The POSTECH Face Database (PF07) and Performance Evaluation

Hyoung-Soo Lee[1], Sungsoo Park[1], Bong-Nam Kang[2], Jongju Shin[1]
Ju-Young Lee[1], Hongmo Je[1], Bongjin Jun[1],Daijin Kim[1]

[1]Department of Computer Science and Engineering, POSTECH
San 31, Hyoja-Dong, Nam-Gu, Pohang, 790-784, Korea
{sooz, sungs, jjshin, zealota, invu71, simple21, dkim}@postech.ac.kr

[2]Department of Computer Science, Yonsei University
134,Shinchon-Dong, Seodaemun-Gu, Seoul, 126-749, Korea
bongnam@yonsei.ac.kr

Abstract

We constructed a face database POSTECH face database (PF07). PF07 contains the true-color face images of 200 people, 100 men and 100 women, representing 320 various images (5 pose variations $\times$ 4 expression variations $\times$ 16 illumination variations) per person. All of the people in the database are Korean. We also present the results of face recognition experiments under various conditions using three baseline face recognition algorithms in order to provide an example evaluation protocol on the database. The database is expected to be used to evaluate the algorithm of face recognition for Korean people or for people with systematic variations.

1. Introduction

In modern life, the need for personal security and access control becomes important issue. Biometrics is a technology which is expected to replace traditional authentication methods which is easy to be stolen, forgotten and duplicated. Fingerprints, face, iris, and voiceprints are commonly used biometric features. Among these features, face provides more direct, friendly and convenient identification way and is more acceptable compared with individual identification ways of other biometrics features. Thus, face recognition is one of the most important parts in biometrics. For the development of face recognition algorithm, we need the face image database to evaluate the performance of face recognition algorithm.

Many face databases have been reported. Face databases can be categorized according to their purposes. Some databases can be used for face recognition technology, other for face expression recognition or for pose estimation. In this paper, we concentrate on face databases for face recognition. The representative databases are FERET database, XM2VTS database, Yale DB, AR Face database, PIE database, CAS-PEAL database, and KFDB. We explain the representative databases briefly.

FERET database [8] The FERET database was constructed to develop automatic face recognition capabilities that can be employed to assist security, intelligence and law enforcement personnel, sponsored by the Department of Defenses Conterdrug Technology Development Program through the Defense Advanced Research Products Agency. It has 14051 eight-bit grayscale images of human faces. It includes face images of various poses, including profiles of alternative expressions and of different illuminations. For some people, it includes face images with eye glasses worn, with different hair length, and both. Pose variations are systematic and various, but other variations are not. Nevertheless, since it includes many face images of many people, it is one of the best-known face databases.

XM2VTS database [7] The XM2VTS database is a large multi-modal database to test multi-modal face verification algorithms. It was made by the Center for Vision, Speech and Signal Processing, at the University of Surrey, England. It contains four recordings of 295 people over four months. Each recording contains a speaking head shot and a rotating head shot. The shots were stored in high quality color images, 23KHz 16-bit sound files of speech, video sequences of rotating heads, and a 3D model of a face.

978-1-4244-2154-1/08/$25.00 ©2008 IEEE

Yale database [1] The Yale face database was made by the Center for Computational Vision and Control, at Yale University. It contains gray face images of 15 people, where there are images of 11 variations for each person. Images for each person are normal images, images with glasses and without glasses, images with light variations (such as center-light, left-light and right-light), and images with expression variations (such as happy, sad, sleepy, surprised and winking).

AR face database [6] AR face database was made by Computer Vision Center at UAB(Universitat Autonoma de Barcelona). It contains the color images of 126 people (70 men and 56 women), where there are frontal faces with different facial expressions, illumination conditions, and occlusions (sun glasses and scarf). The pictures were taken at the CVC under strictly controlled conditions. No restrictions on wear (clothes, glasses, etc.), make-up, hair style, etc. were imposed on participants. Each person participated in two sessions, separated by two weeks (14 days). The same pictures were taken in both sessions.

PIE database [9] The CMU PIE database was made by the robotics institute, at CMU. It contains 41,368 images of 68 individuals. It systematically samples a large number of pose and illumination conditions along with a variety of facial expressions. Each subject were imaged using 13 cameras and 21 flashes. Each subject were asked to display a neutral face, to smile, and to close their eyes in order to simulate a blink.

