Eigenfaces and Fisherfaces – A comparison of face detection techniques

Pradyumna Desale
SCPD, NVIDIA
pdesale@nvidia.com

Angelica Perez
Stanford University
pereza77@stanford.edu

Abstract

In this project we compare different subspace based techniques of face recognition. Face recognition is considered a relatively mature problem with decades of research behind it and there is a lot of interest because face recognition, in addition to having numerous practical applications such as access control, Mug shots searching, security monitoring, and surveillance system, it is a fundamental human behavior that is essential for effective communications and interactions among people.

In the literatures, face recognition problem is defined as: given static (still) or video images of a scene, identify or verify one or more persons in the scene by comparing with faces stored in a database. We focus on the classification problem which is a superset of identification problem.

1. Motivation

Face recognition is a mature problem and even though computers can not pick out suspects from thousands of people NCIS style but the ability of computers to differentiate among a small number of family member and friends is considered better than humans. Face recognition has additional applications, including human-computer interaction (HCI), identity verification, access controls etc.

Feature based face recognition methods rely on extracting processing of input image to identify and extract distinctive facial features such as the eyes, mouth, nose etc., and the geometric relationship among the facial points, thus reducing the input facial image to a vector of geometric features. Standard statistical pattern recognitions are then employed to match faces using these measurements.

The distinct disadvantage of feature based techniques is that since the extraction of feature points precedes the training and classification, the implementer has to make an arbitrary decision about which features are important and that's why we are evaluating statistical methods of face recognition in this work.

We use various methods in our two-stage face recognition systems: PCA (Principal Component Analysis), 2D PCA, and LDA (Linear Discriminant Analysis) for feature extraction and SVM (Support Vector Machines) for classification.

Specifically, we compare the methods for accuracy of face identification between different statistical methods +SVM in presence of following variations

a. Training size variation.
b. Variations in number of principal components.
c. Presence of noise in image – Gaussian, salt-and-pepper, speckle noise.
d. Variations in facial expressions of subjects
e. Variations in angle and structural changes to face such as beard, glasses.
f. Variations in illumination.

Accuracy can be measured for identification or for classification of a person against training set. We use the latter as metric of accuracy.

2. Technical solution

We describe PCA, 2D PCA, LDA, and SVM briefly in this section. The mathematics behind all these techniques is proven so we don't include any derivations. References can be reviewed for mathematical proofs.

2.1. PCA

Sirovich and Kirby were the first to utilize Principal Components Analysis (PCA) to economically represent face images. They argued that any face image can be reconstructed approximately as a weighted sum of a small collection of images that define a facial basis (eigenimages), and a mean image of the face. Turk and Pentland presented the well-known Eigenfaces method for face recognition.

Suppose there are M training face images for each of K subjects. Let each face image **A**(x, y) be a 2-dimensional N-by-N array of pixel values. The image may also be represented as a vector of dimension N^2. Let's denote each face image of the training set as $\mathbf{f}_{ij}$, and corresponding average of all training images as **g**. Principal components per image are eigen values of $\mathbf{w}_i$ = eigen values of covariance matrix of $(\mathbf{f}_{ij} - \mathbf{g})$.

Principal components transform input image to lower dimensional feature vector

$$\mathbf{y}_k = \mathbf{w}_i^T * \mathbf{f}_{ij}$$

2.2. 2D PCA

In the PCA-based face recognition technique, the 2D face image matrices must be previously transformed into 1D image vectors. The resulting image vectors of faces usually lead to a high dimensional image vector space, where it is difficult to evaluate the covariance matrix accurately due to its large size and the relatively small number of training samples. As opposed to conventional PCA, 2DPCA is based on 2D matrices rather than 1D vectors.

Image covariance or scatter matrix is computed from M training images as

$$\mathbf{G} = \frac{1}{M}\sum_{1}^{M}\left(A_j - \acute{A}\right)^T\left(A_j - \acute{A}\right)$$

Eigenvectors of scatter matrix **G** corresponding to the d largest eigen values are the 2D principal components of image **A**j, If we denote the optimal projection vectors as $\{\mathbf{X}_1^{(j)}, \mathbf{X}_2^{(j)}, \ldots, \mathbf{X}_d^{(j)}\}$, corresponding features of image are computed as $\mathbf{Y}_k = \mathbf{A} * \mathbf{X}_k$. The matrix of features is called the feature matrix and $\mathbf{B} = [\mathbf{Y}_{1,} \mathbf{Y}_{2, ..} \mathbf{Y}_d]$.

