Department of Electrical and Computer Systems Engineering

Technical Report MECSE-6-2004

A Study of the Eigenface Approach for Face Recognition

Tat-Jun Chin and David Suter

MONASH UNIVERSITY

A Study of the Eigenface Approach for Face Recognition

Tat-Jun Chin* and David Suter
Dept of ECSE, P.O. Box 35, Monash University, Clayton 3800, Australia, 2004.
{tat.chin, d.suter}@eng.monash.edu.au

Abstract: *Appearance-based approaches in face recognition, specifically the* Eigenface *approach, were one of first successful demonstrations of machine recognition of faces [1]. These methods, such as those proposed in [2, 3], proved to be effective in experiments with large databases. Further development of holistic methods of face recognition and their theoretical background, such as those proposed in [4-9], were focused towards recognizing faces from images with changes caused by illumination effects and pose variations. Although much effort has been made towards this goal, current algorithms are still far away from the capability of the human perception system [1]. This report shall give a detailed description of the fundamentals of appearance-based holistic methods for face recognition, specifically the* Eigenface *approach [3], as well as our experimental results on the Yale Face Database.*

Index terms: *Face recognition, Appearance-based holistic methods,* Eigenface *Approach Illumination effects.*

1. Introduction

Automatic visual recognition of human faces is an extremely attractive research subject. This is motivated by the wide range of commercial and law enforcement applications, as well as the desire to understand the psychophysical nature of the capabilities of face recognition in human beings. The task seems very easy and natural for biological systems, whereas current state-of-the-art face-recognition algorithms, although having reached a certain stage of maturity, is still limited to environments with strict constraints imposed. Noticeably, recognition of face images acquired in a setting with illumination changes, or pose variation of the subject, still remains a largely unsolved problem [1].

After three decades of research effort, the *Eigenface* approach [3] emerged as the first real successful demonstration of automation human face recognition. This is one of the methods which can be classified as appearance-based methods, which uses the whole face region as the raw input to a recognition system. The goal of an appearance-based face recognition algorithm is essentially to create low-dimensional representations of face images to perform recognition. In contrast, geometric feature-based methods attempt to distinguish between faces by comparing properties and relations between facial features, such as eyes, mouth, nose and chin. As a consequence, success of these methods depends heavily on the feature extraction and measurement process.

* Tat-Jun Chin was a holder of the Australia-Asia Awards conferred by the Department of Education, Science and Training (DEST) of the Government of Australia since Feb 2004.

The low-dimensional representation of faces in the *Eigenface* approach is derived by applying Principle Component Analysis (PCA) to a representative dataset of images of faces. The system functions by projecting face images onto a feature space that spans the significant variations among known face images. These significant features are termed "*Eigenfaces*" because they are the principal components of the set of training face images [3]. It should be noted that these features do not necessarily correspond to facial features such as eyes, nose and ears. They merely capture the image points that cause meaningful variations between the faces in the database that allow them to be differentiated. Face images are then classified within the low-dimensional model using a nearest-neighbor classifier.

The *Eigenface* approach works well on test images unaffected by illumination changes. It is a well-know fact that intrapersonal differences (e.g. illumination effects, poses) cause more variations between face images than interpersonal differences (identity) [10]. To handle this variability, methods usually take one of two approaches: measure some property in the image that is invariant or at least insensitive to illumination effects, or model the object in order to predict the variations caused by changes in illumination [5]. Solutions that follow the former approach are so far still elusive, and may never exist at all [11]. This suggests that appearance-based methods, which derive low-dimensional models of the face images used for training, are the only answer to this challenging problem.

The *Eigenface* approach (or other appearance-based methods) can be extended to include test images which are affected by illumination variations, provided that the faces have been recorded under similar illumination conditions. Acquiring face images to create databases that include all possible illumination conditions is unwieldy and impossible, as the space that corresponds to lighting conditions is infinite dimensional. This is a drawback of appearance-based models where, in their original form, they cannot be extrapolated to include novel viewing conditions [9].

Theoretical work along the lines of [4, 6, 12, 13] argued that the set images of objects under all possible illumination can be modeled by a low-dimensional subspace. Their conclusion was made under the assumption that the surfaces of the objects have Lambertian reflectance functions. Empirical results in [6, 13] showed that "only a small number of eigenimages are needed to approximately represent the intensity changes caused by variations in the lighting conditions". Furthermore, it was suggested that for a wide range of objects, at most 7 eigenimages suffice to capture the effects of lighting variations on images. Elaborate theoretical analyses published in [4, 12] proved that the set of all Lambertian reflectance functions obtained with arbitrary light sources lies close to a 9D linear subspace.

What remains now is a means to obtain a set of images of the same face but subjected to all possible illumination conditions: the *Eigenface* approach (or other appearance-based methods) is readily extendable to include face images with different illumination conditions to create a low-dimensional representations that models lighting changes as well, and the low-dimensional representation has been proven empirically and theoretically to be at most 9-dimensional. A distant variant of the *Eigenface* approach, the Illumination Subspace Method [5, 8, 9] proved to be effective in filling in this gap. With fundamentals rooted in photometric stereo methods, especially those along the lines of [14], this method estimates an "illumination cone" that contains "all possible images of a convex Lambertian surface created by varying the direction and strength of an arbitrary number of point light sources at infinity". The illumination cone is a generative model that uses a small number of training images to synthesize novel images under changes in lighting. The cone is then sampled and used to create a low-dimensional representation of a face image under all possible illumination conditions.

Section 2 provides an introduction of the *Eigenface* approach. Section 3 gives a detailed mathematical description of the training procedures involved in the *Eigenface* approach. Section 4 presents an illustration on how the *Eigenface* approach is used to classify a human face. Our experimental results on the Yale Face Database are inserted according to relevance in the previous sections. We present our conclusion in Section 5.

