İSTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY

FACE RECOGNITION UNDER VARYING ILLUMINATION

Master Thesis by
Erald VUÇINI, B.Sc.

Department : Computer Engineering

Programme: Computer Engineering

Supervisor: Prof. Dr. Muhittin GÖKMEN

JUNE 2006

PREFACE

This study is a result of a successful joint-venture with my adviser Prof. Dr. Muhittin Gökmen. I am thankful to him for his continuous assistance on preparing this project. Special thanks to the assistants of the Computer Vision Laboratory for their steady support and help in many topics related with the project.

May, 2006 Erald VUÇINI

CONTENTS

ABBREVIATIONS

SSS	: Small Sample Size
PCA	: Principal Component Analysis
LDA	: Linear Discriminant Analysis
FRS	: Face Recognition System
PC	: Principal Components
FLD	: Fisher's Linear Discriminant
D-LDA	: Direct Linear Discriminant Analysis
BHE	: Block-based Histogram Equalization
HE	: Histogram Equalization
AHE	: Adaptive Histogram Equalization

TABLE LIST

FIGURE LIST

SYMBOL LIST

w	**:** width of image
h	**:** height of image
W	**:** projection matrix
μ	**:** arithmetic average of face images
$\mathbf{S_T}$	**:** covariance matrix
$\mathbf{W_{opt}}$	**:** optimal projection matrix
Φ	**:** eigenvector matrix
Λ	**:** eigenvalues matrix
ρ	**:** albedo (surface texture)
s	: light source direction
Q	: Quotient Image
F	: Gaussian Filter
B	: Blurred image with a Gaussian Filter
ε	: error
m	: number of selected eigenvectors
R	: radiance map
X	: matrix of input image
Y	: projection of X in feature space

ÖZET

Günümüzde, görüntü işleme ve bilgisayarla görü alanlarındaki araştırmacılar, insan yüzlerinin makineler tarafından tanınması konusunda oldukça yoğun çalışmalar yapmaktadır. Buna rağmen özellikle gerçek zamanlı işlemin gerekli olduğu durumlarda otomatik bir tanıma sistemin gerçeklenmesi hala zorluklar içermektedir. Bu zorlukların nedenleri arasında yüz görüntülerinin yüzün 3boyutlu geometrisi dolayısıyla aydınlatma ve poz değişimleri olduğunda büyük değişimler göstermesi ve yüz ifadesi, yaşlanma, sakal, bıyık, gözlük ve makyajdan oluşan değişimler yer almaktadır. Aynı yüzün farklı ışıklandırma ve görüş açısı durumlarında gösterdiği değişim farklı insan yüzlerinin arasında görülen değişimden daha fazla olduğu söylenebilir. Özellikle değişken aydınlatma koşulları, yüz tanımayı günümüzdeki en zor problemlerden birisi haline getirmiştir.

Bu çalışma üç anahat üzerinde kurulmuştur. Sözü edilen hatlar yüz tanımadaki aydınlanma değişimlerle ilgili üç sorunu aşmak üzere geliştirilmiştir:

- Temel Bileşen Analizi ve Doğrusal Ayrışım Analizi kullanıldığında değişken aydınlatma koşullarının yarattığı sorunlarının giderilmesi
- Küçük örneklem probleminden kaynaklanan yüz nitelik çıkarma sorununun çözülmesi
- Herhangi bir ışıklandırma koşulunda alınmış görüntülerin önden aydınlatılmış şekilde tekrar oluşturulması

Bu çalışma boyunca alınan deneysel sonuçlar, bu üç yaklaşımın birleşiminden oluşan yöntemin, var olan yöntemlerinin başarımını geçtiğini göstermektedir. Aynı zamanda yöntem, aydınlanma değişimlerine dayanıklı gürbüz ve gerçek zamanlı bir yüz tanıma sistemi oluşturmaya olanak vermektedir.

SUMMARY

Face Recognition, being one of the most active areas of image analysis and understanding, has received special attention during the past several years. Researchers in psychology, neural sciences and engineering, image processing and computer vision have been investigating a number of issues related to face recognition by human beings and machines. Despite this it is still difficult to design an automatic system, especially when real time identification is required. The reasons for this difficulty are two-fold: 1) Face images are highly variable and 2) Sources of variability include individual appearance, three dimensional pose, facial expression, facial hair, make-up, and so on and these factor change from time to time. It has been observed that, "The variations of the same face due to illumination and viewing direction are almost always larger than image variations due to changes in the face identity". All these problems make face recognition and especially dealing with variable illumination on human face, one of the most challenging problems for robust face recognition systems.

This study is based in three main axes which have been developed in an effort to overpass three problems of Face Recognition especially dealing with varying illumination:

- Difficulty of Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) with varying illumination
- Difficulty of feature extraction due to the "small sample size " (SSS) problem
- Reconstruction of images illuminated by an arbitrary angle in order to have a frontal illumination

During this study, experimental results have shown that the combination of these three approaches outperforms existing ones and in the same time offers the possibility of creating o robust real time Face Recognition system when dealing with illumination.

1. INTRODUCTION

Face Recognition Systems (FRS) have been receiving recently special attention. This can be evidenced by the emergence of Face Recognition conferences, systematic empirical evaluations of face recognition techniques, protocols, and many commercially available systems. This introductory section offers a general view of the applications, the configuration of a generic FRS, and finally a background on the different approaches related with FRS under varying illumination.

1.1 Application Areas

The problem of machine recognition of human faces is attracting more and more researchers not only from areas as image processing, pattern recognition and neural networks but also from neuroscience and psychology. There is a strong need to secure our possessions and in the same time to protect our privacy. An important point is that while the other biometric identification methods such as fingerprint analysis or retinal scans rely on the cooperation of the participants, a FRS is often effective without the participant's cooperation or knowledge. Table 1.1 lists some of the applications of face recognition.