2.3. Segmented PCA

The PCA based face recognition method is not very effective under the conditions of varying pose and illumination, since it considers the global information of each face image and represents them with a set of principal components. Under the variation of pose and illumination, statistical features will vary considerably from the weight vectors of the images with normal pose and illumination; hence it is difficult to identify them correctly. On the other hand if the face images were divided into smaller regions and the weight vectors are computed for each of these regions, the weights will be more representative of the local information of the face. When there is a variation in the pose or illumination, only some of the face regions will vary and rest of the regions will remain the same as the face regions of a normal image. Hence weights of the face regions are not affected by varying pose and illumination will closely match with the weights of the same individuals face regions under normal conditions. We implemented segmented PCA variations of 1D and 2D PCA methods that assign equal weight to all the sub images.

2.4. LDA Fisher's Linear Discriminant Analysis

PCA methods reduce the dimension of input data by a linear projection that maximizes the scatter of all projected samples. Fisher's Linear Discriminant (FLD) shapes the scatter with the aim to make it more suitable for classification. A computation of the transform matrix results in maximization of the ratio of the between-class scatter.

In choosing the projection which maximizes total scatter, PCA retains some of the unwanted variations due to lighting and facial expression. The variations between images of the same face due to illumination and viewing direction are almost always larger than image variations due to change in face identity. Thus while PCA projections are optimal for reconstruction from a low dimension basis, they may not be optimal from a discrimination standpoint.

In this project we implement a variant of LDA called D-LDA. The basic premise behind the D-LDA approaches is that the information residing in (or close to) the null space of the within-class scatter matrix is more significant for discriminant tasks than the information out of (or far away from) the null space. Generally, the null space of a matrix is determined by its zero eigenvalues. However, due to insufficient training samples, it is very difficult to identify the true null eigenvalues. As a result, high variance is often introduced in the estimation for the zero (or very small) eigenvalues of the within-class scatter matrix. Note that the eigenvectors corresponding to these eigenvalues are considered to be the most significant feature bases in the D-LDA approaches.

2.5. Support Vector Machines

The goal of SVM classifiers is to find a hyperplane that separates the largest fraction of a labeled data set $\{(x^{(i)}, y^{(i)}); x(i) \in R^N; y^{(i)} \in \{-1,1\}; i=1,2\ldots,N\}$. The most important requirement, which the classifiers must have, is that they must maximize the distance or the margin between each class and the hyperplane. In most of real applications, the data cannot be linearly classified. To deal with this problem, we transform data into a higher dimensional feature space and assume that our data in this space can be linearly classified.

The discriminant hyperplane is defined as the following

$y(\mathbf{x}) = \sum_{1}^{N} \alpha y^{(i)} K\left(x^{(i)}, x\right) + b$ where K is the kernel function for SVM. In this paper we use a radial basis function kernel for SVM classification.

2.6. Algorithm description

PCA or LDA method is used to identify features of training images. To apply SVM for classification, we use one-against-all decomposition to transform a multi-class problem into a two-class problem.

Training set D= $\{(x(i), y(i)); x(i) \in R^N; y(i) \in \{1,2,3,\ldots,K\}$; is transformed into a series of $D_k = \{(x(i), yk(i)); x(i) \in R^N; y(i) \in \{k, 0\}$.

We use the MATLAB function svmtrain to compute discriminant function of SVM and in the classification phase, we use the following rule with MATLAB function svmclassify to identify the class of test probe image x.

3. Experimental Setup

We use two databases for our experiments

- AT&T/ ORL face database
- Yale face database

The AT&T Face database, sometimes also known as ORL Database of Faces, contains ten different images of each of 40 distinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement).

The AT&T Face database is good for initial tests, but it's a fairly easy database. The Eigen faces method already has a 90+% recognition rate, so we didn't expect to see considerable improvements with other algorithms. The Yale Face database A is a more appropriate dataset for our experiments, because the recognition problem is harder. The database consists of 15 people (14 male, 1 female) each with 11 grayscale images sized 320 × 243 pixel. There are changes in the lighting conditions (center light, left light, right light), facial expressions (happy, normal, sad, sleepy, surprised, wink) and glasses (glasses, no-glasses).

We select first M images of each subject for computing features of that subject class. These features are then used for training SVM with RBF kernel and the images not used for training are used for classification testing.