2. The Eigenface Approach

Proposed in 1991 by Turk and Pentland, this was the first genuinely successful system for automatic recognition of human faces. It was a breakaway from contemporary research trend on face recognition which focused on detecting individual features such as eyes, nose, mouth, and head outline, and defining a face model based on position and size of these features, as well as geometrical relationship between them. Despite being economical representations of a face, these methods are quite sensitive to the feature extraction and measurement process [9]. Lack of existing techniques for effective extraction and measurement of facial features presents a drawback for such methods [15].

It was argued in [3] that previous work has ignored the issue of "just what aspects of the face stimulus are important for identification". It suggested an information theory approach that decomposes face images into a small set of characteristic feature images called "*Eigenfaces*", which are nothing more but principle components of the set of training images. These components span a subspace within the image space where each different face in the training set has a unique position. The *Eigenfaces* emphasize significant discriminatory "features" that causes large variations between the faces used for training, which subsequently allows them to be differentiated. It should be noted that these features may not necessarily correspond to facial features mentioned earlier. Recognition is then performed by projecting the test image into the subspace spanned by the *Eigenfaces* and classified based on the distance of the projection from positions of known faces.

Turk and Pentland were motivated by a technique developed by Sirovich and Kirby first published in [2, 16]. Based on the Karhunen-Loève expansion (which goes by other names such as Hotelling Transform or Principle Component Analysis), Kirby and Sirovich demonstrated that "any particular face can be economically represented in terms of a best coordinate system" and the system was termed "*eigenpictures*". Eigenpictures are eigenfunctions of the averaged covariance of the ensemble of faces. In other words, they showed that in principle, a collection of face images can be approximately represented by a small set of standard pictures (the *eigenpictures*) with a small set of weights for each of the standard pictures.

2.1 Procedures of the Eigenface Approach to Face Recognition

As proposed by Turk and Pentland, the system was initialized or trained with the following operations:

1. An initial set of face images were acquired. This was the training set.
2. The *Eigenfaces* were calculated from the training set. Only M *Eigenfaces* corresponding to the M largest eigenvalues were retained. These *Eigenfaces* spanned the face space which constituted of the training set.
3. The M *Eigenface*-weights were calculated for each training image by projecting the image onto face space spanned by the *Eigenfaces*. Each face image then will be represented by M weights- an extremely compact representation.

After initialization, the following steps were performed to recognize test images:

4. The set of M weights corresponding to the test image were found by projecting the test image onto each of the *Eigenfaces*.
5. The test image was determined if it was a face at all by checking whether it was sufficiently close to the face space. This was done by comparing the distance between the test image and the face space to an arbitrary distance threshold.
6. If it was sufficiently close to the face space, compute the distance of the M weights of the test image to the M weights of each face image in the training set. A second arbitrary threshold was put in place to check whether the test image corresponded at all to any known identity in the training set.
7. If the second threshold was overcome, the test image was assigned with the identity of the face image with which it had the smallest distance.
8. (*Optional*) For a test image with a previously unknown identity, the system was retrained by adding this image to the training set.

3. The Training Procedure of Eigenface Approach

A face image, I(x,y), is a two-dimensional N by N matrix of intensity values, which are usually quantized to 8-bit values. Each x and y pair denotes a position in the image. For the purpose of exposition, it is convenient to represent the matrix of intensity values as a vector, where each row is concatenated. Now, instead of having a matrix of dimension N by N, we have a vector of dimension N^2. As an example, a typical image with size 220 by 220 pixels becomes a point in a 48400-dimensional space.

To obtain the *Eigenfaces* for a training set, it is crucial to first determine the mean vector, deviation-from-mean vectors and the covariance matrix for the particular training set. Let the images in the training set be represented by $\{T_1, T_2, T_3, \dots, T_M\}$, where each T_n is a vector of N^2-dimension. The value M is the number of images in the training set. With this representation, the mean vector is:

$$\Psi = \frac{1}{M}\sum_{n=1}^{M} T_n \tag{1}$$

The set of deviation-from-mean vectors, $\{\Phi_1, \Phi_2, \Phi_3, \dots, \Phi_M\}$ contains the individual difference of each training image from the mean vector. Kirby and Sirovich [16] refer to these vectors as *caricatures*. They are simply defined as:

$$\Phi_i = T_i - \Psi \tag{2}$$

As a concrete illustration, the training set in (Figure 1), which is a subset of the Yale Face Database, yields the average face in (Figure 2). The caricatures corresponding to the training set is displayed in (Figure 3).

As described previously, the *Eigenfaces* are the set of principal components of the training set. To obtain the eigenface description of the training set, the training images are subjected to Principal Component Analysis (PCA), which seeks a set of vectors (the principal components) which significantly describes the variations of the data. Mathematically, the principal components of the training set are the eigenvectors of the covariance matrix of the training set [17]. The covariance matrix is given by:

$$C = \frac{1}{M}\sum_{n=1}^{M} \Phi_n \Phi_n^T \tag{3}$$

It is from this matrix that we are interested in finding the set of vectors $\mathbf{u_k}$ and scalars λ_k that satisfy the relations

$$Cu_k = \lambda_k u_k \tag{4}$$

$$u_l^T u_k = \begin{cases} 1, & if\ l = k \\ 0, & if\ l \neq k \end{cases} \tag{5}$$

It is clear from (5) that the vectors $\mathbf{u_k}$ are orthonormal. Another way of representing the covariance matrix is by writing

$$A = \begin{bmatrix} \Phi_1 & \Phi_2 & \dots & \Phi_M \end{bmatrix} \tag{6}$$

$$C = \frac{1}{M} A A^T \tag{7}$$

A closer look at (7) reveals that matrix C has a dimension of N^2 by N^2, and determining N^2 eigenvectors and eigenvalues from a matrix this large (48400 by 48400 for our example) is unwieldy. Furthermore, the purpose of employing PCA in the first place is to obtain a low-dimensional representation that can succinctly describe the training set, and using N^2 eigenvectors for that will defeat the purpose. In fact, if the number of data points in the image space for which we wish to find a compact representation is less than the dimension of the image space (i.e. $M << N^2$), only M-1 eigenvectors will be meaningful.