Table 1.1 Applications of Face Recognition

Areas	**Specific Application**
Entertainment	Video game, virtual reality, training programs
	Human-robot-interaction, human-computer-interaction
Smart Cards	Drivers' licenses, entitlement programs
	Immigration, national ID, passports, voter registration
Information Security	TV Parental control, personal device logon, desktop logon
	Application security, database security, Intranet security, internet access
	Witness face reconstruction, Electronic mug shots
Surveillance	Advanced video surveillance, CCTV control
	Portal control, post-event analysis
	Shoplifting, suspect tracking and investigation

Commercial and surveillance applications of FRS range from controlled still images to uncontrolled video images, posing a wide range of technical challenges and requiring an equally wide range of techniques from image processing, analysis, understanding, and pattern recognition. Generally, FRS can be divided into two groups depending on whether they make use of static images or video. Even within these groups, significant differences exist, depending on the specific application. The differences are in terms of image quality, amount of background clutter (posing challenges to segmentation algorithms), variability of the images of a particular individual that must be recognized, availability of a well-defined recognition or matching criterion, and the nature, type, and amount of input from a user.

In this study we will deal with still images received from a digital camera such as an IP camera, webcam and so on.

1.2 Configuration of a Generic FRS

A general statement of the problem of face recognition can be formulated as follows: given still or video images of a scene, match or verify one or more persons in the scene using a stored database of faces (learning or training set). Available extra information such as race, age, gender or facial expressions can be used as factors to speed the searching process. A general configuration of a FRS can be formulated over the following steps:

- Face Detection (Face Tracking, Pose Estimation, Compression)
- Feature Extraction (Facial Feature Tracking, Emotion Recognition, Gaze Estimation)
- Face Recognition (Holistic Templates, Feature Geometry, Hybrid)

In identification problems, the input is unknown face, and the system reports back the determined identity matched with a database of known individuals. On the other side, in verification problems the system needs to accept or reject the claimed identity of the input face. It is clear that the second case is the most problematic because we have to deal also with the rejection conditions of an input face.

In this study we will deal will identification problems, with intention that in future works the matching criteria can be expanded in a way that can allow rejection.

1.3 Face Recognition Research Survey

Face perception is an important part of the capability of human perception system and is a routine task for humans, while building a similar computer system is still an on-going research area. Research on automatic machine recognition of faces started with Kelly' work [1] in 1970, and after the seminal work done by Kanade [2] in 1973. As for the early researches, both geometric feature based methods and template-matching methods were regarded as typical technologies. The experimental results revealed that template matching outperforms the geometric feature based methods.

Therefore since 1990s, appearance based methods have been playing a dominant role in the area, from which a number of approaches were derived: holistic appearance feature based, and analytical local feature based. Popular methods belonging to the first paradigm include Eigenface [3], Fisherface [4], Probabilistic and Bayesian matching [5], and Active Shape/Appearance Models (ASMs/AAMs) [6, 7] based methods. Local Feature Analysis (LFA) [8] and Elastic Bunch Graph Matching (EBGM) [9] are typical instances of the latter category, among which LFA has been developed to one of the most successful commercial face recognition system, named FaceIt®, by Identix Corporation. FERET evaluation has provided extensive comparisons of these algorithms [10] as well as a kind of evaluation protocol for face recognition systems. More recently, Support Vector Machine (SVM) has also been applied to face recognition successfully [11].

However, face recognition remains a difficult, unsolved problem in general. The performance of almost all current face recognition systems, both the best academic results and the most successful commercial systems, is heavily subject to the variations in the imaging conditions. It has been discovered by the FERET and FRVT2000 test that pose and illumination variations are among the several bottlenecks for a practical face recognition system [10].

In this study we will deal with normalization and handling of varying illumination. The changes induced by illumination are often larger than the differences between individuals, causing systems based directly on comparing images to misclassify input images [12]. So far, no revolutionary or practical solutions are available for these

problems. However, some approaches have been proposed to handle illumination. These can be divided into four categories:

- Extraction of illumination invariant features
- Transformation of images with variable illuminations to a canonical representation
- Modeling the illumination variations
- Utilization of some 3D face models whose facial shapes and albedos are obtained in advance

The approaches of the first category use illumination invariant face features to represent face images and for face recognition. Examples of these features include edge maps, image intensity derivations, and images convolved with the 2-D Gabor-like filters. Edge maps can serve as relatively robust representations of illumination changes for some classes of objects such as telephones, tables, but for other subjects such as faces, parts of the edges do not remain stable. In [12], Adini's empirical study shows that none of these representations is sufficient by itself to overcome image variations caused by a change in the illumination direction.

The approaches in the second category transform face images with variable illuminations to canonical forms. Zhao and Chellappa [13] proposed a symmetrical shape from shading (SSFS) method based on a single face image of each person. The method uses a simple and generic 3D face model to estimate the pose and light source direction of the input image. Then, the input image is warped to a virtual image with frontal illumination, which is matched to the prototype images in a gallery. Shashua and Tammy [14] proposed the Quotient Image, which is defined as the ratio between the albedo functions of two objects under the assumption that all the faces have the same shape. Therefore, the Quotient Image is an illumination invariant signature of a face. This technique can produce an illumination invariant prototype image, which can be used for face recognition.

The approaches in the third category represent illumination variations by means of an appropriate model based on many images of each person under various lighting conditions. Shashua [15] proposed that the set of images with variable illuminations of a face subject lies in a 3D linear subspace in the absence of shadowing. Efforts were made to enhance the 3D linear subspace model to deal with the shadows by segmenting a face image into some sub-regions that have surface normals with

directions close to each other. Georghiades et al. [16] proposed the individual illumination cone model that can be approximated by a low-dimensional linear subspace whose basis vectors are estimated using a large number of images under different lighting conditions. The images of each person are generated from a 3D facial shape and albedos which are reconstructed using seven frontal face images under different lighting conditions. However, the methods are computationally intensive.

In general, the methods belonging to the fourth category can achieve good recognition results, and are more commonly in use at present. However, the methods require for each face subject either many images captured under different lighting conditions or many synthesized images with different illuminations by virtue of a set of bootstrap images. 3D Morphable Models are also classified into this category. This theory utilized a 3D Morphable Model to synthesize new face images with different lighting directions, and illumination variations are modeled using the Phong illumination model. However, the 3D face models are usually captured by a 3D Laser Scanner.

