Conf. On Computer Vision and Pattern Recognition. 2001

[11] A. Lanitis, A. Hill, T. F. Cootes , C. J. Locating Facial Features using Genetic Algorithms. Proc. International Conference on Digital Signal Processing, Cyprus, 1995. pp. 520-525. 1995

[12] M. Sadeghi, J. Kittler, K. Messer. Modelling and segmentation of lip area in face images IEE Proc. Vis. Image Signal Process., Vol 149, No 3, , pp 179-184. June 2002

[13] A. Yilmaz and M. Shah. Automatic Feature Detection and Pose Recovery for Faces. Asian Conf. on Computer Vision (ACCV2002), pp. 284-289. January 2002

[14] Y. Tian, T. Kanade, and J. Cohn. Recognizing action units for facial expression analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 2, pp. 97 – 115. February, 2001

[15] T. Moriyama, T. Kanade, J. Cohn , J. Xiao, Z. Ambadar, J. Gao and M. Imanura M. Automatic recognition of eye blinking in spontaneously occurring behavior. Proceedings of the International Conference on Pattern Recognition (ICPR '2002). August, 2002

[16] M. Pantic and L. J. M Rothkrantz. Expert system for automatic analysis of Facial Expression. Image and Vision Computing Journal, vol. 18, no.11, pp. 881-905.August 2000

[17] Kapoor Ashish. Automatic Facial Action Analysis. Masters Thesis in Media Arts & Sciences. MIT Media Lab. June 2002

[18] Tanzeem Khalid Choudhuty. FaceFacts: Study of Facial Features for Understanding Expression. Masters Thesis in Media Arts & Sciences. MIT Media Lab. September 2002

[19] Nuria M Oliver. Towards Perceptual Intelligence: Statistical Modeling of Human Individual and Interactive Behaviors. Thesis in Media Arts & Sciences. MIT Media Lab. April 2000

[20] Ira Cohen. Automatic Facial Expresión Recognition from video sequences using temporal information. Thesis in the University Illinois at Urbana-Champaign, 2000.

[21] J.A. Colmenarez. Facial Análisis from continuos video with application to human-computer interface. Thesis in the University Illinois at Urbana-Champaign, 1999

[22] A. R. Martinez and R. Benavente. The AR face database. Technical Report 24, Computer Vision Center (CVC) Technical Report, Barcelona, Spain, June 1998

[23] http://pics.psych.stir.ac.uk/help.html

[24] M. Turk, A. Pentland, Eigenfaces for recognition. Journal of Cognitive Neuroscience, vol. 3, pp. 71-86. 1991.

[25] Yang, D. Zhang, A. F. Frangi and Jing-yu Yang. "Two-dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, no.1, pp. 131-137, Jan. 2004.

[26] J. Yang, W. Lu, and A. Waibel. Skin-color modeling and adaptation. Technical report, CMU-CS-97-146, May 1997.

[27] L.M. Bergasa. Seguimiento facial, mediante visión artificial, orientado a la ayuda a la movilidad. Tesis doctoral. Universidad de Alcalá. 1999

[28] T. Kohonen. Self-Organizating Maps. Springer-Verlag, 1997

[29] A. Selinger and D. Socolinsky. Appearance-Based Facial Recognition Using Visible and Thermal Imagery: A Comparative Study. Technical Report 02-01, Equinox Corportation, February 2002.

[30] Peter N. Belhumeur, João P. Hespanha and David J. Kriegman "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no.7, Jul. 2004.

[31] B. Schölkopf, A. Smola, K. R. Muller, "Nonlinear Component Analysis as a Kernel Eigenvalue Problem", Neural Comput. 10 (5) (1998) 1299-1319.

[32] M. H. Yang, "Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition using kernel methods", Proceedings of the fifth 1EEE International Conference on Automatic Face and Gesture Recognition (RGR'02), Washington DC, May 2002, pp 215-220.