To circumnavigate the problem, Turk and Pentland proposed the following solution. Consider the eigenvectors $\mathbf{v_i}$ of $A^T A$ such that

$$A^T A v_i = \mu_i v_i \tag{8}$$

The scalars μ_i are the corresponding eigenvalues of $\mathbf{v_i}$. Multiplying $\frac{1}{M} A$ from the left for both sides of the equation yields

$$\frac{1}{M} A A^T A v_i = \frac{1}{M} \mu_i A v_i \tag{9}$$

$$C A v_i = \frac{1}{M} \mu_i A v_i \tag{10}$$

which implies that $A\boldsymbol{v_i}$ are the eigenvectors of the covariance matrix. With this treatment, we have effectively reduced the dimension of the matrix on which we have to work on from N^2 by N^2 to M by M.

Following this method, we should first construct the matrix $L = A^T A$ of M by M dimensions and find the M eigenvectors, $\boldsymbol{v_i}$, of L. The first M eigenvectors of the covariance matrix can be obtained by finding $A\boldsymbol{vi}$, and the corresponding eigenvalues allow us to rank the eigenvectors according to their significance. As described in detail previously, these eigenvectors are termed *Eigenfaces* for our purpose, and the eight most significant *Eigenfaces* for our training set are displayed in (Figure 4).

The following question arises: *how many Eigenfaces should be used?* We obviously want to capture as much variations as possible of the training set with as fewer numbers of *Eigenfaces* as possible. For a small sized training set of about 10-20 individuals, it was found that less than 10 *Eigenfaces* were enough to account for more than 90% of the variations among the training set, i.e.

$$\frac{\sum_{i=1}^{M'} \mu_i}{\sum_{j=1}^{M} \mu_j} > 0.9 \tag{11}$$

where $M' < M << N^2$

For our particular training set, it was found that the first 8 *Eigenfaces* were sufficient. It was reported in [3] that for an ensemble of 115 images, 40 *Eigenfaces* were sufficient for a "very good description" of the training set.

Each element of the training set $\{T_1, T_2, T_3, \ldots, T_M\}$ is projected onto "face space" by the following operation

$$\omega_k = (Av_k)^T (T_i - \Psi) \tag{12}$$

$$1 \leq k \leq M', 1 \leq i \leq M$$

Therefore, for each face image in the training set, we would have a set of *M'* weights, $\Omega_i = \{\omega_1 \quad \omega_2 \quad \ldots \quad \omega_{M'}\}$, $1 \leq i \leq M$, which describes the contribution of each *Eigenface* to the face image.

4. Classifying a Face Image

With each training image represented by the set of weights, standard pattern recognition methods can be used to classify input images into known identity classes. For this case, the Euclidean distance was used as the measure for classification. Before the value can be calculated, the test image, T_P, has to be projected onto the face space as well, using equation (12), yielding the set Ω_P. The test image is assigned to the class *k* which minimizes

$$\varepsilon_{C,k}^2 = \left\| \Omega_P - \Omega_k \right\|^2 \tag{13}$$

with $1 \leq k \leq M$

Since recognition is performed by projection first, any image similar-sized can be fed into the system. Images of individuals not previously seen in the training set, as well as non-face images, can be projected onto face space, yielding the set of weights Ω_P. Hence, a competent face recognition must be able differentiate between a face image and non-face image, and if a face image is received, whether it corresponds one or none of the individuals in the training set. For this purpose, the distance between the input image and face space,

$$\varepsilon_F^2 = \left\| \Phi_P - \Phi_I \right\|^2 \tag{14}$$

with

$$\Phi_P = T_P - \Psi \tag{15}$$

$$\Phi_I = \sum_{i=1}^{M'} \omega_i (Av_i) \tag{16}$$

is proposed by Turk and Pentland to countercheck whether an input image is indeed a face image. The value of Φ_I is simply the reconstructed image (less the mean vector) of the projection of the input image onto the face space spanned by the eigenvectors.

For the evaluations whether the distances ε_C^2 and ε_F^2 are sufficiently close, two arbitrarily chosen thresholds, θ_C and θ_F, were used to define the maximum allowable distance to any face class, and the maximum allowable distance to face space.

With the treatment presented, for every input image to a trained system, we would four possible scenarios:

(1) $\varepsilon_F^2 < \theta_F$ and $\min [\varepsilon_{C,k}^2 : 1 \le k \le M] < \theta_C$
(2) $\varepsilon_F^2 < \theta_F$ and $\min [\varepsilon_{C,k}^2 : 1 \le k \le M] > \theta_C$
(3) $\varepsilon_F^2 > \theta_F$ and $\min [\varepsilon_{C,k}^2 : 1 \le k \le M] < \theta_C$
(4) $\varepsilon_F^2 > \theta_F$ and $\min [\varepsilon_{C,k}^2 : 1 \le k \le M] > \theta_C$

For the first two cases, the input image is found to be a face image. For scenario (1), the input image should be assigned the class k for which $\varepsilon_{C,k}^2$ is the minimum, whereas for case (2), it should be concluded that the input image is a face which is unknown. For the last 2 cases, the results indicate that the input image is not a face image at all. Scenario (3) is a false recognition which might have been undetected if not for the ε_F^2 measure.

For the system trained with the set in (Figure 1), there were 12 individual faces, subject01 to subject12, from different ethnicity and gender. These faces were carefully chosen to have neutral expression as well as the same lighting conditions. They were then manually centered and cropped to be of the same size. The following test sets were obtained to evaluate the effectiveness of the *Eigenface* approach:

(1) Test Set 1: Subject01 with varying expressions and with glasses.
(2) Test Set 2: Subject09 with varying expressions and with glasses.

(3) Test Set 3: Subject11 with varying expressions and with glasses.
(4) Test Set 4: Subject13, Subject14 and Subject15 with varying expressions.
(5) Test Set 5: Face and non-face images.

The first three test sets constitutes known identities whose faces were trained previously. Test set 4 were made up of individuals not seen previously in the training set, and test set 5 were generated from face as well as non-face images. Recognition results for Test Set 1 to 4 are shown in (Figure 5), (Figure 6), (Figure 7) and (Figure 8) respectively.