1.4 Outline of proposed approach

This study is based in three main axes which have been developed in an effort to overpass three problems of Face Recognition dealing with varying illumination:

- Difficulties of Principal Component Analysis (PCA) [3] and Linear Discriminant Analysis (LDA) [4] with varying illumination
- Difficulty of feature extraction due to the "small sample size " (SSS) problem
- Reconstruction of images illuminated by an arbitrary angle in order to have a frontal illumination

For the first problem, PCA and LDA are two powerful tools for dimensionality reduction and feature extraction. Still PCA is presumed to fail when dealing with variable illumination; and LDA methods, which outperform PCA, are supposed to loose some important discriminatory information during the effort to overcome the scatter matrix singularity by discarding the null space. To surpass this problem a combination of Fisher-LDA and Direct-LDA has been used and the results are overall superior to the other two methods.

For the second problem, by focusing on Lambertian surface classes and by creating analytically an image signature of an arbitrary illuminated input image, we are able to create the image space of this image for different illumination directions. This approach makes possible the multiplication of images in the training databases where only one image per person is available and so to overcome the SSS problem.

For the third problem, many approaches have been tried to represent images with some canonical forms which are insensitive to illumination. Recently it has been showed that the ratio of two images of the same class is simpler than if the images are from different objects. This concept has been used to create a reconstructionist approach in order to restore arbitrary illumination of input images to frontal ones.

2. FACE SPACE AND DIMENSIONALITY REDUCTION

In this section we will give an insight on the image space, a small part of which represents the face space. Furthermore, we will introduce how instead of dealing with such a huge space we can process the data retained by dimensionality reduction methods such as PCA and LDA.

2.1 Image space vs. face space

When using appearance based methods, we usually represent an image of width $\boldsymbol{w}$ and height $\boldsymbol{h}$ as a vector in a $w \cdot h$–dimensional space. In practice, this ($w \cdot h$-dimensional) space, i.e. the full image space, is too large to allow robust and fast object recognition. A common way to attempt to resolve this problem is to use dimensionality reduction techniques. This can be done also because we know that the face space is concentrated only in a small part of the image space. Two of the most popular techniques for this purpose are: PCA and LDA (also known as Fisher Discriminant Analysis).

2.2 Principal Component Analysis

Principal component analysis is a technique that is useful for the compression and classification of data. It is well known that there exist significant statistical redundancies in natural images. For a limited class of objects such as face images that are normalized with respect to scale, translation and rotation the redundancy is even greater. The purpose is to reduce the dimensionality of a data set (sample) by finding a new set of variables, smaller than the original set of variables that nonetheless retains most of the sample's information. By information we mean the variation present in the sample, given by the correlations between the original variables. The new variables, called principal components (PCs), are uncorrelated, and are ordered by the fraction of the total information each retains.

To better understand the role of the PCA we can give a graphical example in which we have n observations in the 2-D space $x=(x_1,x_2)$. The goal is to account for the variation in a sample in as few variables as possible, to some accuracy. As we see in Fig 2.1 PCA aims the better representation of these samples.

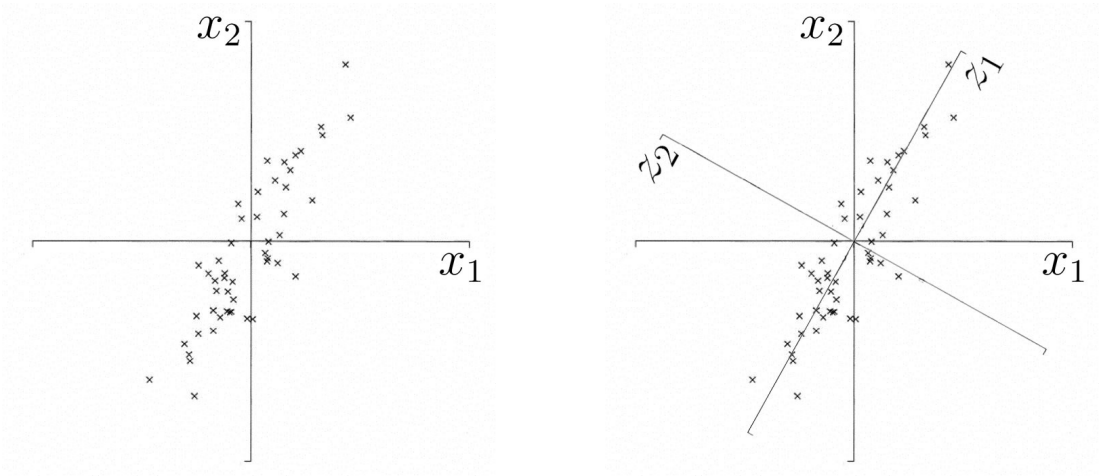

Fig 2.1: N observations in the 2-D space (left). Their presentation in the PCA space

More formally let us consider a set of N sample images $\{x_1,x_2,...,x_N\}$ taking values in an n-dimensional image space, and assume that each image belongs to one of c classes $\{X_1,X_2,...,X_c\}$. We consider a linear transformation mapping the original n-dimensional image space into an m-dimensional feature space, where (m < n). The new feature vectors $y_k \in \mathbb{R}^m$ are defined by the following linear transformation:

$$y_k = W^T x_k \qquad k = 1,2,...,N \tag{2.1}$$

where $W \in \mathbb{R}^{nxm}$ is a matrix with orthonormal columns. If the total scatter matrix S_T (covariance matrix) is defined as:

$$S_T = \sum_{k=1}^{N} (x_k - \mu)(x_k - \mu)^T \tag{2.2}$$

where N is the number of sample images and $\mu \in \mathbb{R}^n$ is the mean image of all samples, then after applying the linear transformation W^T, the scatter of the transformed feature vectors $\{y_1,y_2,...,y_N\}$ is $W^T S_T W$. In PCA, the projection W_{opt} is

chosen to maximize the determinant of the total scatter matrix of the projected samples, i.e.:

