# Experiments with COIL database using PCA-based Object Recognition Techniques

M.A. Vicente, O.Reinoso, N.García, J.M. Sabater, C. Fernández, L.Jiménez,

*Abstract*—**Principal component analysis (PCA) or eigeinimages has become a succesful tool for face recognition. Recently eigenrepresentantions are being used to implement appearance-based object recognition systems, although the results are highly sensitive to clutter and partial occlusion.**

**The goal of this paper is to rigorously compare two diferent PCA techniques in object recognition tasks. The objects used for the system have been taken from the COIL image database. A new approach to object recognition using PCA is presented, and experimental results show that the combination of a particular PCA technique with a predetection of image borders gives very good results even in the presence of occlusions.**

*Index Terms*— **Object recognition, Principal component analysis, Appearance-based methods**

## I. Introduction

OVER the last years, principal component analysis or eigeinimages has become a succesful tool for face recognition. Sophisticated commercial systems have been developed achieving high recognition rates [4][7]. Nevertheless, this technique is seldom used in the recognition of general three-dimensional objects, although there are some important exceptions [2],[3],[16],[17] but most of them show the weakness of this technique: the high sensibility to clutter and partial occlusion. The appearance of a 3D object in a 2D image depends on its shape, its colour, its pose in the global scene, its reflectance properties and the sensor and illumination characteristics. Murase et al [16][17] proposed an appearance representation to recognize 3D objects from 2D images based in eigenrepresentations. In [2] an appearance-based matching method using eigenimages is used too. Instead of computing the coefficients by a projection of the data onto the eigenimages, they are extracted by a hypothesize and test paradigm using subsets of image points. Competing hypothesis are then subject to a selection procedure based on the Minimum Description Length (MDL) principle. Huang and Camps [3] introduced a new representation using appearance-based parts and relations improving the occlusion problems that other approaches present.

M.A. Vicente, O.Reinoso, N.García, J.M. Sabater, C. Fernández and L.Jiménez, are with the System Engineering and Control area of the Miguel Hernandez University, Elche 03202 Spain (phone: +34 96-665-8620; fax: +34 96-665-8798; e-mail: suni@umh.es).

The goal of this paper is to rigorously compare two diferent PCA techniques in object recognition tasks. The objects used for the system have been taken from the COIL image database [5]. A new approach to object recognition using PCA is presented, and experimental results show that the combination of a particular PCA technique with a predetection of image borders gives very good results even in the presence of occlusions.

The remainder of the paper is as follows: in following section a PCA analysis of the different evaluated methods are presented. Section III shows the characteristics of the image database used in the experiments. In section IV the experiments carried out with these images are detailed. Finally conclusions are summarized in section V.

## II. PCA Analysis

### A. Introduction to subspace methods

An image may be considered as a vector of pixels where the value of each entry in the vector is the grayscale value of the correspondig pixel. For example, a NxN image may be unwrapped and treated as a vector of length $N^2$. The image is said to sit in N-dimensional space, which is considered to be the original space of the image. This space is just one of infinitely many spaces in which the image can be examined Others subspaces are the subspace created by the eigenvectors of the covariance matrix of the training images (Principal Component Analysis)[1], the basis vectors obtained using Linear Discriminant Analysis [8],[9] (also known as Fisher Discriminants) or the subspaces computed by Independent Component Analysis[10]-[12].

The original PCA optimizes variance among the images, while LDA optimizes discrimination characteristics. ICA obtains statistically independent vectors from the images. The basic algorithm for identifying images by projecting them into a subspace is the following: first a subspace is obtained from the training images, and these images are projected into this subspace. Next, each test image is also projected into the subspace, and the training projections are compared with the test projections by a similarity or distance measure.

### *B. PCA1*

Eigenspace is calculated by finding the eigenvectors of the covariance matrix created from the set of training images. The eigenvectors correspondig to non-zero eigenvalues of the covariance matrix represent an orthonormal basis that projects the original images in the N-dimensional space.

Each image is stored as a vector of size N:

$$\vec{\phi}^i = \left[\phi_1{}^i \ldots \phi_N{}^i\right]^T \qquad (1)$$

The images are mean centered by substracting the mean image:

$$\tilde{\phi}^i = \vec{\phi}^i - \vec{m} \qquad (2)$$

where

$$\vec{m} = \frac{1}{P}\sum_{i=1}^{P}\vec{\phi}^i \qquad (3)$$

and P is the number of images.