Test Set 5 was used to estimate the best value of ε_F^2 based on equation (14), with the results shown in (Figure 9). The results show that no satisfactory value of ε_F^2 can be established to discriminate between face and non-face images. Based on the measure of distance between input image and face space, images 3 and 4 which are face images were as distant from the face space as other non-face images, while images 17, 19 and 21 which are non-face images were too close to the face space to be confidently judged as non-face images.

From equations (14), (15) and (16), we can rewrite equation of the value of distance between input image and face space as

$$\varepsilon_F^2 = \left\| T_P - T_I \right\|^2 \tag{17}$$

with

$$T_I = \left(\sum_{i=1}^{M'} \omega_i (Av_i) + \Psi\right) \tag{18}$$

From equation (18), we can see that T_I is the reconstruction of the projection of T_P onto the first *M' Eigenfaces*. The first *M' Eigenface* were found to be able to account for more than 90% of the variations in the training set, and the reconstruction is very good approximation of T_P if the image has a position in the image space close to the subspace defined by the *Eigenfaces*. This means that as long as an input image lies near the subspace defined by the *Eigenfaces*, regardless of whether the position of the image in the image space R^L (with $L = N^2$) is close to the positions of the face images, T_P and T_I will be fairly similar, and ε_F^2 will have a small value. This causes face space distances defined by equation (14) to be close for both face and non-face images.

To visualize this problem, refer to the 2D analogy in (Figure 10). The mean-adjusted points which correspond to the mean-adjusted face-images, $\{\Phi_1, \Phi_2, \Phi_3, \ldots, \Phi_M\}$, lie near the origin. After a PCA was performed on these points, it was found that the first principal component was sufficient to capture the major variations among the points, i.e. all points can be discriminated based on their projections onto the first principal component. When two test points, P and Q, which correspond to non-face images are presented, it can be observed that their projections onto the principal component are ω_P and ω_Q respectively. The reconstruction (minus the mean) of the representation of points P and Q based on the first principal component only yields the points P' and Q'. The distance between points P and P', as well as Q and Q', is analogous to the distance measure defined by equation (14). It can be seen that although both P and Q correspond to non-face images, the values of $\varepsilon^2_{F,P}$ and $\varepsilon^2_{F,Q}$ are starkly different, and $\varepsilon^2_{F,Q}$ will be in the range of ε^2_F values for points corresponding to face images. In other words, the system will classify point Q as a face image.

A better measure for face space distance would be to use deviation from mean directly. As it was reported in [3] that "Images of faces, being similar in overall configuration, will not be randomly distributed in the huge image space…", in can be conjectured that face images are situated near the average face. Therefore, we can simply use the following as a measure of face space distance:

$$\varepsilon^2_F = \left\| T_P - \Psi \right\|^2 \tag{19}$$

Accordingly, a threshold θ_F can be established to differentiate face and non-face images. The 2D analogy of this measure is shown in (Figure 11). Based on this measure, the results for Test Set 5 are shown in (Figure 12). It can be seen that a threshold value can be easily determined for discrimination between face and non-face images.

Based on the training set in (Figure 1), the results from Test Set 1 to Test Set 5 can be used to estimate the optimal values for θ_F and θ_C. Both values are approximated to be

$$\theta_F = 3.5\text{x}10^8,\ \theta_C = 5.38\text{x}10^{15} \tag{20}$$

By using these values again, images of Subject01 to Subject12 and their variations, Subject13 to Subject15 and their variations, as well as non-face images are ran through the trained system to obtain recognition results. It was found that the system was able to produce 78% of correct results.

The results above did not take into account that different losses are associated with different classification results. For example, it is undesirable for the system to classify a face image from class i to class j for it is a catastrophic error, but it still acceptable for the system to refuse making a decision and classify a face as unknown. The values in equation (20) have been optimized for this purpose, with the loss matrix as follows:

Classification Result	Face Image Class i	Unknown Face Image	Non-face Image
Class j $j = i$ $j \neq i$	 0 1	 1 1	 1 1
Unknown Face	0.5	0	0.3
Non-face Image	0.3	0.3	0

Table 1. The loss matrix

A loss value of between 0 and 1 would mean that the classification result is not entirely wrong or right. Based on the loss matrix described above, the classification result was improved to 88% accurate.

5. Conclusions

The *Eigenface* approach performs satisfactorily for our training set of faces. Although the results were not good enough for practical purposes, much can be done to improve it. For example, a well-devised preprocessing stage on the face images can be done to obtain a training set which is consistent in terms of spatial location of facial features. In our experiments, the face images were manually cropped from their original form by approximately placing the eye levels in the middle of the image. Another pre-training process which can improve the performance is to normalize the input images to have zero mean and unit variance. In our experiments, the training images were used without much alterations.

The method was found to be robust enough to account for changes in facial expressions and addition of accessories, as our extensive experiments have shown. It was capable of classifying known faces as well as discarding unknown face images. The approach proposed by the originators to differentiate face and non-face images was unsatisfactory, and the flaw was explained in our report. More importantly, we developed a simpler but more effective method for the task. Experimental results show that the proposed method improved the results considerably.