$$W_{opt} = \arg\max_{W} \left| W^T S_T W \right| = [w_1, w_2, ..., w_m] \tag{2.3}$$

where $\{w_i \mid i = 1, 2, ...m\}$ is the set of n-dimensional eigenvectors of S_T corresponding to the m largest eigenvalues. Since these eigenvectors have the same dimension as the original images, they are referred to as Eigenpictures or Eigenfaces. If classification is performed using a nearest neighbor classifier in the reduced feature space and N is chosen to be the number of images in the training set, then the Eigenface method is equivalent to a simple correlation method. So we had that:

$$S_T \Phi = \Phi \Lambda \tag{2.4}$$

where Λ is a diagonal matrix whose diagonal elements are the eigenvalues of S_T with their magnitudes in descending order, and Φ is a matrix whose i^{th} column is the i^{th} eigenvector of S_T. In order to obtain the eigenspace we generally choose m eigenvectors corresponding to the m largest eigenvalues, which capture over 95% of the variations of the images. Fig 2.2 and Fig 2.3 shows a training set of 10 person faces taken from the YaleB database[17] and their respective eigenfaces (eigenpictures) sorted in descending order according to their eigenvalues(there are only nine eigenfaces). The eigenpictures have been normalized for display purpose:

Fig 2.2 Face examples of 10 different persons from the YaleB database

Fig 2.3 Eigenfaces calculated on basis of the 10 examples from the YaleB database

After calculating the eigenfaces the projection is the only process left to done. Let Φ_S be the matrix consisting of the m eigenvectors, and I_n be a new face image. The projection of I_n onto the eigenspace is represented as follows:

$$a = \Phi_S^T (I_n - \mu) \tag{2.5}$$

Where ***a*** is m*x*1 vector containing the m projections coefficients. The reconstructed image $I_n^{'}$ is then given as follows:

$$I_n^{'} = \Phi_s a + \mu \tag{2.6}$$

The reconstructed image is the best approximation of the raw input image in the mean square sense. An advantage of using such representations is their reduced sensitivity to noise. Some of this noise may be due to small occlusions, as long as the topological structure does not change. For example, good performance under blurring, partial occlusion and changes in background has been demonstrated in many eigenpictures based systems. This should not come as a surprise, since the PCA reconstructed images are much better than the original distorted images in terms of their global appearance.

For better approximation of face images outside the training set, for example using an extended training set that adds mirror-imaged faces showed to achieve lower approximation error. Using such an extended training set, the eigenpictures are either

symmetric or antisymmetric, with the most leading eigenpictures typically being symmetric.

A drawback of this approach is that the scatter being maximized is due not only to the between-class scatter that is useful for classification, but also to the within-class scatter that, that for classification purposes is unwanted information. Thus if PCA is presented with images of faces under varying illumination, the projection matrix W_{opt} will contain principal components (i.e., Eigenfaces) which retain, in the projected feature space, the variation due lighting. Consequently, the points in the projected space will not be well clustered, and worse, the classes may be smeared together. It has been suggested that by discarding the three most significant principal components, the variation due to lighting is reduced [4]. The hope is that if the first principal components capture the variation due to lighting, then better clustering of projected samples is achieved by ignoring them. Yet, it is unlikely that the first several principal components correspond solely to variation in lighting; as a consequence, information that is useful for discrimination may be lost.

2.3 Fisher's Linear Discriminant and Fisher Faces

Fisher's Linear Discriminant Analysis (FLDA) uses an important fact of photometric stereo: In the absence of shadowing, given three images of a Lambertian surface from the same viewpoint taken under three known, linearly independent light source directions, the albedo and surface normal can be recovered. For classification, this fact has great importance: It shows that, for a fixed viewpoint, the images of a Lambertian surface lie in a 3D linear subspace of the high-dimensional image space. One can perform dimensionality reduction using linear projection and still preserve linear separability. Fisher's Linear Discriminant (FLD) is an example of a class specific method, in the sense that it tries to "shape" the scatter in order to make it more reliable for classification.

More formally, let us continue our previous study of PCA from this new point of view. The LDA selects W in such way that the ratio of the between-class scatter and the within-class scatter is maximized. Let the between-class scatter matrix be defined as:

$$S_B = \sum_{i=1}^{c} N_i (\mu_i - \mu)\ (\mu_i - \mu)^T \tag{2.7}$$

and the within-class scatter be defined as

$$S_w = \sum_{i=1}^{c}\ \sum_{k=1}^{n} (x_k - \mu_i)\ (x_k - \mu_i)^T \tag{2.8}$$

where μ_i is the mean image of class X_i , and N_i is the number of samples in class X_i. If S_W is nonsingular, the optimal projection W_{opt} is chosen as the matrix with orthonormal columns which maximizes the ratio of the determinant of the between-class scatter matrix of the projected samples to the determinant of the within-class scatter matrix of the projected samples, i.e.,

$$W_{opt} = \arg\max_{w} \frac{\left|W^T S_B W\right|}{\left|W^T S_W W\right|} = [w_1\ w_2 \ldots w_m] \tag{2.9}$$

where $\{w_i \mid i = 1,2,\ldots,m\}$ is the set of generalized eigenvectors of S_B and S_W corresponding to the largest eigenvalues $\{\lambda i \mid i = 1,2,\ldots,m\}$, i.e.,

$$S_B w_i = \lambda_i S_W w_i \qquad i = 1, 2, \ldots, m \tag{2.10}$$

Note that there are at most c - 1 nonzero generalized eigenvalues, and so an upper bound on m is c - 1, where c is the number of classes.

A graphical example can better explain the role of LDA. If we imagine that the samples labelled ω_1 fall more or less into one cluster while those labelled ω_2 all in another, we want the projections falling onto the line to be well separated, not thoroughly intermingled. Figure 2.4 illustrates the effect of choosing two different values for ω for a two-dimensional example. It should be abundantly clear that if the original distributions are multimodal and highly overlapping, even the "best" W_{opt} is unlikely to provide adequate separation, and thus this method will be of little use.