Graph 1: PCA method accuracy vs number of training images

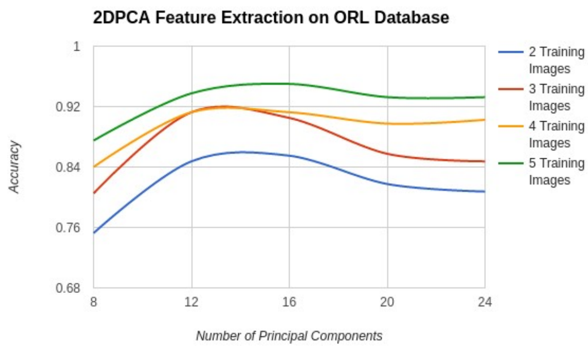
4. Results

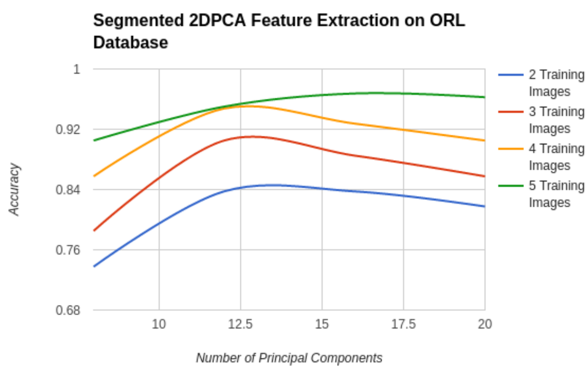
In this section we present and discuss a comparison of each of the previously mentioned feature extraction techniques with SVM classifiers.

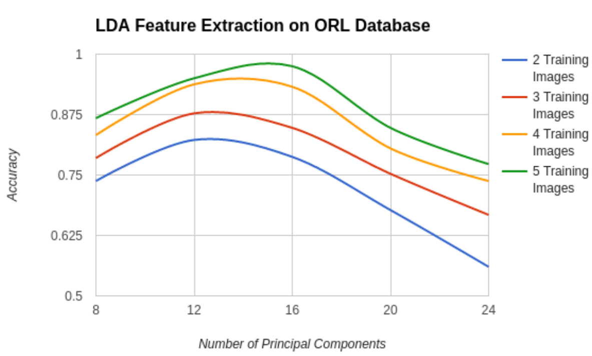
4.1. Effect of training size variations and number of principal components variations

Training size significantly affects the accuracy of all methods but there is a ceiling on the highest level of accuracy obtained from any of the methods. ORL database of faces achieves 90% accuracy within 3 training images for 2D PCA and LDA methods achieve 95% accuracy with 4 training images as shown in the graphs below.

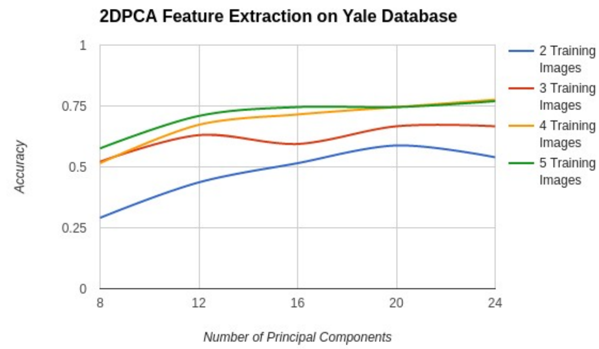
Graphs 2,3,4: Graph 2:2D PCA against number of principal components for various training sets from ORL
Graph 3: Segmented PCA against number of principal components for various training sets ORL
Graph 4: LDA against number of principal components for various training sets ORL