References:

1. Zhao, W., et al., *Face Recognition: A Literature Survey.* ACM Computing Surveys, 2003. **35**(4): p. 399-458.
2. Kirby, M. and L. Sirovich, *Application of the Karhunen-Loève Procedure for the Characterization of Human Faces.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990. **12**(1): p. 103-108.
3. Turk, M. and A. Pentland, *Eigenfaces for Recognition.* Journal of Cognitive Neuroscience, 1991. **3**(1): p. 71-86.
4. Basri, R. and D.W. Jacobs, *Lambertian Reflectance and Linear Subspaces.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003. **25**(2): p. 218-233.
5. Belhumeur, P.N. and D.J. Kriegman, *What Is the Set of Images of an Object Under All Possible Illumination Conditions.* International Journal of Computer Vision, 1998. **28**(3): p. 245-260.
6. Epstein, R., P.W. Hallinan, and A.L. Yuille, *5±2 Eigenimages Suffice: An Empirical Investigation of Low-Dimensional Lighting Models.* Proceedings IEEE Workshop Physics-Based Vision, 1995: p. 108-116.
7. Epstein, R., A.L. Yuille, and P.N. Belhumeur, *Learning Object Representations from Lighting Variations.* ECCV Object Recognition Workshop, 1996.
8. Georghiades, A.S., D.J. Kriegman, and P.N. Belhumeur. *Illumination Cones for Recognition Under Variable Lighting: Faces*. in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*. 1998.
9. Georghiades, A.S., P.N. Belhumeur, and D.J. Kriegman, *From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001. **23**(6): p. 643-660.
10. Adini, Y., Y. Moses, and S. Ullman, *Face Recognition: The Problems of Compensating for Changes in Illumination Direction.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997. **19**(7): p. 721-732.
11. Moses, Y. and S. Ullman. *Limitation of Non-Model-Based Recognition Schemes*. in *ECCV 92*. 1992.
12. Basri, R. and D.W. Jacobs. *Lambertian Reflectance and Linear Subspaces*. in *International Conference on Machine Vision*. 2001.
13. Hallinan, P.W. *A Low-dimensional Representation of Human Faces for Arbitrary Lighting Conditions*. in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*. 1994.
14. Hayakawa, H., *Photometric Stereo Under a Light Source with Arbitrary Motion.* Journal of the Optical Society of America A, 1994. **11**(11): p. 3079-3089.
15. Cox, I., J. Ghosn, and P. Yianilos. *Feature-based Face Recognition Using Mixture Distance*. in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*. 1996.
16. Kirby, M. and L. Sirovich, *Low-dimensional Procedure for the Characterization of Human Faces.* Journal of the Optical Society of America A, 1987. **4**: p. 519.
17. Jolliffe, I.T., *Principal Component Analysis*. 2nd ed. Springer series in statistics. 2002: New York : Springer.

List of Figures

Figure 1. The training set

Figure 2. The average face

Figure 3. The caricatures corresponding to the training set.

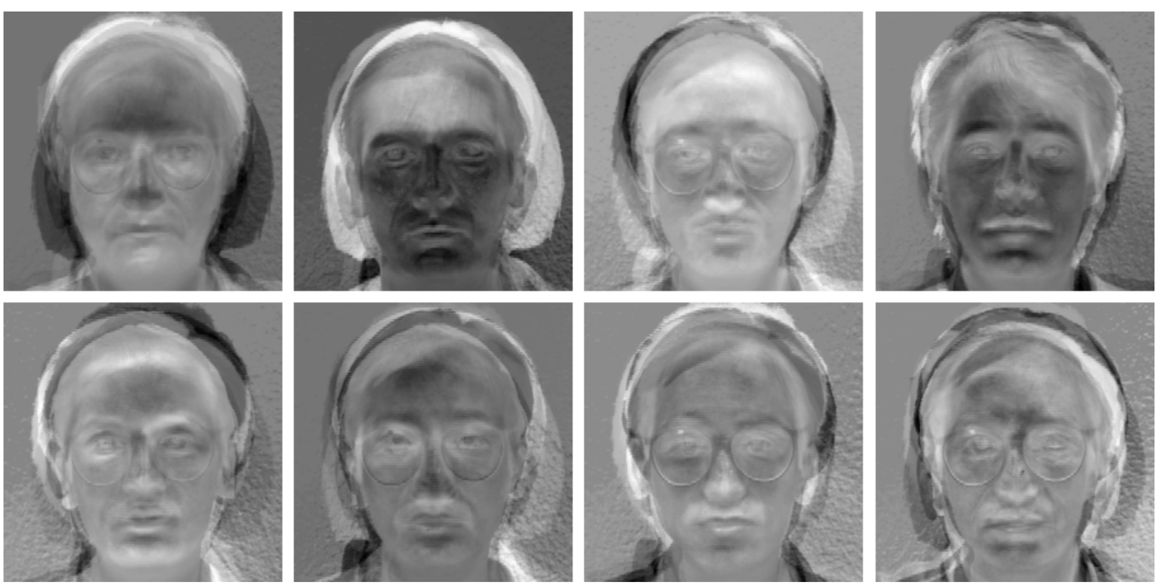

Figure 4. Eight most significant eigenface for the training set.