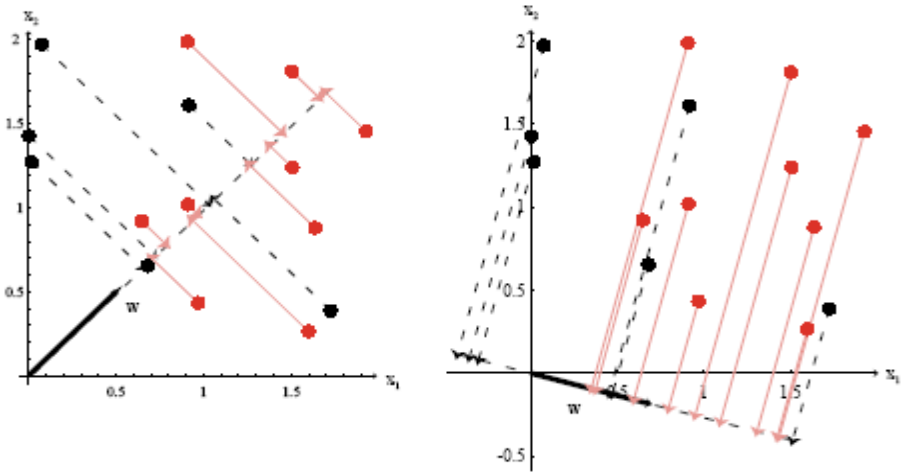

Figure 2.4: Projection of samples onto two different lines. The figure on the right shows greater separation between the red and black projected points.

In the face recognition problem, one is confronted with the difficulty that the within-class scatter matrix $S_W \in \mathbb{R}^{nxn}$ is always singular. This stems from the fact that the rank of S_W is at most (N – c), and, in general, the number of images in the learning set N is much smaller than the number of pixels in each image n. This means that it is possible to choose the matrix W such that the within-class scatter of the projected samples can be made exactly zero.

In order to overcome the complication of a singular S_W, it has been proposed an alternative to the criterion in (2.9). This method, which is called Fisherfaces, avoids this problem by projecting the image set to a lower dimensional space so that the resulting within-class scatter matrix S_W is nonsingular. This is achieved by using PCA to reduce the dimension of the feature space to N - c, and then applying the standard FLD defined by (2.9) to reduce the dimension to c - 1. More formally, W_{opt} is given by:

$$W_{opt}^T = W_{fld}^T \cdot W_{pca}^T \tag{2.11}$$

where

$$W_{pca} = \arg\max_W \left| W^T S_W W \right| \tag{2.12}$$

$$W_{fld} = \arg\max_W \frac{\left| W^T W_{pca}^T S_B W_{pca} W \right|}{\left| W^T W_{pca}^T S_W W_{pca} W \right|} \tag{2.13}$$

Note that the optimization for W_{pca} is performed over $n \times (N-c)$ matrices with orthonormal columns, while the optimization for W_{fld} is performed over $(N-c) \times m$ matrices with orthonormal columns. In computing W_{pca}, only the smallest (c – 1) principal components have been thrown away.

The Fisherface method appears to be the best simultaneously handling variation in lighting and has lower error rate than the PCA method. But to achieve better results some preprocessing as normalization to zero mean, unit standard deviation, and alignment of face textures is needed.

2.4 A variant of Direct LDA

As we stated in formula (2.9) assuming that S_W is non-singular, the basis vectors W correspond to the first m eigenvectors with the largest eigenvalues of ($S_W^{-1} S_B$). The m-dimensional representation is then obtained by projecting the original face images onto the subspace spanned by the m eigenvectors. However, a degenerated S_W in (2.9) may be generated due to the SSS problem widely existing in most FR tasks. As we said in the previous section a possible solution is to apply a PCA step in order to remove the null space of S_W prior to the maximization in (2.9). Nevertheless, it recently has been shown that the null space of S_W may contain significant discriminatory information. As a consequence, some of the significant discriminatory information may be lost due to this pre-processing PCA step. An approach that is a successful variant of D-LDA [18] will be introduced next.

The basic premise of the D-LDA methods that attempt to solve the SSS problem without a PCA step is, that the null space of S_W contains significant discriminant information if the projection of S_B is not zero in that direction, and that no significant information will be lost if the null space of S_B is discarded. Assuming that A and B represent the null space of S_B and S_W, while $A' = \square^N - A$ and $B' = \square^N - B$ are the complement spaces of A and B respectively, the optimal discriminant subspace sought by D-LDA is the intersection space ($A' \cap B$). So a first approach is to find A' when seeking the solution of (2.9) by diagonalizing S_B. Fortunately the rank is determined by $rank(S_B) = \min(N, C-1)$ with C the number of image classes. Then ($A' \cap B$) can be obtained by solving the null space of projection of S_W into A'. Based on the analysis given above, it can be known that the most significant discriminant

information exist in the intersection subspace ($A' \cap B$), which is usually low-dimensional so that it becomes possible to further apply some sophisticated techniques, such as the rotation strategy of the LDA subspace used in F-LDA, to derive the optimal discriminant features from the intersection subspace.

The maximization process in (2.9) is not directly linked to the classification error which is the criterion of performance used to measure the success of the FR procedure. Modified versions of the method, such as the F-LDA approach, use a weighting function in the input space, to penalize those classes that are close and can potentially lead to misclassifications in the output space. Thus, the weighted between-class scatter matrix can be expressed as:

$$\hat{S}_B = \sum_{i=1}^{C} \Phi_i \Phi_i^T \tag{2.14}$$

where $\Phi_i = (N_i / N)^{1/2} \sum_{j=1}^{C} (w(d_{ij}))^{1/2} (\mu_i - \mu_j)$ and $d_{ij} = \left\| \mu_i - \mu_j \right\|$ is the Euclidian distance between the mean of class ***i*** and ***j***. The weighting function $w(d_{ij})$ is a monotonically decreasing function of the distance d_{ij}. The only constraint is that the weight should drop faster than the Euclidean distance between the means of class ***i*** and class ***j***. In this study instead of the conventional Fisher criterion as in (2.9) a new metric is proposed:

$$W = \arg\max_{W} \frac{\left| W^T \hat{S}_B W \right|}{\left| \eta (W^T \hat{S}_B W) + W^T S_W W \right|} \tag{2.15}$$

where $0 \le \eta \le 1$ is a regularization parameter. It can be proven that although Eq. (2.15) looks quite different from the conventional Fisher's criterion, it can be show that they are exactly equivalent.