CAS-PEAL database [3] The CAS-PEAL (pose, expression, accessory, lighting) database wad constructed by the Chinese Academy of Sciences (CAS). It contains 99,594 images of 1040 subjects, where there are pose, expression, accessory, lighting, background, distance, and time variations. For the pose variations, nine cameras are used around the subject. For the illumination variations, nine lamps were placed around the subject. Subjects were asked to smile, to frown, to look surprised, to close their eyes, and to open the mouth for facial expressions. Some subjects were allowed to wear three types of glasses and three types of hats.

Korean face database [5] The Korean face database (KFDB) was collected by the center for artificial vision research at Korea university. It contains 52,000 images of 1,000 subjects (500 male and 500 female), where there are pose, expression, illumination variations. The database consists of only Korean subjects. The cameras were placed between $45\circ$ off frontal in both directions at $15\circ$ increments for seven poses. The lights were placed in a full circle around the subject at $45\circ$ interval for nine illuminations. The subjects were asked to display five facial expressions - neutral, happy, surprise, anger, and blink.

There are two drawbacks in the existing face databases including those mentioned above. The first drawback is that most databases do not contain many Asian people. The KFDB and CAS-PEAL database contain many asian people, but the KFDB is not publicly distributed and the CAS-PEAL database is partly available by now. The second drawback is that there is no database which has simultaneously changing variations. Some databases, such as PIE database and CAS-PEAL database, contain two simultaneously changing variations, but not for other variations.

To overcome these drawbacks, we have built a new database called POSTECH face database (PF07). It contains face images of Korean people. There are face variations for pose, expression, and illumination. Variations are systematic and changing simultaneously. We also present the results of face recognition experiments under various conditions using three baseline face recognition algorithms such as principal component analysis (PCA), PCA without 3 most significant basis vectors (PCA/wo3) and linear discriminant analysis (LDA) in order to provide an example evaluation protocol on the database.

2. POSTECH Face Database

2.1. Overview

PF07 database includes 64,000 images because there are 100 male and 100 female subjects with 320 images per each subject. Each image is recorded in BMP format. The image is of size 640×480 and of true color. The database also includes the information for normalization, which consists of two eye points. We converted true color images into the gray color images and normalized them using the two eye points. The normalized images are of size 40×50. We provide the normalized images as a part of the database. Figure 1 shows the full normalized image set of a subject in the database. In this figure, four image groups represent the different expressions, the five columns in each image group represent the different poses, and five rows in each image group represent some of 16 different illuminations due to the limitation of page space.

2.2. Face Variations

The face image database should include many variations, since it is used to evaluate the performance of face recognition algorithms appropriately. The sources of variability in facial appearance can be either intrinsic or extrinsic. Intrinsic variables for face variations are facial expression, clothing, facial occlusion, hair, and so on. Extrinsic variables are alignment, background, camera adjustment, illumination, imaging device, pose, scale, and so on. In the existing

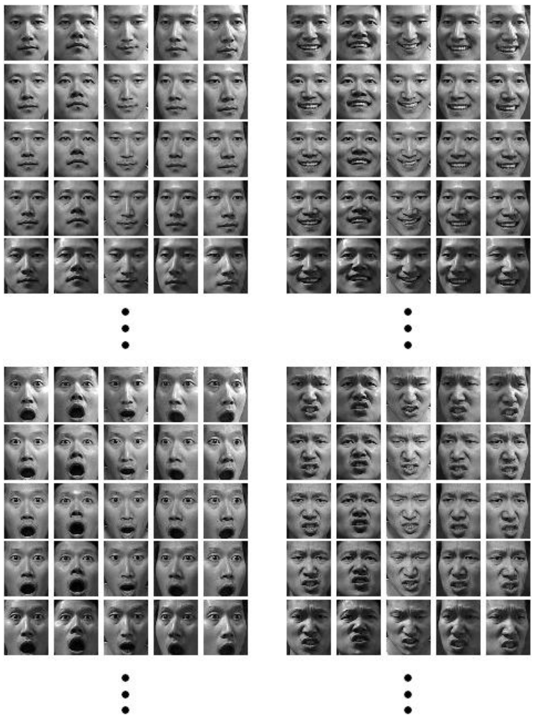


Figure 1. Example images of PF07 database.

face image databases, there are many variations, including illumination, pose, expression, hair, glasses, earrings, etc. Some databases include many kinds of variations, but other databases do not. Our database includes one intrinsic variation of expression and two extrinsic variations of illumination and pose. The variations in our database are systematic.