These images are combined to create a NxP data matrix:

$$\Phi = \left[\tilde{\phi}^1 \mid \tilde{\phi}^2 \mid \ldots \mid \tilde{\phi}^P\right] \qquad (4)$$

The data matrix $\Phi$ is multiplied by its transpose to calculate the covariance matrix:

$$\Omega = \Phi \cdot \Phi^T \qquad (5)$$

This covariance matrix has up to P eigenvectors associated with non-zero eigenvalues, assuming P<N. The eigenvectors are sorted, high to low, according to their associated eigenvalues. The eigenvector associated with the largest eigenvalue is the eigenvector that finds the highest variance in the images. The eigenvector associated with the second largest eigenvalue is the eigenvector that finds the second highest variance in the images.

This trend continues until the smallest eigenvalue is associated with the eigenvector that finds the lowest variance in the images.

The method outlined above can lead to extremely large covariance matrices. For example, in our experiments we use images of size 128x128 combined to create a data matrix of size 16384xP and a covariance matrix of size 16384x16384. This is a problem because calculating the covariance matrix and the eigenvectors/eigenvalues of the covariance is computationally expensive. But there is a solution: a theorem in linear algebra states that the eigenvalues of $\Omega = \Phi \cdot \Phi^T$ are the same that the eigenvalues of $\Omega' = \Phi^T \cdot \Phi$, so the eigenvectors of $\Omega$ are the same as the eigenvectors of $\Omega'$ multiplied by the matrix $\Phi$ and normalized [1]. Using this property, the covariance matrix $\Omega'$ has just PxP components rather than the original with NxN. This method will be called PCA1 in advance.

We resume PCA1 in the following steps:

1. Training:
   a. A set of training images is selected, and each image is mean centered, "(2)".
   b. The data matrix is created, "(4)"
   c. The covariance matrix $\Omega'$ is obtained:
   $$\Omega' = \Phi^T \cdot \Phi \qquad (6)$$
   d. The eigenvalues and eigenvectors of $\Omega'$ are computed:
   $$\Omega' \cdot V' = \vec{\lambda} \cdot V' \qquad (7)$$
   where:
   $$\vec{\lambda} = [\lambda_1 \ldots \lambda_k] \qquad (8)$$
   e. The eigenvectors of $\Omega$ are computed and ordered according to the correspondig eigenvalues from high to low.
   $$\hat{V} = \Phi \cdot V' \qquad (9)$$
   f. The training centered images are projected into the eigenspace to obtain the classes:
   $$\vec{\omega}^i = \hat{V}^T \cdot \tilde{\phi}^i \qquad (10)$$
   There are as many classes as training images, and each class is a vector whose size is the number of non-zero eigenvalues.

2. Test:
   a. The centered test image $\tilde{\varphi}$ is projected into the eigenspace:
   $$\vec{t} = \hat{V}^T \cdot \tilde{\varphi} \qquad (11)$$
   b. The projected test image ( $\vec{t}$ ) is compared to every projected training image (class).

This technique of principal component analysis enables us to create and use a reduced set of variables. A reduced set (the classes obtained from the training images) is much easier to analyze and interpret than the original variables (the training images).

### *C. PCA2*

It is possible to introduce a modification on the aproach described above. This new technique it will be denoted as PCA2.

In the PCA1 method one image is chosen from each object to create the data matrix, and there is just one subspace to represent the object recognition system. In the PCA2 method one subspace is obtained for each object, so there are as many subspaces as objects to recognize. Then, for each subspace the data matrix is created from different images taken from the same object. In our experiments a set of 20 different objects was used, so with the PCA2 approach 20 different subspaces were computed.

### *D. DISTANCE MEASURES*

Once the new images are projected into a subspace, the goal is to find the closer images. Usually there are two ways to determine how alike images are. One is to measure the distance between the images in N-dimensional space. The second way is to measure how similar two images are. When measuring distance, one wishes to minimize distance, so two images that are alike produce a small distance.

When measuring similarity, one wishes to maximize similarity, so that two like images produce a high similarity value. There are many possible similarity and distance measures [14],[15], most used are L1 and L2 norm, covariance or cosine, correlation and Mahalanobis distance. We used the following:

*a) L1 norm*

The L1 norm is also known as the city block norm or the sum norm. It sums up the absolute difference between pixels. The L1 norm of two vectors, $\vec{\omega}$ and $\vec{t}$ is:

$$L1 = \sum_{i=1}^{M} |\omega_i - t_i| \qquad (12)$$

where M is the size of both vectors.