The next steps of the approach can be given by the following algorithm:

1. Calculate those eigenvectors of $\Phi_b^T \Phi_b$ with non-zero eigenvalues: $E_m = \left[e_1 ... e_m \right]$ where $m \le C - 1$ and Φ_b is from $\hat{S}_B = \Phi_b \Phi_b^T$.
2. Calculate the first **m** most significant eigenvectors and their corresponding eigenvalues of $\hat{S}_B$ by $V = \Phi_b E_m$ and $\Lambda_b = V^T \hat{S}_B V$.

3. Let $U = V\Lambda_b^{-1/2}$. Calculate eigenvectors of $(U^T S_W U)$, $P = [p_1, ..., p_m]$ in increasing order.
4. Choose the first $M (\leq m)$ eigenvectors in **P**. Let P_M and Λ_w be the selected eigenvectors and their corresponding eigenvalues, respectively.
5. Let $W = HP_M (\eta I + \Lambda_W)^{-1/2}$ be the basis of the dimensionality reduction.
6. The optimal discriminant feature representation of $\boldsymbol{x}$ can be obtained by $y = \varphi(x) = W^T x$.

It can be seen from the algorithm that varying the values of η within [0, 1] leads to a set of intermediate D-LDA representations variants in the feature space. Since the subspace spanned by W may contain the intersection space ($A' \cap B$), it is possible that there exist zero or very small eigenvalues in Λ_W, which have been shown to be high variance for estimation in the SSS environments [18]. As a result, any bias arising from the eigenvectors corresponding to these eigenvalues is dramatically exaggerated due to the normalization process ($P_M \Lambda_W^{-1/2}$). Against the effect, the introduction of the regularization helps to decrease the importance of these highly unstable eigenvectors, thereby reducing the overall variance. Also, there may exist the zero eigenvalues in Λ_W, which are used as divisors due to η = 0 so that the LDA process can not be carried out. However, it is not difficult to see that the problem can be avoided simply by setting the regularization parameter $\eta > 0$.

In order to display the effect of LDA in dimensionality reduction we take a training set form the YaleB database by choosing only three images per person (Fig 2.5).

Fig. 2.5 Training set used for Linear Discriminant Analysis example

As we see the faces belonging to a person are not very different and they illumination angles also differ by at most 15'. So we expect that the eigenvectors of this training set will be similar to those extracted by the PCA method in Fig 2.3, because of the means of each class are quite near the value of each sample of that class. Indeed we can see in figure 2.6 that they are quite similar. On the other hand they differ in the values of most important non-zero eigenvalues and we must take into consideration that they have been normalized for display purposes. So we arrive in the conclusion that in order to apply successfully D-LDA we have to train classes with some constituent differences.

Fig 2.6: Eigenvectors resultant that will be used in the D-LDA dimension projection

The performance of this algorithm depends critically in the choosing of the regularization parameter and the number of the selected eigenvectors. It has been observed that the optimal regularization parameter decreases monotonously as the number of training samples per object increase. It seems that the relationship is not linear. In addition in the experiments executed the η value increases as the number of subjects increase.

2.5 Comparison between Linear Projection Algorithms

It is helpful to make comparisons among these linear projection algorithms. In all the projection algorithms, classification is performed by projecting the input x into a subspace via a projection/basis matrix W in the form $y = W^T x$ comparing the projection coefficient vector y of the input to all the prestored projection vectors of

labeled classes to determine the input class label. The vector comparison varies in different implementations and can influence the system's performance dramatically. For example, PCA algorithms can use either the angle or the Euclidean distance (weighted or unweighted) between two projection vectors. For LDA algorithms, the distance can also be unweighted or weighted.

Furthermore PCA requires only one picture for individual while LDA-based methods require at least 2. As we showed in the previous example if the classes' variance is not prominent then there will no much difference between PCA and LDA projection matrix. In the extreme case where only one sample per class is available the LDA method is roughly transformed in plain PCA. We have to say that the PCA-based method is optimal in terms of image reconstruction, thereby provides some insight on the original image structure or image distribution, which is highly complex and non-separable. In spite of that D-LDA methods greatly increase the separability of subjects.

3. IMAGE SYNTHESIS BASED ON QUOTIENT IMAGE

Face Recognition Systems' performance is strongly related with the databases on which these systems are based for training or learning. This fact becomes even more important when class-based discrimination methods are used. As we said for the LDA method in the previous chapter many samples per class have to be provided in order to successfully obtain the high results of this method. Unfortunately in many FRS only one image per person is provided and the use of LDA-based methods becomes impossible. In order to surpass this difficult obstacle many proposals have been done. Following, we will introduce some successful techniques [14, 19] in image synthesis and re-rendering. These approaches make possible the synthesis of the image space of a given input image being based on small bootstrap of images (training database).

Nearly all approaches to view synthesis or image-based rendering take a set of images gathered from multiple viewpoints and apply techniques related to structure from motion, stereopsis, image transfer, image warping, or image morphing. Each of these methods requires the establishment of correspondence between image data (e.g. pixels) across the set. Since dense correspondence is difficult to obtain, most methods extract sparse image features (e.g. corners, lines), and may use multi-view geometric constraints or scene-dependent geometric constraints to reduce the search process and constrain the estimates. By using a sequence of images taken at nearby viewpoints, incremental tracking can further simplify the process, particularly when features are sparse.

For these approaches to be effective, there must be sufficient texture or viewpoint-independent scene features, such as albedo discontinuities or surface normal discontinuities. From sparse correspondence, the epipolar geometry can be established and stereo techniques can be used to provide dense reconstruction. Underlying nearly all such stereo algorithms is a constant brightness assumption, that is, the intensity (irradiance) of corresponding pixels should be the same. In turn, constant brightness implies two seldom stated assumptions: (1) The scene is

Lambertian, and (2) the lighting is static with respect to the scene – only the viewpoint is changing.