2.2.1 Pose

Five poses, which consist of frontal, upper, lower, left, and right, are considered. The data could be used for evaluating the robustness of face recognition algorithms across pose and evaluating the pose estimation algorithms. Figure 2 shows the five pose variations of the database, where each image corresponds to frontal, upper, lower, left, and right, respectively.

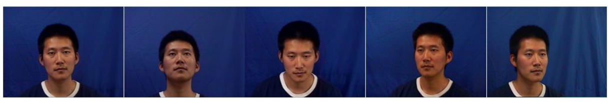


Figure 2. Example images of a subject in different poses.

2.2.2 Expression

Four kinds of facial expressions, which are neutral, happy, surprise, and anger, are considered. The data could be used for evaluating the robustness of face recognition algorithms to expressions and evaluating expression recognition algorithms. Figure 3 shows the four expression variations of the database, where each image corresponds to neutral, happy, surprise, and anger, respectively.

Figure 3. Example images of a subject with different expressions.

2.2.3 Illumination

Sixteen different illumination conditions, which consist of no light condition and 15 different light directions, are considered. The data could be used for evaluating the robustness of face recognition algorithms to illumination. Figure 4 shows the sixteen expression variations of the database.

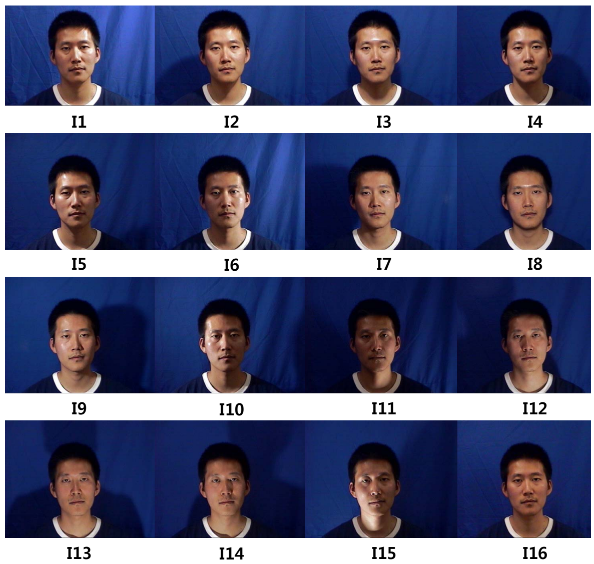


Figure 4. Example images of a subject under different illuminations.

2.3. Capturing Environment

We describe the capturing environment. First, we explain the camera configuration. For obtaining images of a subject from five poses, we used five CNB-AN202L CCD cameras capturing images simultaneously, where the angle between the frontal camera and other cameras is 22.5°. Second, we explain the illumination configuration. We occluded the windows in a small studio room by dark curtain. We used the fluorescent light lamps partially occluded by dark papers in the ceiling for an ambient light effect. For obtaining images of a subject under different illumination conditions, we used 15 fluorescent lights as a directional light source. The location of five cameras and 15 lights are shown in Fig. 5. Third, we explain capturing procedure. We asked the subject to sit in the chair and look at the frontal camera with one of four expressions (neutral, happy, surprise, anger). Then we captured images from five camera with turning on/off each light. For one expression, we were able to obtain 80 images from five cameras, where each camera captured 16 images which have different illumination. Consequently, we obtained $5 \times 4 \times 16 = 320$ images for a subject.

3. Performance Evaluation

In order to show the difficulty of the database for face recognition algorithms and provide an example evaluation protocol on the database, we evaluated some baseline face recognition algorithms on the PF07 database. We chose the principal component analysis (PCA), PCA without 3 most significant basis vectors (PCA/wo3) and linear discriminant analysis (LDA) as the baseline face recognition algorithms.