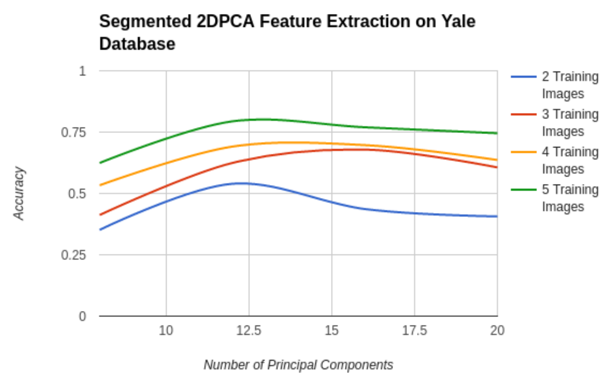


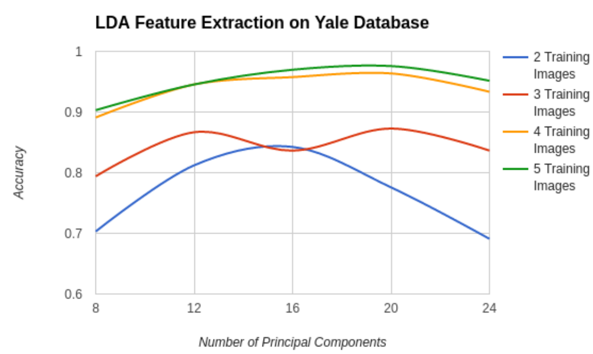




Graphs 5,6,7: Graph 5:2D PCA against number of principal components for various training sets from Yale
Graph 6: Segmented PCA against number of principal components for various training sets Yale
Graph 7: LDA against number of principal components for various training sets Yale







It is hard to compare the accuracy of 2DPCA and LDA methods with just ORL database but Yale database (graphs above) clearly shows that LDA method is superior to 2D PCA when in class variation is large.

Statistical learning methods PCA and LDA-based ones seem to perform very well with small sample and test size but often suffer from the so-called “small sample size” (SSS) problem if number of test samples is very large.

Looking closely at the inaccurate face detection in Yale database provides us insight into this problem. When first 4 images of each subject are picked for training, only subject 15 and 3 has dark glasses while other subjects don't have glasses. When test images are run against such training set, all the probe images of subjects with glasses are categorized into either class 3 or class 15. Since all subject images with glasses were not used for training, our training model is heavily biased towards two classes. That being said, increasing the number of training images improves the accuracy and we get around the bias.

4.2. Variations in facial expression of subjects, structural changes to the faces, and variations in pose

Yale database also includes facial expressions of subjects and graphs 5,6,7 show that LDA method is more immune to variations in facial expressions but PCA method's accuracy is not substantially lower than LDA. LDA method is definitely superior when

Overall we are very surprised by how resilient both sets of algorithms are to significant changes in facial poses.

4.3. Presence of noise in probe image

Noise and distortions in face images can seriously affect the performance of face recognition systems. Analog or digital capturing the image has come a long way and very good quality photo captures are possible even with cell phone camera but biometric system will need to be resilient to tampering so we explore noise immunity of different algorithms now. Noise in probe image degrades the performance of all algorithms substantially. In this section we only evaluate the 2D PCA and LDA based methods since 1D PCA has lower accuracy compared to 2DPCA and segmented PCA methods are essentially just PCA methods so the effect of noise on their performance is easily studied by understanding the effect of noise on PCA methods only. We also restrict the description of results to only Yale database even though ORL database was also studied and we found the effects of noise had less dependency on the database.

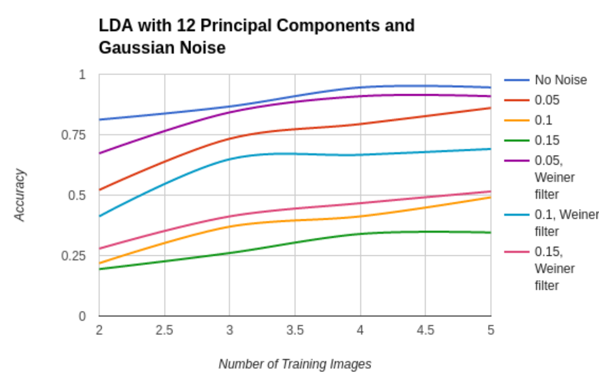
4.3.1 Gaussian noise

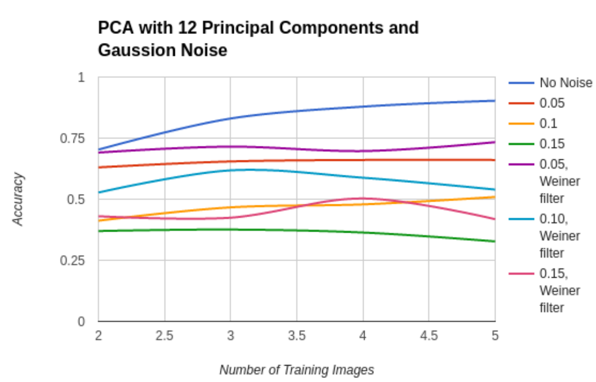
Gaussian noise is the most common noise occurring in everyday life. We use zero mean Gaussian noise of different variance in the experiment.