3.1 Image Synthesis using Illumination Cone [19]

In this illumination-based approach, it is assumed that the surface is Lambertian, although this assumption is very explicit. Geometric correspondence is trivially established, and so the method can be applied to scenes where it is difficult to establish multi-viewpoint correspondence, namely scenes that are highly textured (i.e. where image features are not sparse) or scenes that completely lack texture (i.e. where there are insufficient image features).

At the core of this approach for generating novel viewpoints is a variant of photometric stereo which simultaneously estimates geometry and albedo across the scene. However, the main limitation of classical photometric stereo is that the light source positions must be accurately known, and this necessitates a fixed lighting rig as might be possible in an industrial setting. Instead, the proposed method does not require knowledge of light source locations, and so illumination could be varied by simply waiving a light around the scene. In fact, this method derives from work by Belhumeur and Kriegman in [20] where they showed that a small set of images with unknown light source directions can be used to generate a representation, the illumination cone, which models the complete set of images of an object (in fixed pose) under all possible illumination.

Generated images from the illumination cone representation accurately depict shading and attached shadows under extreme lighting. Furthermore the cone representation was extended to include cast shadows (shadows the object casts on itself) for objects with non-convex shapes. Unlike attached shadows, cast shadows are global effects, and their prediction requires the reconstruction of the object's surface. In generating the geometric structure, multi-viewpoint methods typically estimate depth directly from corresponding image points. It is well known that without sub-pixel correspondence, stereopsis provides a modest number of disparities over the effective operating range, and so smoothness or regularization constraints are used to interpolate and provide smooth surfaces. This illumination-based method estimates surface normals which are then integrated to generate a surface.

3.1.1 Illumination Modeling

In [20] it is shown that, for a convex object with a Lambertian reflectance function, the set of all images under an arbitrary combination of point light sources forms a convex polyhedral cone in the image space $\mathbb{R}^n$ which can be constructed with as few as three images.

Let $x \in \mathbb{R}^n$ denote an image with n pixels of a convex object with a Lambertian reflectance function illuminated by a single point source at infinity. Let $B \in \mathbb{R}^{n\times 3}$ be a matrix where each row in B is the product of the albedo with the inward pointing unit normal for a point on the surface projecting to a particular pixel in the image. A point light source at infinity can be represented by $s \in \mathbb{R}^3$ signifying the product of the light source intensity with a unit vector in the direction of the light source. A convex Lambertian surface with normals and albedo given by B, illuminated by s, produces an image x given by:

$$x = \max(Bs, 0) \tag{3.1}$$

where *max(Bs,0)* sets to zero all negative components of the vector *Bs*. The pixels set to zero correspond to the surface points lying in an attached shadow. By firstly assuming that objects are convex it is noted that when no part of the surface is shadowed, x lies in the 3-D subspace L given by the span of the columns of B. If an object is illuminated by k different light sources, then the image is given by the superposition of the images which would have been produced by the individual sources:

$$x = \sum_{i=1}^{k} \max(Bs_i, 0) \tag{3.2}$$

where s_i is a single light source. Due to the inherent superposition, it follows that the set of all possible images C of a convex Lambertian surface created by varying the direction and strength of an arbitrary number of point light sources at infinity is a convex cone. It is also evident from Equation (3.2) that this convex cone is completely described by matrix B.

So to construct the illumination model for an individual: gather three or more images of the face without shadowing illuminated by a single light source at unknown locations but viewed under fixed pose, and use them to estimate the three-

dimensional illumination subspace L. This can be done by first normalizing the images to unit length and then estimating the best three-dimensional orthogonal basis B^* using a least-squares minimization technique such as singular value decomposition (SVD).

Unfortunately, using SVD in the above procedure leads to an inaccurate estimate of B^*. For even a convex object whose Gaussian image covers the Gauss sphere, there is only one light source direction (the viewing direction) for which no point on the surface is in shadow. For any other light source direction, shadows will be present. If the object is non-convex, such as a face, then shadowing in the modeling images is likely to be more pronounced. When SVD is used to find B^* from images with shadows, these systematic errors bias its estimate significantly. In [19] an alternative way has been proposed to find B taking into account the fact that some data values should not be used in the estimation and finds a basis B^* for the 3-D linear subspace L from image data with missing elements.

3.1.2 Surface Reconstruction and Image Synthesis

After estimating B^* the next step followed in this approach is the generation of the object's surface with the condition of integrability constraints. Then the corresponding surface $\hat{z}(x, y)$ is estimated by fitting it to the given components of the gradient p and q by minimizing the functional

$$\iint_{\Omega} (\hat{z}_x - p)^2 + (\hat{z}_y - q)^2 \, dxdy \tag{3.3}$$

the Euler equation of which reduces to $\nabla^2 z = p_x + q_y$. By enforcing the right natural boundary conditions and employing an iterative scheme that uses a discrete approximation of the Laplacian, the desired surface $\hat{z}(x, y)$ is obtained.

The estimated surface can now be used to generate images of an object under novel illumination conditions. To determine cast shadows ray-tracing is employed. Specifically, a point on the surface is in cast shadow if, for a given light source direction, a ray emanating from that point parallel to the light source direction intersects the surface at some other point. With this extended image formation model, the generated images exhibit realistic shading and have strong attached and cast shadows.

3.2 Ideal Class Assumption and Image Re-rendering

This approach is based on a recent result showing that the set of all images generated by varying lighting conditions on a collection of Lambertian objects all having the same shape but differing in their surface texture (albedo) can be characterized analytically using images of a prototype object and a (illumination invariant) "signature" image per object of the class. The Cartesian product between the signature image of an object *y* and the linear subspace determined by the images of the prototype object generates the image space of *y*. The second result is on how to obtain the signature image from a data base of example images of several objects while proving that the signature image obtained is invariant to illumination conditions.