*b) L2 norm*

The L2 norm is also known as the Euclidean norm or the Euclidean distance when its square root is calculated. It sums up the squared difference between pixels. The L2 of two vectors, $\vec{\omega}$ and $\vec{t}$ is:

$$L2 = \sum_{i=1}^{M} (\omega_i - t_i)^2 \qquad (13)$$

*c) Mahalanobis distance*

The Mahalanobis distance calculates the product of the pixels and the eigenvalue of a specific dimension and sums all these products.The Mahalanobis distance of two vectors, $\vec{\omega}$ and $\vec{t}$ is:

$$MAH = -\sum_{i=1}^{N} \omega_i \cdot t_i \cdot h_i \qquad (14)$$

where

$$h_i = \frac{1}{\sqrt{\lambda_i}} \qquad (15)$$

and $\lambda_i$ is described in "(8)".

## III. Objects

The Columbia Object Image Library (COIL-20) [5] was used in our experiments, which is a database of 1440 grayscale images of 20 objects (72 images per object). The objects have a wide variety of complex geometric and reflectance characteristics (see Fig. 1).

During the acquisition process each object was placed on a turntable that rotates in steps of 5 degrees.

The two PCA approaches were performed, namely the PCA1 and the PCA2, to the original COIL images and to the binary and border images taken from the COIL-20 (see Fig. 2). The well-known Canny detector [13] was used to obtain the border images.

Fig. 1. The Columbia Object Image Library (COIL-20) contains 1440 images of 20 objects.

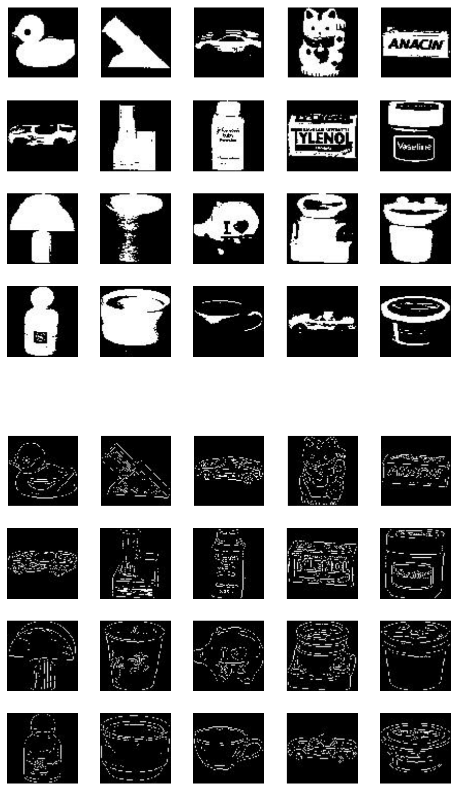

Fig. 2. Binary and border images from the COIL-20

## IV. Experiments

In this section a report on the extensive testings performed to compare the PCA approaches can be found.

All the algorithms were computed in MATLAB using LAPACK routines [6] to compute eigenvalues and eigenvectors. Using a Pentium IV with 256 Mbytes of RAM, to compute just one subspace (PCA1) takes three seconds, so obtaning all subspaces in PCA2 takes one minute of computation.

PCA2 takes one second to identify a new image; while PCA1 performs the same task in just a few hundredths of a second. Naturally PCA1 is faster than PCA2, but both approaches are fast enough to use them in a real time application.

In the following subsections will be shown the recognition rate results obtained during the experiments.

### A. PCA1 Results

In this technique just one image is used in training for each object, and the subspace is computed using the first image (the more representative image) of each object set or choosing one image randomly. So, in PCA1 approach 19 non-zero eigenvalues (#vectors) were obtained from the covariance matrix $\Omega'$ and the higher recognition rate was reached when using all the eigenvectors.

Table 1 shows the best results obtained during the PCA1 experiments. The distance measures that performed better were L2 norm and Mahalanobis distance. Almost the same recognition rate was reached for the binary set than for the original images.

