



University of
Salford
MANCHESTER

The evaluation of camera motion, defocusing and noise immunity for linear appearance based methods in face recognition

Hajiarbabi, M, Askari, J, Sadri, S and Saraee, MH

Title	The evaluation of camera motion, defocusing and noise immunity for linear appearance based methods in face recognition
Authors	Hajiarbabi, M, Askari, J, Sadri, S and Saraee, MH
Publication title	Proceedings of the World Congress on Engineering 2007 Vol I WCE 2007, July 2 - 4, 2007, London, U.K
Publisher	WCE
Type	Conference or Workshop Item
USIR URL	This version is available at: http://usir.salford.ac.uk/id/eprint/18851/
Published Date	2007

USIR is a digital collection of the research output of the University of Salford. Where copyright permits, full text material held in the repository is made freely available online and can be read, downloaded and copied for non-commercial private study or research purposes. Please check the manuscript for any further copyright restrictions.

For more information, including our policy and submission procedure, please contact the Repository Team at: library-research@salford.ac.uk.

The Evaluation of Camera Motion, Defocusing and Noise Immunity for Linear Appearance Based Methods in Face Recognition

M. Hajiarbabi, J. Askari, S. Sadri, and M. Saraee

Abstract— Face recognition has assigned a special place to itself because of its low intrusiveness, low cost and effort and acceptable accuracy. There are several methods for recognition and appearance based methods is one of the most popular one. Unfortunately most of the papers that have been published these years have just shown the results on the databases that are all without any noise and all of focus. But it is clear that for a real system all these problems can happen, so finding methods that are robust to such problems is important. In this paper we show that linear appearance based methods are robust to an acceptable degree to problems such as, when the camera is moving or it is defocus and when the image is influenced with Gaussian noise. For linear appearance based methods we chose Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Multiple Exemplar Discriminant Analysis (MEDA) that has shown better performance than other appearance based methods.

Index Terms— Linear Discriminant Analysis, Principal Component Analysis, Multiple Exemplar Discriminant Analysis.

I. INTRODUCTION

Human identification recognition has attracted the scientists from so many years ago. Due to increasing in terrorism the needs for such systems have increased much more. The most important biometric systems which have been used during these years we can name fingerprint recognition, speech recognition, iris, retina, hand geometry and face recognition. For comparing biometric systems four features have been considered: intrusiveness, accuracy, cost and effort. The investigation has shown that among the other biometric systems, face recognition is the best one [1].

Among the various methods that has been applied to face recognition, appearance based methods have shown better

results, and have become the most popular method in face recognition. But a real system may encounter with some difficulties which are unexpected, for example when the camera or the subject is moving, when the camera is defocus or when the image has been influenced by noise, we need methods to be robust to such problems.

The goal of this paper is show the degree of robustness to camera moving, defocusing and noise immunity for linear appearance based methods.

The rest of the paper is organized as follows: in section 2 PCA, in section 3 LDA and in section 4 MEDA will be reviewed. In section 5 the RBF classifier and distance based classifier are reviewed. And finally in section 6 our work and the results are presented on the ORL [2] database.

II. PRINCIPAL COMPONENT ANALYSIS

PCA is a method to efficiently represent a collection of sample points, reducing the dimensionality of the description by projecting the points onto the principal axes, where an orthonormal set of axes points in the direction of maximum covariance in the data. These vectors best account for the distribution of face images within the entire image space. PCA minimizes the mean squared projection error for a given number of dimensions, and provides a measure of importance (in terms of total projection error) for each axis.

Let us now describe the PCA algorithm [3]. Consider that Z_i is a two dimensional image with size $m \times m$. First we convert the matrix into a vector of size m^2 . The training set of the n face can be written as:

$$Z = (Z_1, Z_2, \dots, Z_n) \subset \mathfrak{R}^{m^2 \times n} \quad (1)$$

Each of the face images belongs to one of the c classes. In face recognition the total images that belong to one person is considered as one class. For the training images the covariance matrix can be computed by:

$$\Gamma = \frac{1}{n} \sum_{i=1}^n (Z_i - \bar{Z})(Z_i - \bar{Z})^T = \Phi \Phi^T \quad (2)$$

where $\Phi = (\Phi_1, \Phi_2, \dots, \Phi_n) \subset \mathfrak{R}^{m^2 \times n}$ and $\bar{Z} = \left(\frac{1}{n}\right) \sum_{i=1}^n Z_i$ is the average of the training images in the database. As can be

Manuscript received March 3, 2007.