We also use the Wiener filter which is a MSE-optimal stationary linear filter for suppressing the degradation caused by additive noise and blurring. Fourier transforms are unable to recover components for which Fourier transform of point spread function is 0. This means they are unable to undo blurring caused by band limiting of Fourier transform. We can see that some of the accuracy of face recognition is recovered when Weiner filters are used for correction of additive noise in Graphs 8 and 9.

Weiner filter noise suppressed images look like below

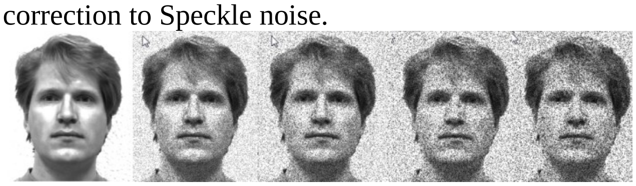
Graphs 8, 9:



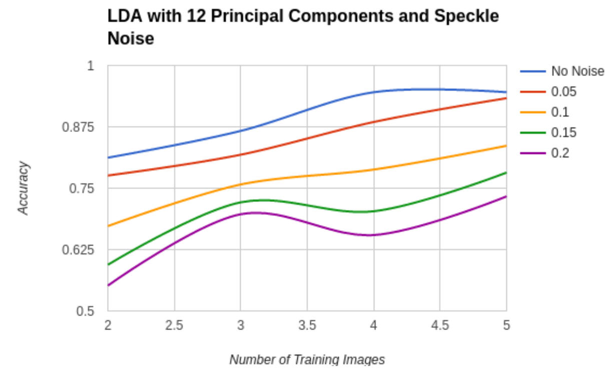


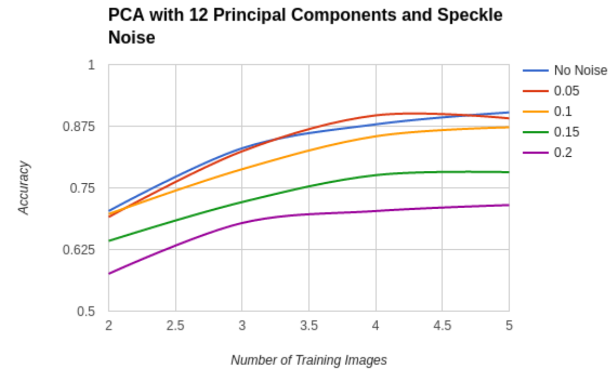
4.3.2 Speckle Noise

This granular noise occurs in ultrasound, radar and X-ray images and images obtained from the magnetic resonance. The multiplicative signal dependent noise is generated by constructive and destructive interference of detected signals. The wave interference is a reason of multiplicative noise occurrence in the scanned image. The speckle noise is image dependent. Therefore it is considered hard to find a mathematical model that describes the removal of this noise, especially if we expect the randomness of the input data Structural changes to the face image such as beard, glasses etc. We had identified Lee's filter as the method to counter the Speckle noise but speckle noise is not greatly affecting the recognition performance so we prioritized the study of Gaussian and S&P noise over correction to Speckle noise.



Graphs 10, 11: Additive speckle noise effect on accuracy for PCA and LDA.



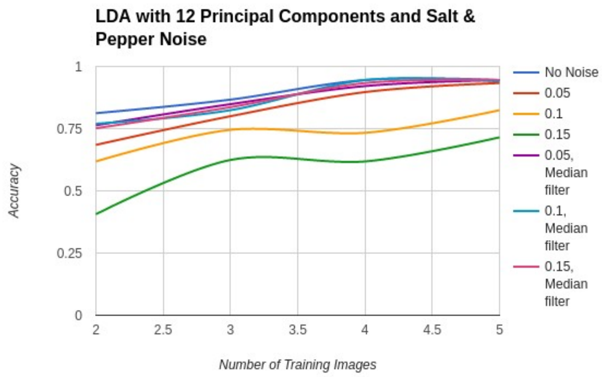


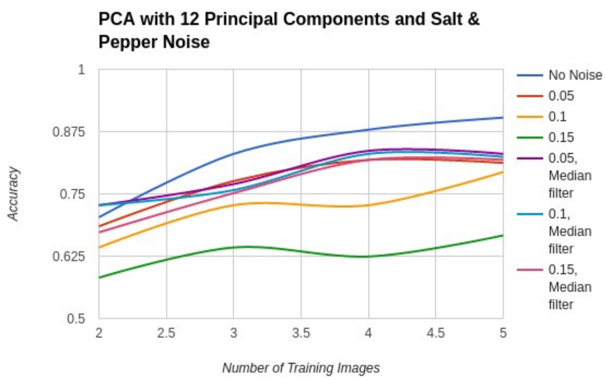
4.3.3 Salt & Pepper Noise

Salt & pepper noise is perceived as a random occurrence of black and white pixels in a digital image.