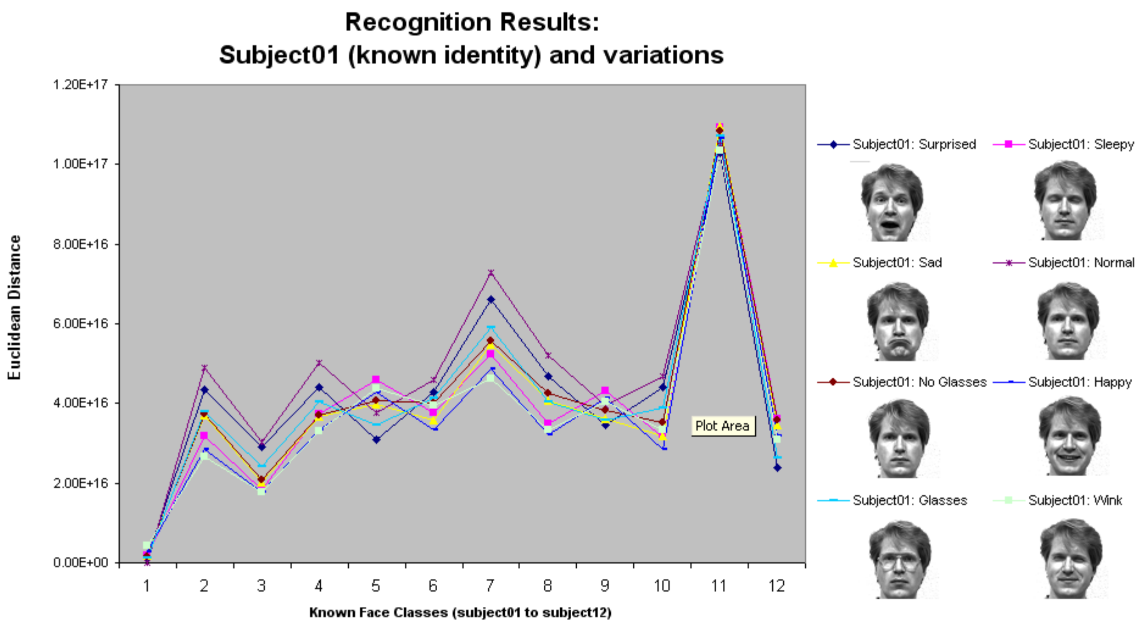

Figure 5. Recognition Result: Test Set 1

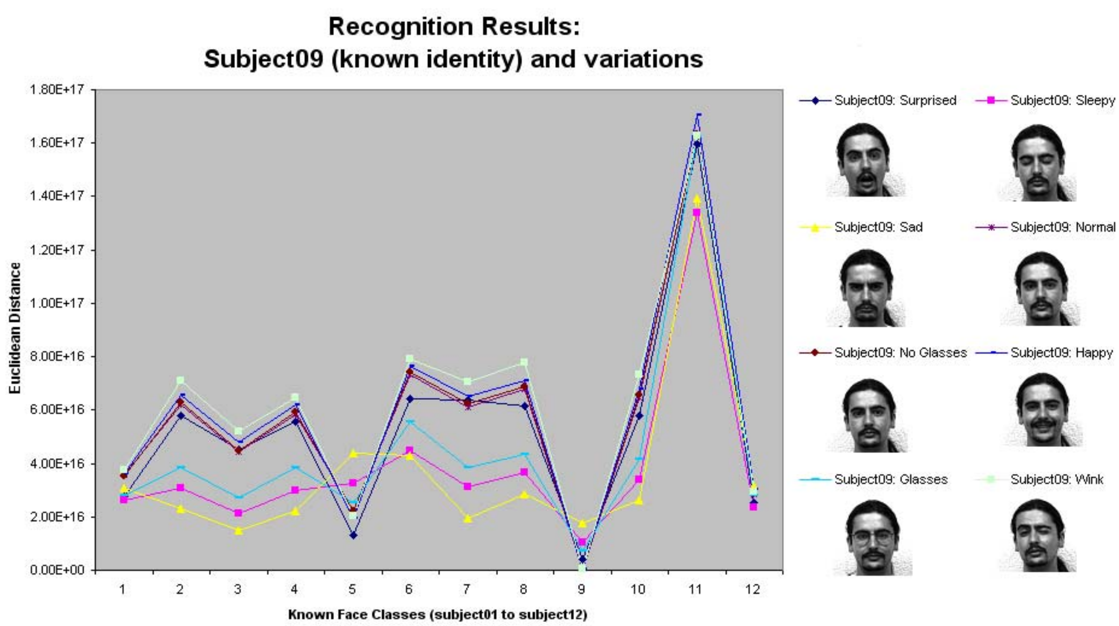

Figure 6. Recognition Result: Test Set 2

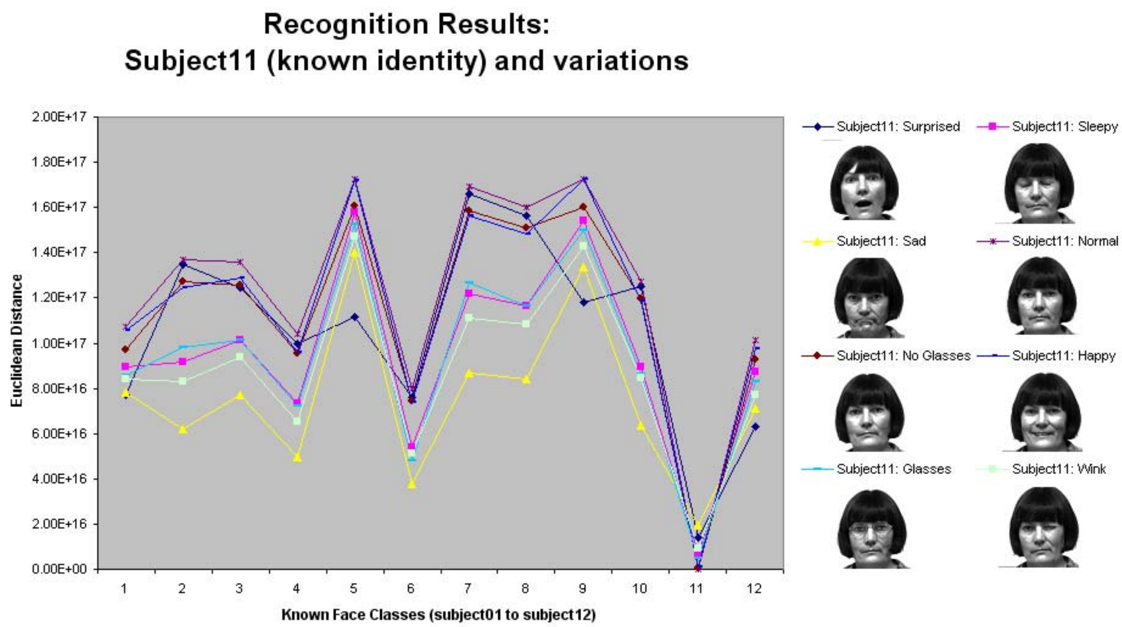