M. Hajiarbabi, is an M.S. student in the Electrical and Computer Engineering Department, Isfahan University of Technology, Iran (corresponding author to provide phone: 09151107128; e-mail: m_arbabi@ec.iut.ac.ir).

J. Askari, is with the Electrical and Computer Engineering Department, Isfahan University of Technology, Iran. (e-mail: j-askari@cc.iut.ac.ir).

S. Sadri is with the Electrical and Computer Engineering Department, Isfahan University of Technology, Iran. (e-mail: sadri@cc.iut.ac.ir).

M. Saraee is with the Electrical and Computer Engineering Department, Isfahan University of Technology, Iran. (e-mail: saraee@cc.iut.ac.ir).

easily seen the face images have been centered which means that the mean of the images is subtracted from each image.

After computing covariance matrix, the eigenvectors and eigenvalues of the covariance matrix will be computed.

Consider that $U = (U_1, U_2, \dots, U_r) \in \mathfrak{R}^{m^2 \times r}$ ($r < n$) be the eigenvectors of the covariance matrix, only small parts of this eigenvectors, for example r , that have the larger eigenvalues will be enough to reconstruct the image. So by having an initial set of face images $Z \in \mathfrak{R}^{m^2 \times n}$ the feature vector corresponding to its related eigenface $X \in \mathfrak{R}^{r \times n}$ can be calculated by projecting Z in the eigenface space by

$$X = U^T Z \quad (3)$$

III. LINEAR DISCRIMINANT ANALYSIS

LDA is used for projecting a set of training data. In face recognition and because of singularity problem it is common to use PCA first on the image and then apply LDA or MEDA to it. So consider that $X = (X_1, X_2, \dots, X_n)$ is a matrix containing the vectors in the training set. X_i is a vector that has been calculated from an image after applying PCA to it. In LDA two matrixes within class scatter matrix and between class scatter matrixes is defined. This method finds an optimal subspace in which the between class scatter matrix to the within class scatter matrix will be maximized [4]. The between class scatter matrix is computed by

$$S_B = \sum_{i=1}^c n^i (\bar{X}^i - \bar{X})(\bar{X}^i - \bar{X})^T \quad (4)$$

Where $\bar{X} = \left(\frac{1}{n}\right) \sum_{j=1}^n X_j$ is the mean of the vectors in the training set and $\bar{X}^i = \left(\frac{1}{n^i}\right) \sum_{j=1}^{n^i} X_j^i$ is the mean of class i ,

and c is the number of the classes. The between class scatter matrix defines the average scattering of one class across the average of the total classes. The within class scatter matrix is computed by

$$S_W = \sum_{i=1}^c \sum_{X_i \in n^i} \left(X_i - \bar{X}^i \right) \left(X_i - \bar{X}^i \right)^T \quad (5)$$

The within class scatter matrix defines the data of one class across the average of the class. The optimal subspace is computed by

$$E_{optimal} = \operatorname{argmax}_E \frac{\|E^T S_B E\|}{\|E^T S_W E\|} = [c_1, c_2, \dots, c_{c-1}] \quad (6)$$

Where $[c_1, c_2, \dots, c_{c-1}]$ is the set of eigenvectors of S_B and S_W corresponding to $c - 1$ greatest generalized eigenvalue λ_i and $S_B E_i = \lambda_i S_W E_i$ $i = 1, 2, \dots, c - 1$

$E_{optimal}$ is an optimal matrix which maximize the proportion of between class scatter matrix to the within class scatter matrix. This means that it maximize the scattering of the data that

belongs to different classes and minimize the scattering of the data belonging to the same class.

Thus the most discriminant answer for face vectors X would be [4]:

$$P = E_{optimal}^T \cdot X \quad (8)$$

IV. MULTIPLE EXEMPLAR DISCRIMINANT ANALYSIS

The problem of face recognition differs from other pattern recognition problems and so it needs different discriminant methods other than LDA. In LDA the classification of each class is based on just one sample and that's the mean of each class. Because of lacking samples in face recognition problem it is better to use all the samples instead of the mean of each class for classification. Rather than minimizing the within class distance while maximizing the between class distance, multiple exemplar discriminant analysis (MEDA) finds the projection directions along which the within class exemplar distance (i.e. the distances between exemplars belonging to the same class) is minimized while the between-class exemplar distance (i.e. the distances between exemplars belonging to different classes) is maximized [5].