TABLE I
PCA1 RESULTS

| images set | original COIL | | | |
|---|---|---|---|---|
| training | first | random | first | random |
| norm | L2 | L2 | Mahalanobis | Mahalanobis |
| # vectors | 19 | 19 | 19 | 19 |
| recognition rate | 57 % | 52 % | 64 % | 60 % |
| images set | border COIL | | | |
| training | first | random | first | random |
| norm | L2 | L2 | Mahalanobis | Mahalanobis |
| # vectors | 19 | 19 | 19 | 19 |
| recognition rate | 23% | 20% | 52% | 53% |
| images set | binary COIL | | | |
| training | first | random | first | random |
| norm | L2 | L2 | Mahalanobis | Mahalanobis |
| # vectors | 19 | 19 | 19 | 19 |
| recognition rate | 56% | 55% | 64% | 61% |

### B. PCA2 Results

In this approach, two types of training were performed, namely best and random. In best training 36 ordered images were selected for training and 36 for testing. In this best training we get a different view of the object with a 10 degrees separation. So, each training image x is very similar to test image x'. In random training the 36 training images and the 36 test images were selected randomly, so not all the relevants views of the object are present in the data matrix.

Not all the objects are easy to recognize; there are objects very similar between them as the three toy cars, or very different to the rest as the cup, so the whole set was split in three subsets from the easier to the most dificult to classify (the objects are numbered by rows begining with the duck image):

a. set I: 7,12, 15,16,17,18,20
b. set II: 1,8, 10, 11, 13
c. set III:2,3,5,6,9,14,19

Table II shows the recognition rates obtained by PCA2 for different distance measures changing the number of eigenvectors used to represent the subspace for the original COIL images (see Fig. 1). In this experiment L2 norm performs much better than Mahalanobis distance (85% recognition rate versus 19% recognition rate).

Table II also shows how very different rates are associated to the three different object subsets.

The same experiment was performed with random training (see Table III). Obviously best training performs better than random, but the difference is smaller than expected obtaining a 81% in random training.

TABLE II
PCA2 RESULTS FOR DIFFERENT DISTANCE MEASURES
BEST TRAINING

| images set | original COIL | | | | | | |
|---|---|---|---|---|---|---|---|
| # vectors | 5 | 10 | 15 | 20 | 25 | 30 | 35 |
| L2 norm | recognition rate | | | | | | |
| I | 92 | 98 | 99 | 99 | 100 | 100 | 100 |
| II | 60 | 89 | 85 | 89 | 92 | 92 | 94 |
| III | 42 | 55 | 63 | 62 | 61 | 60 | 61 |
| total | 65 | 81 | 82 | 83 | 84 | 84 | 85 |
| L1 norm | recognition rate | | | | | | |
| I | 93 | 96 | 98 | 98 | 98 | 99 | 100 |
| II | 58 | 83 | 80 | 84 | 87 | 90 | 92 |
| III | 39 | 53 | 62 | 62 | 56 | 51 | 54 |
| total | 63 | 77 | 80 | 81 | 80 | 80 | 82 |
| Infinite norm | recognition rate | | | | | | |
| I | 91 | 95 | 97 | 97 | 98 | 98 | 98 |
| II | 55 | 83 | 81 | 88 | 86 | 86 | 86 |
| III | 40 | 56 | 60 | 60 | 56 | 56 | 56 |
| total | 62 | 78 | 79 | 81 | 80 | 80 | 80 |

Both experiments were developed fixing the number of training images (36 images) and incrementing the number of eigenvectors (5-35), so the recognition rate is increased as the number of eigenvectors used is incremented as is well-known in eigenspaces representation. But how many images do we need to train a recognition system is still unresolved. To find out this parameter the experiment shown in Fig. 3 was performed. PCA2 was computed for different training images set: the smaller having just 5 training images for each object , and the bigger having 65 training images. In all cases the images were projected in the subspace composed by the 90% of all eigenvectors, so for the smaller set there were 4 eigenvectors and for the bigger, 58.

TABLE III

PCA2 RESULTS FOR RANDOM TRAINING

| images set | original COIL | | | | | | |
|---|---|---|---|---|---|---|---|
| # vectors | 5 | 10 | 15 | 20 | 25 | 30 | 35 |
| L2 norm | recognition rate | | | | | | |
| I | 85 | 89 | 92 | 96 | 96 | 96 | 97 |
| II | 57 | 72 | 72 | 79 | 81 | 84 | 88 |
| III | 38 | 53 | 54 | 54 | 55 | 57 | 58 |
| total | 60 | 71 | 73 | 76 | 77 | 79 | 81 |

Fig.3 shows a saturation in the training around the experiment with 25 images, associated to 22 eigenvectors. This saturation is even more clear for the most dificult subset.