Figure 7. Recognition Result: Test Set 3

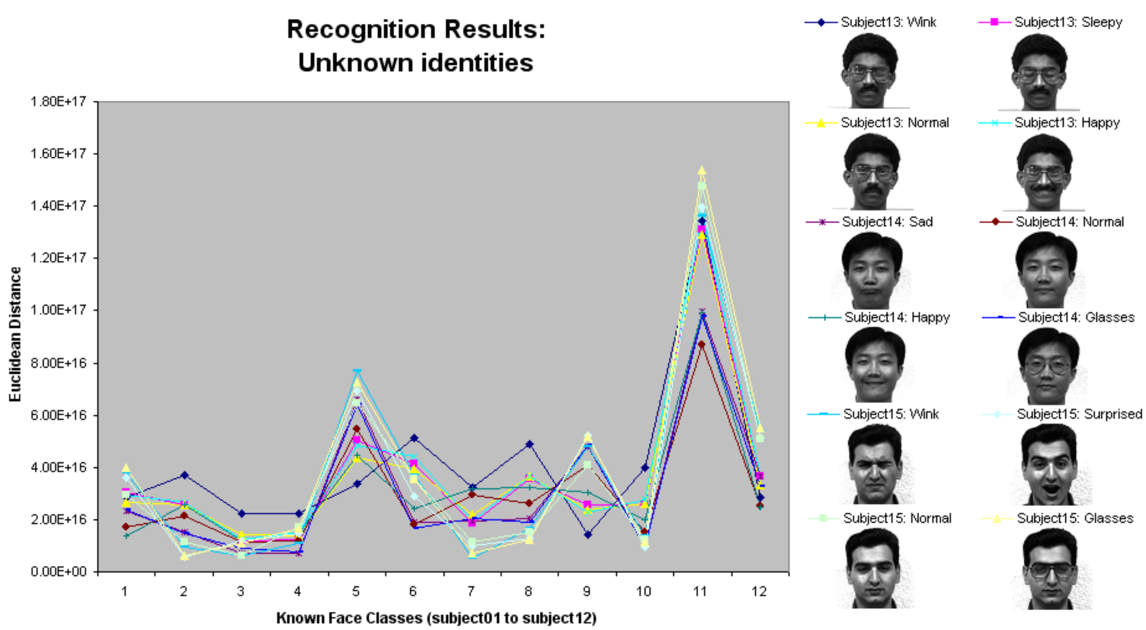

Figure 8. Recognition Result: Test Set 4

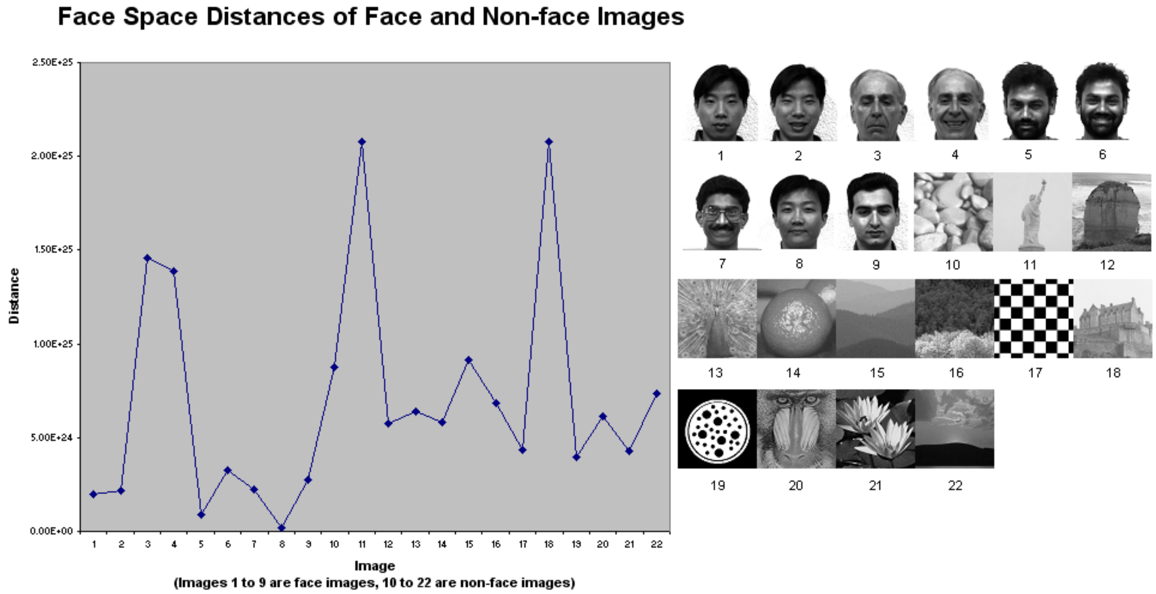

Figure 9. Result of Test Set 5: Face Space Distance based on equation (14)