In MEDA the within class scatter matrix is computed by

$$S_W = \sum_{i=1}^c \frac{1}{n^i} \sum_{j=1}^{n^i} \sum_{k=1}^{n^i} (X_j^i - X_k^i)(X_j^i - X_k^i)^T \quad (9)$$

Where X_j^i is the j th image vector of i th class. By comparing it with the within class scatter matrix of LDA we see that in this method all the images in a class has participated in making the within class scatter matrix instead of using just the mean of the class as in LDA method. The between class scatter matrix is computed by

$$S_B = \sum_{i=1}^c \sum_{j=1, j \neq i}^c \frac{1}{n^i n^j} \sum_{k=1}^{n^i} \sum_{l=1}^{n^j} (X_k^i - X_l^j)(X_k^i - X_l^j)^T \quad (10)$$

Contrary to LDA in which the means of each class and means of all samples made the between class scatter matrix, in MEDA all the samples in one class will be compared to all samples of the other class. The computation of $E_{optimal}$ is the same as LDA.

V. CLASSIFIERS

For classification we used distance measures and also RBF neural network in order to compare their classification power. For distance measures we selected the following distances:

Euclidean distance (L_2):

$$d(x, y) = \|x - y\|^2 = \sum_{i=1}^k (x_i - y_i)^2 \quad (11)$$

City block distance (L_1):

$$d(x, y) = |x - y| = \sum_{i=1}^k |x_i - y_i| \quad (12)$$

Mahalanobis distance:

$$d(x, y) = -\sum_{i=1}^k x_i y_i z_i \quad (13)$$

$$z_i = \frac{1}{\lambda_i^{1/2}}$$

Where λ_i is the eigenvalue of i th eigenvector.

RBF neural network is a powerful classification method for pattern recognition problems. It doesn't have the drawbacks of multi layer perceptron neural networks and trains much faster than it. Fig. 1 shows an RBF neural network.

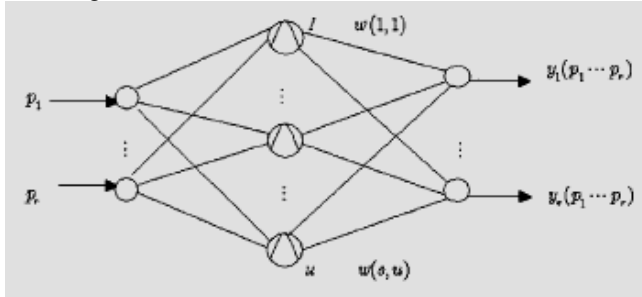


Fig. 1: RBF neural network

Let $P \in \mathcal{R}^r$ be the input vector and $C_i \in \mathcal{R}^r$ ($1 \leq i \leq u$) be the prototype of the input vectors. The output of each RBF units is as follows:

$$R_i(P) = \exp\left\{-\frac{\|P - C_i\|^2}{\sigma_i^2}\right\} \quad (14)$$

Where σ_i is the width of the i th RBF unit. The j th output $y_j(P)$ of an RBF neural network is

$$y_j(p) = \sum_{i=1}^u R_i(P) * w(j, i) + w(j, 0) \quad (15)$$

Where $R_0 = 1$, $w(j, i)$ is the weight of the i th receptive field to the j th output. The weights of first layer are all equal to one. The number of nodes in the second layer at first equals to the number of classes. Whenever two classes have intersection with each other a node is added to the second layer and a class is split into two subclasses. For further knowledge about RBF neural network the reader can refer to neural network references.

VI. EVALUATION OF CAMERA MOTION, DEFOCUSING AND NOISE IMMUNITY

In order to evaluate the algorithms mentioned above we used ORL database. ORL database contains 400 images that belong to 40 people with variety in scale and small variety in pose head. 5 images from every person was used as training set and the rest used as test set Fig. 2 shows a sample of this database.



Fig. 2: ORL database

We extracted 35 features for each method then by using RBF neural network and also Euclidean, City block and Mahalanobis distance we classified the data. As mentioned before, for LDA and MEDA we first applied PCA on the images and then applied LDA and MEDA on the new vectors in order to avoid the singularity problem. The input size of the neural network is equal to the size of the vector and the output size of it is 40 equal to number of classes.