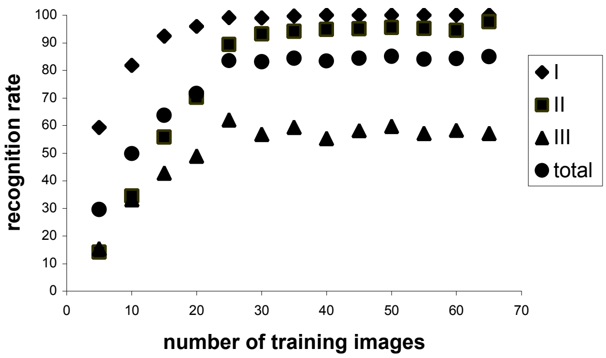


Fig. 3. PCA2 recognition rates for different training sets.

Tables IV-V show the same experiment as Table II but using binary and border images (see Fig.2).

Using binary images the best distance measure was the L2 norm and the higher recognition rate (72%) was obtained for just 10 eigenvectors . If all the eigenvectors are used to project the training images a lower rate (50%) is obtained, this behavior is opposite to that obtained with the original COIL images, where the rate increases as the number of eigenvectors.

Using border images the best success rate is obtained, a 95% for the total subsets, even better than using the original images. So, it can be concluded that using a border detector before the subspace computation improves the recognition results.

TABLE IV

PCA2 RESULTS FOR BINARY COIL

| images set | binary COIL | | | | | | |
|---|---|---|---|---|---|---|---|
| # vectors | 5 | 10 | 15 | 20 | 25 | 30 | 35 |
| L2 norm, best training | recognition rate | | | | | | |
| I | 80 | 84 | 78 | 77 | 76 | 74 | 68 |
| II | 81 | 79 | 71 | 65 | 58 | 54 | 50 |
| III | 53 | 54 | 52 | 44 | 39 | 36 | 31 |
| total | 71 | 72 | 67 | 62 | 58 | 55 | 50 |

TABLE V

PCA2 RESULTS FOR BORDER COIL

| images set | border COIL | | | | |
|---|---|---|---|---|---|
| # vectors | 5 | 10 | 20 | 33 | 35 |
| Mahalanobis, best training | recognition rate | | | | |
| I | 90 | 95 | 98 | 100 | 100 |
| II | 67 | 83 | 95 | 99 | 99 |
| III | 50 | 54 | 74 | 87 | 87 |
| total | 69 | 77 | 89 | 95 | 95 |

### *C. Occlusion results*

Experiments with partially occluded test images were performed in order to prove the robustness of the two methods (see Figs. 4-5).

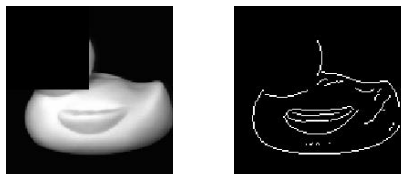

Fig. 4. Border images with 25% occlusion

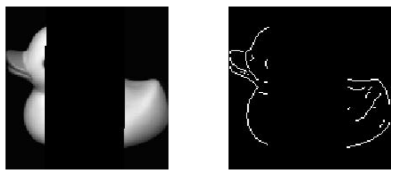

Fig. 5. Border images with 50% occlusion

The results indicate that the PCA2 method using border images is the more robust approach in occlusion problems.

Different recognition rates were also obtained depending on the distance measure used. For original images using the PCA2 approach, the L2 norm works better than the others, while in the rest of trainings the best rates are obtained using the Mahalanobis distance (see Tables VI - VII).

Fig. 6 shows the increase in the succes rates according to the increase in the number of eigenvectors that created the subspaces for the best approach using the 50% partially occluded border images.

TABLE VI
RESULTS FOR 25% OCCLUSION

| method | distance measure | recognition rate |
|---|---|---|
| original COIL | Mahalanobis | 51 |
| PCA1 | L2 norm | 9 |
| border COIL | Mahalanobis | 39 |
| PCA1 | L2 norm | 6 |
| original COIL | Mahalanobis | 5 |
| PCA2 | L2 norm | 30 |
| border COIL | Mahalanobis | 89 |
| PCA2 | L2 norm | 17 |

TABLE VII
RESULTS FOR 50% OCCLUSION

| method | distance measure | recognition rate |
|---|---|---|
| original COIL | Mahalanobis | 16 |
| PCA1 | L2 norm | 7 |
| border COIL | Mahalanobis | 32 |
| PCA1 | L2 norm | 11 |
| original COIL | Mahalanobis | 2 |
| PCA2 | L2 norm | 8 |
| border COIL | Mahalanobis | 59 |
| PCA2 | L2 norm | 5 |