3.1. Principal Component Analysis

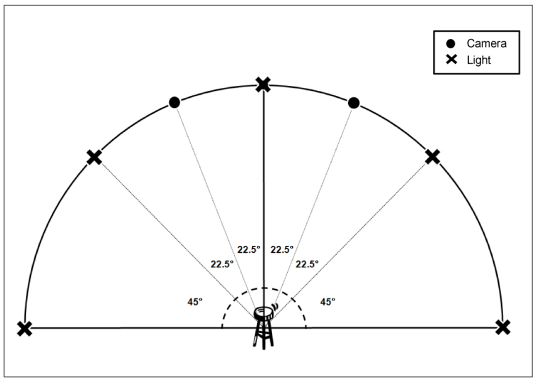
From the viewpoint of both the curse of dimensionality and the optimality of the pattern classification, it is desirable to reduce the dimensionality of feature space of the data. In PCA [4], [10], a set of observed n-dimensional data vector $\vec{X} = \{\vec{x}_p\}$, $p \in \{1, \cdots, N\}$ is reduced to a set of m-dimensional feature vector $\vec{S} = \{\vec{s}_p\}$, $p \in \{1, \cdots, N\}$ by a transformation matrix T as

$$\vec{s}_p = T^t(\vec{x}_p - \mathcal{E}[\vec{x}]), \quad (1)$$

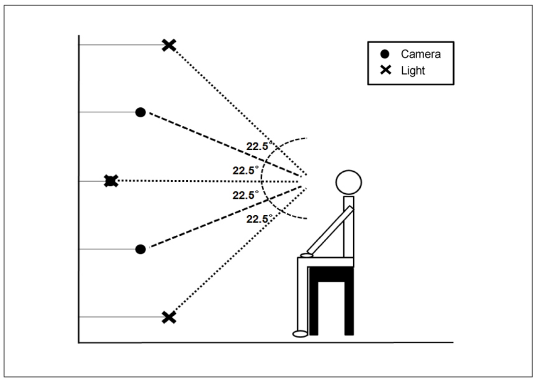
where $m \leq n$, $T = (\vec{w}_1, \cdots, \vec{w}_m)$ and the vector $\vec{w}_j$ is the eigenvector which corresponds to the jth largest eigenvalue of the sample covariance matrix $C = \frac{1}{N}\sum_{p=1}^{N}(\vec{x}_p - \mathcal{E}[\vec{x}])(\vec{x}_p - \mathcal{E}[\vec{x}])^T$, such that $C\vec{w}_k = \lambda_k\vec{w}_k$. The m principal axes T are orthonormal axes onto which the retained variance under projection is maximal. One property of PCA is that projection onto the principal subspace minimizes the squared reconstruction error $\sum \parallel \vec{x}_p - \hat{\vec{x}} \parallel^2$. The optimal linear reconstruction of $\hat{\vec{x}}$ is given by $\hat{\vec{x}} = T\vec{s}_p + \mathcal{E}[\vec{x}]$, where $\vec{s}_p = T^t(\vec{x}_t - \mathcal{E}[\vec{x}])$, and the orthogonal columns of T span the space of the principal m eigenvectors of C.

3.2. Linear Discriminant Analysis

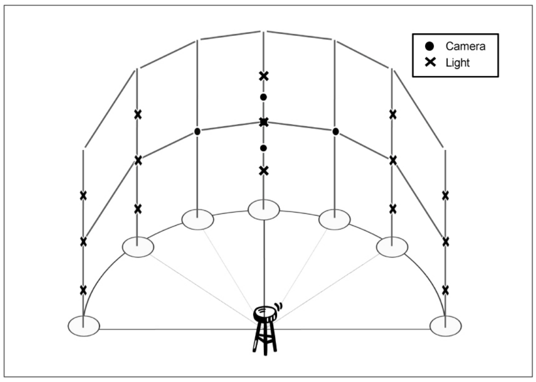
LDA is a well-known classical statistical technique using the projection which maximizes the ratio of scatter among the data of different classes to the scatter within the data of the same class [2]. The face recognition method using LDA is called the Fisherface method. It can be applied directly



(a) Upper view.



(b) Side view.



(c) Quarter view.

Figure 5. The configuration of cameras and lights in the studio room.

to a gray image [1], or to feature vectors of the gray image. Features obtained by LDA are useful for pattern classification since they make the data of the same class closer to each other, and the data of different classes further away from each other. Typically, LDA is compared to PCA because both methods are multivariate statistical techniques for projection. PCA attempts to locate the projection that reduces the dimensionality of a data set while retaining a variation in the data set as much as possible. Since PCA does not use the class information of a data set, LDA usually outperforms PCA for pattern classification. LDA is a well-known method for classification in pattern recognition and simple to understand. The procedure of LDA is based on the eigenvalue problem and gives an exact solution of the maximum of the inertia.