A. Noise immunity

For evaluation the effect of noise in recognition rate, we used Gaussian noise with zero mean and with variances 0.01, 0.05, 0.1, 0.5, and 1; we applied the noise just on the test images. Fig. 3 shows the testing images after applying the proposed noise to them; the first images from left have Gaussian noise with variance 0.01 and the right most images have Gaussian noise with variance 1.

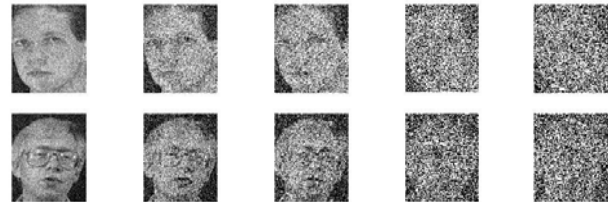


Fig. 3: Test images after applying Gaussian noise

Fig. 4-6 show the recognition rate for PCA, LDA and MEDA respectively. As can be easily seen the three methods are robust to Gaussian noise up to 0.1 variance or even more. The Mahalanobis distance performs quite well as RBF classifier.

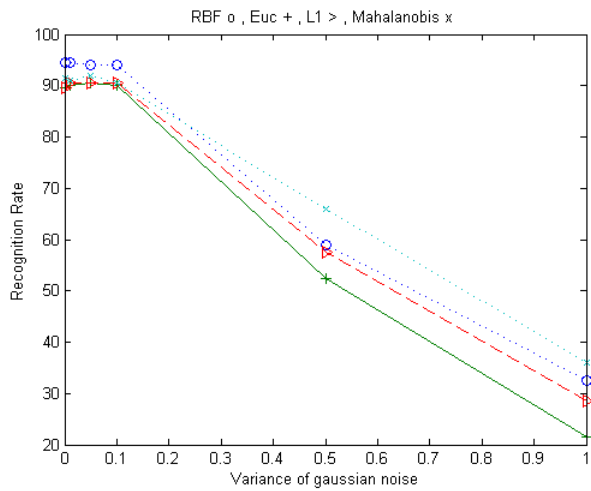


Fig. 4: The effect of Gaussian noise on PCA

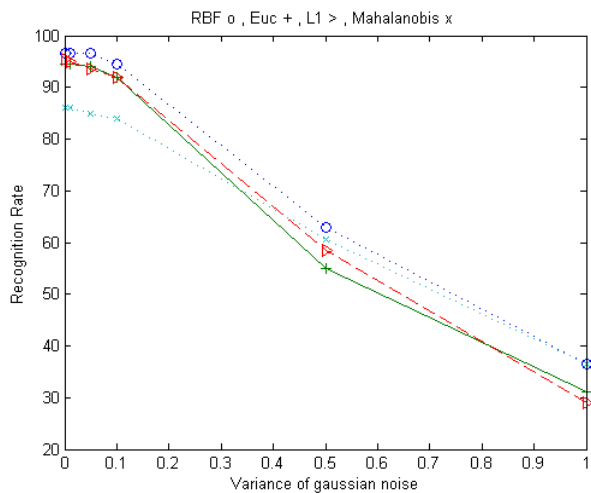


Fig. 5: The effect of Gaussian noise on LDA

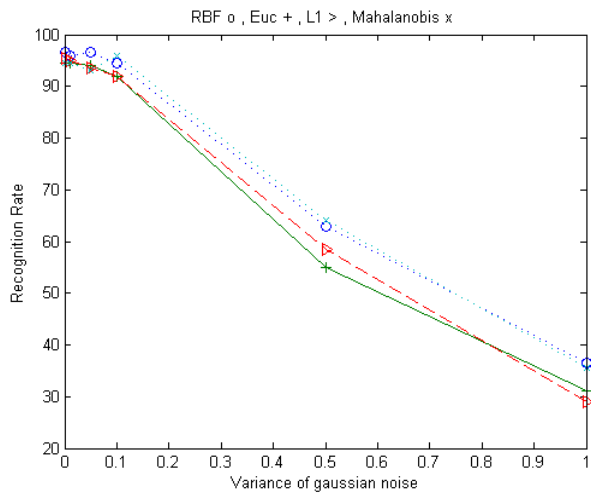


Fig. 6: The effect of Gaussian noise on MEDA

B. Camera motion immunity

For camera motion problem we shift the image to desired pixels with zero degree and then convolve it with the origin picture and so the camera motion is simulated. Fig. 7 shows the

images from left to right after applying the filter with 4, 8, 16, 24 and 40 pixels shifting and convolving with the original image.