It can be caused by incorrect data transmission or by a damage of already received data. In CCD and CMOS sensors or LCD displays, the salt & pepper noise can be caused by permanently turned-on or turned-off pixels. Remaining pixels are unchanged. Usually, the intensity (frequency of the occurrence) of this noise is quantified as a percentage of incorrect pixels. The median filtering (as a specific case of order-statistic filtering) is used as an effective method for elimination of salt & pepper noise from digital images. We can see that almost all of the algorithm performance is recovered by use of median filters.

Graphs 12, 13: Additive S&P noise effect on accuracy for PCA and LDA and corrected with Median filter





https://pereza77@bitbucket.org/pereza77/cs231a_final_project_face_recognition.git

5. Summary

We examined different subspace methods of face recognition in this project. Two-stage recognition systems include PCA, LDA for feature extraction followed by SVM for classification. All methods are significantly influenced by different settings of parameters that are related to the algorithm used (i.e. PCA, LDA or SVM).

For methods working in ideal condition both PCA and LDA achieve greater than 90% accuracy within three training images.

This project dealt with 'closed' image set, so we did not have to deal with issues like detecting people who are not in the training set. On the other hand our two test databases contain images of the same subjects that often differ in face expressions, hairstyles, with or without beard, or wearing glasses and that were taken in different sessions after longer time periods. We also presented recognition results for noisy images and compared them to results for non-distorted images with correction for two types of noises.

We started this project with the intent of implementing face recognition algorithms with SVM and definitely succeeded in that goal. Computer vision and analytics systems performance is far superior when combined with deep learning models CNNs, combining deep-learning and multivariate wide-learning together with improved feature descriptor models can enable extraction of more information from facial images such as facial expression. During the research of this project, we have stumbled upon papers that study separation of a true smile from fake smile using PCAs for feature recognition and CNNs for training and classification. We would like to build fundamentals of CNNs and machine learning and combine the power of statistical models used in this project with deep learning methods to create more interesting projects like machine recognition of facial features.

References

[1] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in Computer Vision and Pattern

[2] Recognition, 1991. Proceedings CVPR '91., IEEE Computer Society Conference on, 1991, pp. 586-591.

[3] Y. Jian, et al., "Two-dimensional PCA: a new approach to appearance-based face representation and recognition," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 26, pp. 131-137, 2004.

[4] A new LDA-based face recognition system which can solve the small sample size problem. Pattern Recognition 33, 1713–1726.

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

[6] B. Moghaddam and Y. Ming-Hsuan, "Gender classification with support vector machines," in Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on, 2000, pp. 306-311

[7] P. Phillips, et al., "The FERET evaluation methodology for face-recognition algorithms," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 22, pp. 1090-1104, 2002.

[8] P. Viola and M. J. Jones, "Robust real-time face detection," International Journal of Computer Vision, vol. 57, pp. 137-154, 2004.

[9] T. H. Le and L. Bui, "Face Recognition Based on SVM and 2DPCA," International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 4, No. 3, September, 2011

[10] MATLAB.com

[11] Brunelli, R., and Poggio, T. Face recognition through geometrical features. In European Conference on Computer Vision (ECCV) (1992), pp. 792–800

[12] Kanade, T. Picture processing system by computer complex and recognition of human faces. PhD thesis, Kyoto University, November 1973.

[13] H. Yu and J. Yang, "A Direct LDA Algorithm for High-dimensional Data with Application to Face Recognition," Pattern Recognition, Vol.34, pp.2067-2070, 2001.

[14] F. Song, D. Zhang, J. Wang, H. Liu, and Q. Tao, "A parameterized direct LDA and its application to face recognition," Neurocomputing, Vol.71, pp.191-196, 2007.

[15] ORL face database - http://www.uk.research.att.com/facedatabase.html

[16] Yale face database - http://cvc.yale.edu/projects/yalefaces/yalefaces.html