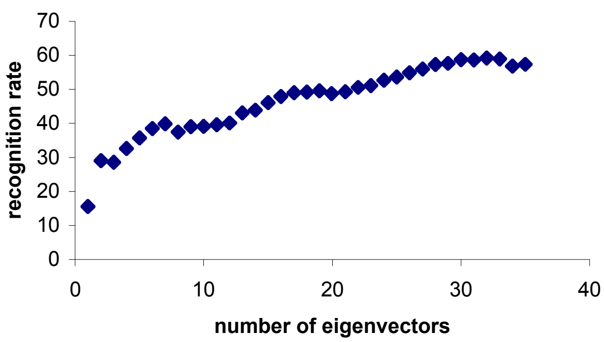


Fig. 6. PCA2 recognition rates for different number of eigenvectors using partially occluded images.

## V. CONCLUSION

The experimental results clearly show how PCA2 outperforms the classical PCA1 method.

A new approach to object recognition using PCA2 has been tested with better results than previous approaches: preprocessing the original images with a border detector and using such border images as input to the PCA2 improves the recognition rate both with complete images and with partially occluded images.

The improvement in recognition rate is particularly high with ocludded images.

## REFERENCES


[1] M.Turk and A.Pentland, *Eigenfaces for Recognition,* Journal of Cognitive Neuroscience, vol. 3, no. 1, pp.71-86,1991.

[2] Ales Leonardis and Horst Bishop. *Robust Recognition Using Eigenimages*. Computer Vision and Image Understanding, 78, no. 1, pages 99-118, 2000.

[3] Chien-Yuan Huang, Octavia I. Camps. *Object Recognition Using Appearance-Based Parts and Relations.*. Proceedings of the IEEE CVPRC 1997.

[4] S. Chandrasekaran, B.S. Manjunath, Y.F. Wang, J.Winkeler, H.Zhang. *An Eigenspace Update Algorithm for Image Analysis*. Graphical Models and Image Processing. Vol 59, Nº 5, Sept., pp 321-332, 1997.

[5] S. A. Nene, S. K. Nayar and H. Murase,*Columbia Object Image Library (COIL-20),* Technical Report CUCS-005-96, February 1996.

[6] Anderson, E., Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, *LAPACK User's Guide* (http://www.netlib.org/lapack/lug/lapack_lug.html), Third Edition, SIAM, Philadelphia, 1999.

[7] The Face Recognition Home Page http://www.cs.rug.nl/~peterkr/FACE/frhp.html

[8] W..Zhao., R.Chellapa and P.J. Phillips. *Subspace Linear Discriminant Analysis for Face Recognition*.1999 (in press)

[9] R.A. Fisher. *The use of Multiple Measures in Taxonomic Problems*. Ann. Eugenics, 7:179-188, 1936.

[10] Pierre Common. *Independent component analysis, A new concept?* Signal Processing, 36:287--314, 1994.

[11] Anthony J. Bell and Terrence J. Sejnowski Howard Hughes,*The `Independent Components' of Natural Scenes are Edge Filters*. Vision Research, 1997

[12] Aapo Hyvärinen, Patrik O. Hoyer, Mika Inki. Topographic Independent Component *Analysis.* Neural Computation, 2000 (in press)

[13] Canny, J. *A computational approach to Edge Detection*. IEEE Trans. Pattern Analysis and Machine Intelligence, 8-6, pp. 679-698, 1986.

[14] Horn R. and Johnson C., Matrix Analysis, New York: Cambridge University Press, 1985.

[15] Kyungim Baek,Bruce A.Draper,J.Ross Beveridge and Kai She. *PCA vs.ICA:A Comparison on the FERET Data Set*, International Conference on Computer Vision, Pattern Recognition and Image Processing in conjunction with the 6th JCIS, Durham, North Carolina, March 8-14, 2002.June 2001.

[16] Murase H. and Nayar S.K. *Visual Learning and Recognition of 3-D Objects from Appearance.* International Journal od Computer Vision, 14:5-24, Januray 1995.

[17] S. K. Nayar, H. Murase, and S. A. Nene, *Parametric Appearance Representation*, in Early Visual Learning, edited by S. K. Nayar and T. Poggio, Oxford University Press, February 1996.