Fig. 7: Test images after applying camera motion

Fig. 8-10 shows the recognition rate for PCA, LDA and MEDA respectively. The results show good immunity in the case of camera motion. Something which is quite interesting in the results is the increasing in recognition rate, although we are increasing the camera motion. The reason can be this: applying some changes to the images, cause the features vectors extracted from the images and that belong to the same class becomes closer to each other in the subspace, so it increases the recognition rate although everyone expect it to decrease. Another thing which is interesting is the RBF neural network behavior. At first and when there is no or little camera motion it quite outperforms other classifiers, but as the camera motion increases the performance of RBF decreases more rapidly than other classifiers.

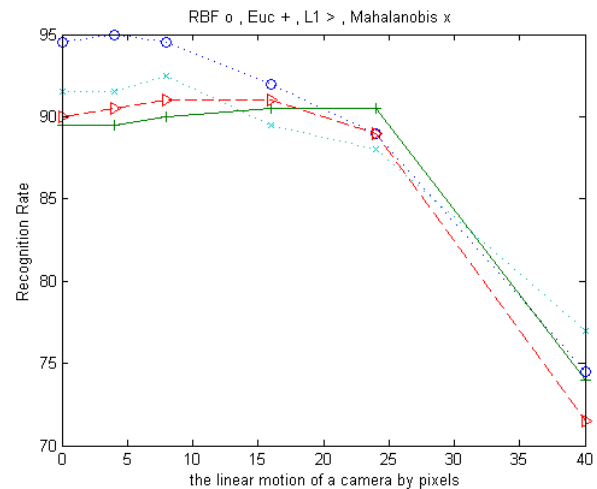


Fig. 8: The effect of camera motion on PCA

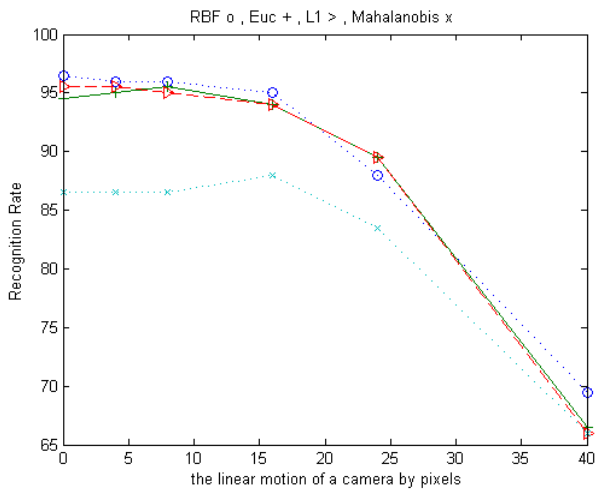


Fig. 9: The effect of camera motion on LDA

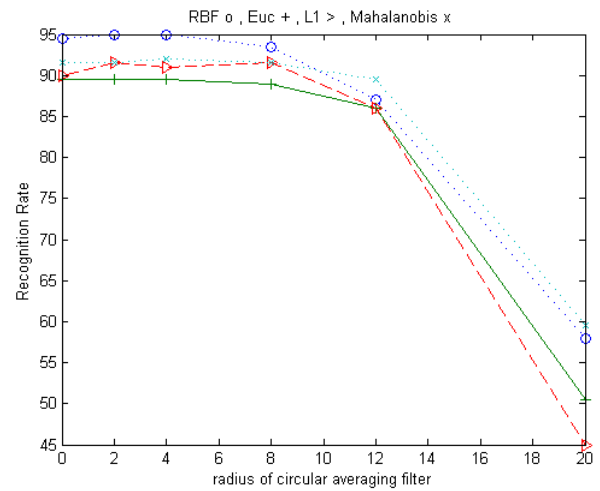


Fig. 12: The effect of defocusing on PCA

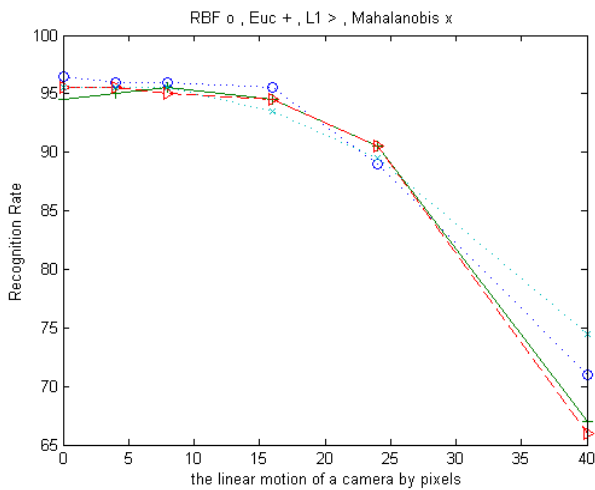


Fig. 10.: The effect of camera motion on MEDA

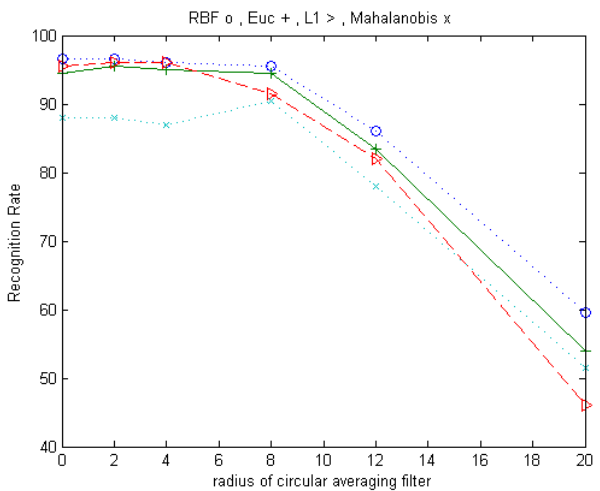


Fig. 13.: The effect of defocusing on LDA

C. Defocusing immunity

For defocusing we used a circular average filter with radius 2, 4, 8, 12 and 20, which was convolved with the origin image. Fig. 11 show the images from left to right after convolving the filters with radius 2, 4, 8, 12 and 20 with them respectively.