3.3. Data Setup

We used the normalized face images which is shown in Fig. 1 for evaluating the baseline face recognition algorithms. We divided the database into three disjoint sets: the training set, the evaluation set, and the test set. Among 200 subjects, we randomly and exclusively took 50 subjects, 50 subjects, and 100 subjects as training set, evaluation set, and test set, respectively. The training set consisting of the randomly chosen 25 male subjects and 25 female subjects was used to train the PCA and LDA and the evaluation set consisting of the randomly chosen 25 male subjects and 25 female subjects was used to determine the optimal number of basis vectors for PCA, PCA/wo3 and LDA. The test set consisting of the remaining 50 male subjects and 50 female subjects was used to evaluate the performance of face recognition algorithms. For all the following experiments, we used the frontal face images with neutral expression under normal illumination condition (I16) of all subjects in the test set as gallery images.

3.4. Evaluation Results

3.4.1 Pose

We evaluated the performance of face recognition algorithms with respect to the pose variations. We used all 5 pose images with neutral expression under normal illumination condition as probe images. Thus, there are one gallery image and 5 probe images for all subjects. Figure 6 shows the recognition rates for different facial poses. When the pose of the probe image is different from the gallery image, the recognition rate declines greatly, especially for PCA and PCA/wo3.

3.4.2 Expression

We evaluated the performance of face recognition algorithms with respect to the expression variations. We used all 4 expression images in frontal pose under normal illumination condition as probe images. Thus, there are one gallery image and 4 probe images for all subjects. Figure 7 shows the recognition rates for different facial expressions.

3.4.3 Illumination

We evaluated the performance of face recognition algorithms with respect to the illumination variations. We used

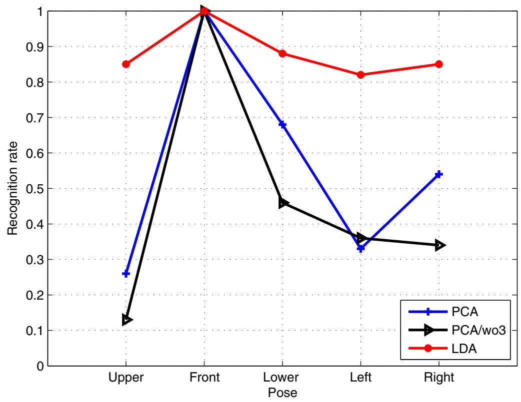


Figure 6. The recognition rates for the pose variations.

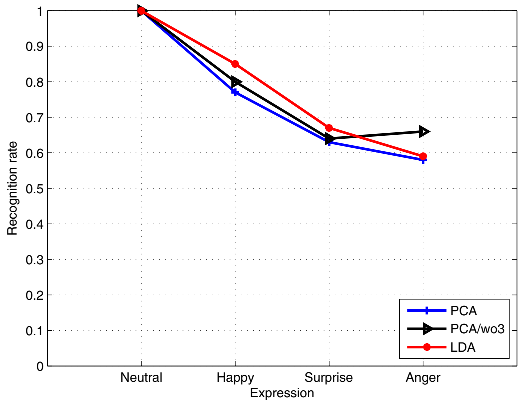


Figure 7. The recognition rates for the expression variations.

all 16 illumination images in frontal pose with neutral expression as probe images. Thus, there are one gallery image and 16 probe images for all subjects. Figure 8 shows the recognition rates for different illumination conditions. From this figure, we know that (1) the lower lights (I11 to I15) degrades the recognition rate greatly, (2) the PCA/wo3 improves the recognition rate of PCA significantly, and (3) the LDA shows the best recognition rate among three different baseline face recognition algorithms.