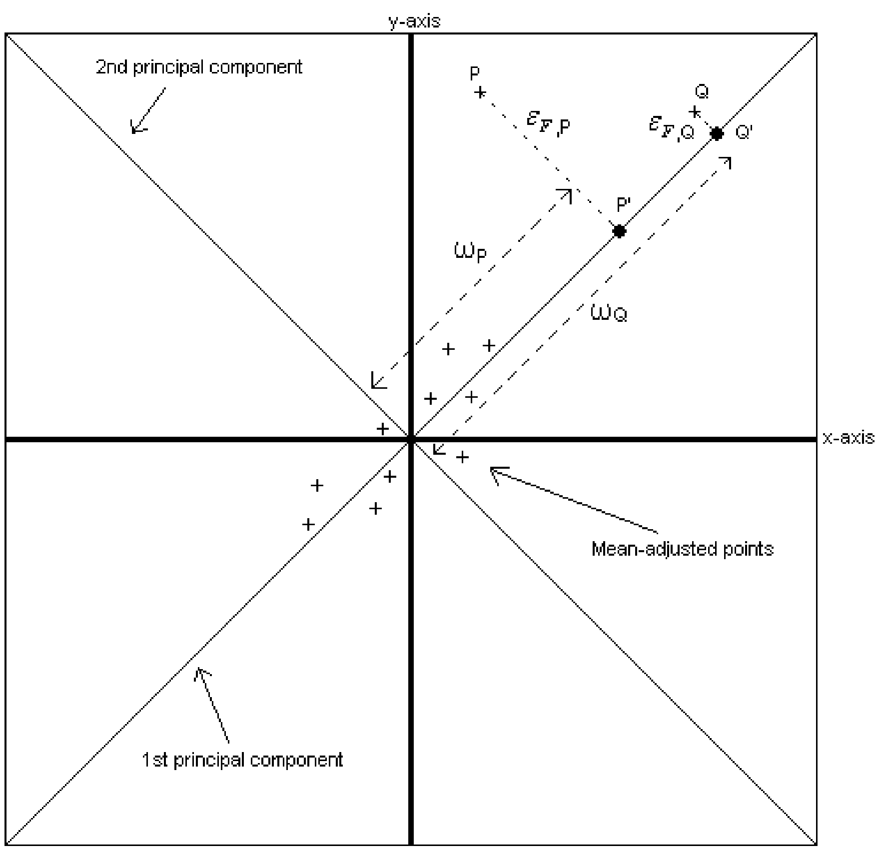

Figure 10. 2D Analogy of the face-space-distance problem

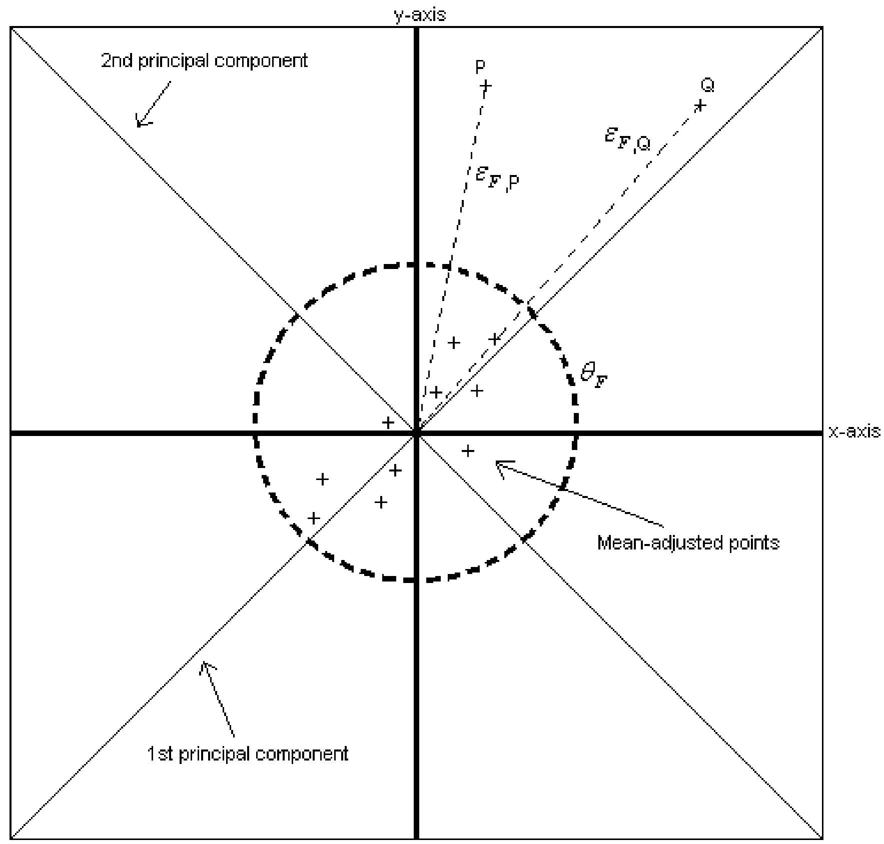

Figure 11. 2D Analogy of the face-space-distance problem: A new measure

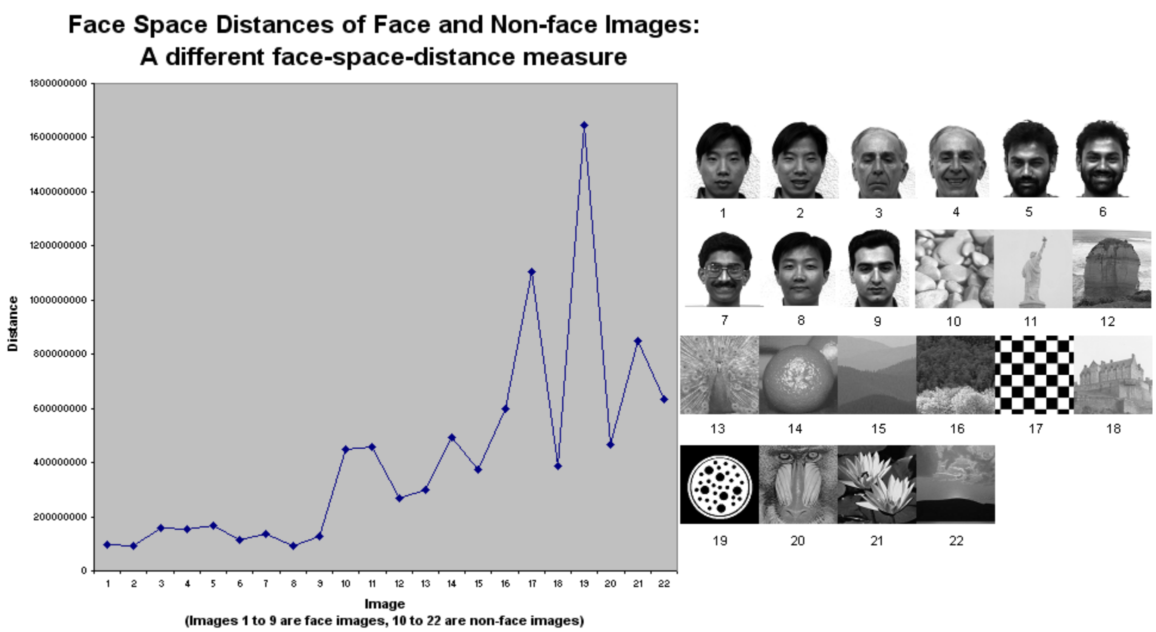

Figure 12. Result of Test Set 5: Face Space Distance based on equation (19).