Fig. 11: Test images after applying defocusing

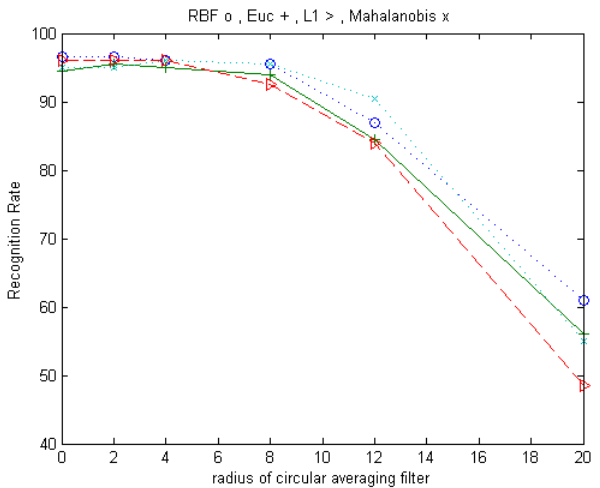


Fig. 14: The effect of defocusing on MEDA

Fig. 12-14 show the recognition rate for PCA, LDA and MEDA respectively. The results show good immunity in the case of defocusing. The results show even when the recognizing of the test images is difficult for human eye, it have been recognized by the appearance based methods.

VII. CONCLUSION

In this paper we showed that the linear appearance based methods are immune to noise, defocusing and camera motion to acceptable degree. This work can be applied to other methods to see their results on such problems.

REFERENCES

- [1] R. Hietmeyer, "Biometric identification promises fast and secure processing of airline passengers," *The Int'l Civil Aviation Organization Journal*, vol. 55, no. 9, 2000, pp. 10–11.
- [2] ORL database: <http://www.camorl.co.uk>.
- [3] M. Turk, and A. Pentland,, "Eigenfaces for recognition," *J. Cognitiv Neurosci.*, vol. 3, 1991, pp. 71–86.
- [4] K. Fukunaga,, *Introduction to statistical pattern recognition*, 2nd ed. San Diego, CA: Academic Press, 1990, pp. 445-450.
- [5] Sh. K. Zhou, and R.Chellappa, "Multiple-exemplar discriminant analysis for face recognition" *Center for Automation Research and Department of Electrical and Computer Engineering University of Maryland, College Park, MD 20742*, 2003.