Figure 9 shows the average recognition rate of three baseline face recognition algorithms for different set of probe images. In the figure, each graph represents the recognition rate when the probe images have (1) pose variation, (2) expression variation, (3) illumination variation, (4) expression and illumination variations, (5) pose and expression variations, (6) pose and illumination variations, and (7) pose, expression, and illumination variations, respectively. From the figure, we know that (1) when the probe set have more variations, the recognition rate is worse, (2) LDA

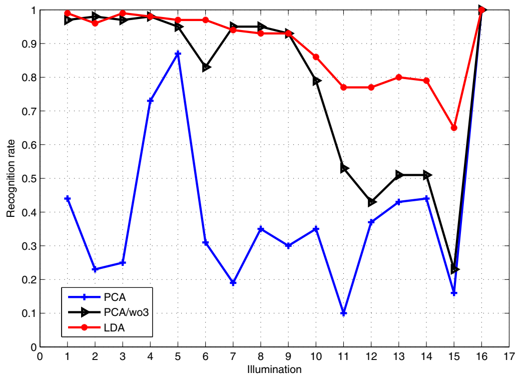


Figure 8. The recognition rates for the illumination variations.

shows the best recognition rate, and (3) when the probe set has all three variations, the recognition rate is under 50%.

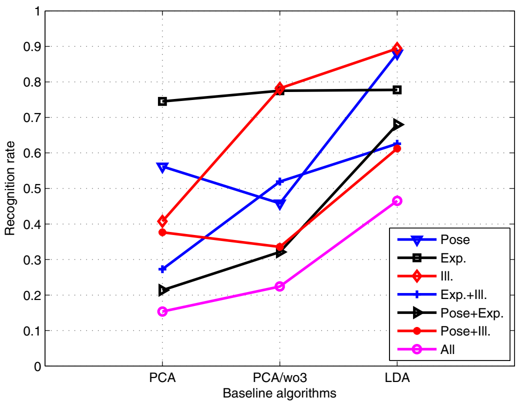


Figure 9. The recognition rates for three baseline algorithms.

4. Conclusion

In this paper, we described the PF07 database, which contains 64,000 images of 200 subjects (100 male and 100 female), where there are pose, expression, illumination variations. We also evaluated three baseline face recognition algorithms on this database. We expect the database to be used to evaluate the algorithm of face recognition for Korean people or for people with systematic variations. The database is now available for research purpose on our web site[1].

5. Acknowledgement

This resarch was supported by the Intelligent Robotics Development Program, one of the 21st Century Frontier R&D Programs funded by the Ministry of Commerce, Industry and Energy. Also it was partially supported by the Ministry of Education and Human Resources Development (MOE), the Ministry of Commerce, Industry and Energy (MOCIE) and the Ministry of Labor (MOLAB) through the fostering project of the Lab of Excellency.

[1]http://imlab.postech.ac.kr/new_imlab/faceDB/PF07/PF07.html

References

[1] P. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs.fisherfaces:class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):711–720, 1997.

[2] R. Duda, P. Hart, and P. Stork. *Pattern Classification*. Wiley, 2001.

[3] W. Gao, B. Cao, S. Shan, D. Zhou, X. Zhang, and D. Zhao. Cas-peal large-scale chinese face database and evaluation protocols. Technical report, JDL-TR-04-FR-001, Joint Research & Development Laboratory, 2004.

[4] H. Hotelling. Analysis of a compltex statistical variables into principal components. *Journal of Educational Psychology*, 24:417–441, 1933.

[5] B.-W. Hwang, H. Byun, M.-C. Roh, and S.-W. Lee. Performance evaluation of face recognition algorithms on the asian face database, KFDB. In *AVBPA*, pages 557–565, 2003.

[6] A. Martinez and A. C. Kak. Pca versus lda. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2), 2001.

[7] J. Matas, M. Hamouz, K. Jonsson, J. Kittler, Y. Li, C. Kotropoulos, A. Tefas, I. Pitas, T. Tan, H. Yan, F. Smeraldi, J. Bigun, N. Capdevielle, W. Gerstner, S. Ben-Yacoub, Y. Abdeljaoued, and E. Mayoraz. Comparison of face verification results on the xm2vts database. In *International Conference on Pattern Recognition*, pages 858–863, 2000.

[8] P. J. Phillips, H. Moon, P. J. Rauss, and S. Rizvi. The feret evaluation methodology for face recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10), 2000.

[9] T. Sim, S. Baker, and M. Bsat. The cmu pose, illumination, and expression database. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(12), 2003.

[10] M. Turk and A. Pentland. Eigenfaces for recognition, journal of cognitive neuroscience. *Journal of Cognitive Neuroscience*, 3(1), 1991.