In this approach the consideration will be restricted to objects with a Lambertian reflectance function, i.e., the image can be described by the product of the albedo (texture) and the cosine angle between a point light source and the surface normal:

$$I(x,y) = \rho(x,y)n(x,y)^T s \tag{3.4}$$

where $0 \le \rho(x,y) \le 1$ is the surface reflectance associated with point (x, y) in the image, n(x, y) is the surface normal direction associated with point (x, y) in the image, and s is the (white) light source direction (point light source) and whose magnitude is the light source intensity.

The basic result that is used in this approach is that the image space generated by varying the light source vector *s* lives in a three-dimensional linear subspace [15]. To see why this is so consider three images I_1, I_2, I_3 of the same object (ρ, *n* are fixed) taken under linearly independent light source vectors s_1, s_2, s_3, respectively. The linear combination $\sum_j \alpha_j I_j$ is an image $I = \rho n^T s$, where $s = \sum_j \alpha_j s_j$. Thus, ignoring shadows, three images are sufficient for generating the image space of the object. The basic principle can be extended to deal with shadows, color images, nonwhite light sources, and non-Lambertian surfaces.

Next will be given two definitions related with what is meant an "ideal class" of objects and the main aim of this approach – the image synthesis.

Definition 1(Ideal Class of objects): An ideal class is a collection of 3D objects that have the same shape but differ in the surface albedo function. The image space of such a class is represented by:

$$\rho_i(x,y)n(x,y)^T s_j \tag{3.5}$$

where $\rho_i(x, y)$ is the albedo (surface texture) of object I of the class, n(x, y) is the surface normal (shape) of the object (the same for all objects of the class), and s_j is the point light source direction, which can vary arbitrarily.

Definition 2(Synthesis (Re-rendering) Problem): Given $N \times 3$ images of N objects of the same class, illuminated under three distinct lighting conditions and a single image of a novel object of the class illuminated by some arbitrary lighting condition, synthesize new images of the object under new lighting conditions.

In practice, objects of a class do have shape variations, although to some coarse level the shape is similar. The ideal class could be satisfied if we perform pixel-wise dense correspondence between images (say frontal images) of the class. The dense correspondence compensates for the shape variation and leaves only the texture variation. By far, the goal of the second definition is to expand linear subspace of images space $\rho n^T s$, where only s varies, to the case where both $\boldsymbol{\rho}$ and $\boldsymbol{s}$ vary.

3.2.1 The Quotient Image

Given two objects y; a, we define the quotient image Q by the ratio of their albedo functions ρ_y/ρ_a. Clearly, Q is illumination invariant. In the absence of any direct access to the albedo functions, we show that Q can nevertheless be recovered, analytically, given a bootstrap set of images. Once Q is recovered, the entire image space (under varying lighting conditions) of object a can be generated by Q and three images of object a.

$$Q_y(u,v) = \frac{\rho_y(u,v)}{\rho_a(u,v)} \tag{3.6}$$

where u,v changes over the image. Thus the image Qy depends only on the relative surface texture information and is independent of illumination. The importance of this ratio becomes clear by the following statement which can be proven very easily:

- Given three images a_1, a_2, a_3 of object a illuminated by any three linearly independent lighting conditions and an image y_s of y illuminated by some light source s, then there exists coefficients x_1, x_2, x_3 that satisfy

$$y_s = (\sum_j x_j a_j) \otimes Q_y \tag{3.7}$$

where $\otimes$ denotes the Cartesian product (pixel by pixel multiplication). Moreover, the image space of object y is spanned by varying the coefficients.

We see that once Q_y is given, we can generate y_s (the novel image) and all other images of the image space of y. The key is obtaining the quotient image Q_y. Given y_s, if somehow we were also given the coefficients x_j that satisfy $s = \sum_j x_j s_j$, then Q_y readily follows: $Q_y = y_s / (\sum_j x_j a_j)$, thus the key is to obtain the correct coefficients x_j. For that reason, and that reason only, we need the bootstrap set. Otherwise, a single object a would suffice.

Let the bootstrap set of 3N pictures be taken from three fixed (linearly independent) light sources s_1, s_2, s_3 (the light sources are not necessarily known). Let A_i, $i = 1,...,N$, be a matrix whose columns are the three pictures of object a_i with albedo function ρ_i. Thus, $A_1,..., A_N$ represent the bootstrap set of N matrices, each is a $m \times 3$ matrix, where m is the number of pixels of the image (assuming that all images are of the same size). Let y_s be an image of some novel object y (not part of the bootstrap set) illuminated by some light source $s = \sum_j x_j s_j$. We wish to recover $x = (x_1, x_2, x_3)$ given the N matrices $A_1,..., A_N$ and the vector y_s. Therefore we have to solve a bilinear problem in the $N+3$ unknowns x and α_i:

$$\min_{x,\alpha_i} \left| y_s - \sum_{i=1}^{N} \alpha_i A_i x \right|^2 \tag{3.8}$$

This problem can be rearranged in order to simplify transformations in the form of function in which the minimum is at x, i.e: $x = \arg\min f(\hat{x})$:

$$f(\hat{x}) = \frac{1}{2} \sum_{i=1}^{N} \left| A_i \hat{x} - \alpha_i y_s \right|^2 \tag{3.9}$$

In order to find the desired global minima we apply the Euler-Lagrange equation according the two variables x and α_i and we get the following relations:

$$x_0 = \sum_{i=1}^{N} \alpha_i v_i \tag{3.10}$$

$$v_i = (\sum_{r=1}^{N} A_r^T Ar)^{-1} A_i^T y_s \tag{3.11}$$

where the coefficients α_i are determined up to a uniform scale as the solution of the symmetric homogeneous linear system of equations derived from $\partial f / \partial \alpha_i = 0$:

$$\alpha_i y_s^T y_s - \left(\sum_{r=1}^{N} \alpha_r v_r \right)^T A_i^T y_s = 0 \tag{3.12}$$

The energy function $f(\hat{x})$ in (3.9) consists of a simultaneous projection of y_s onto the subspaces spanned by the columns of A_1, columns of A_2, and so on. In addition, during the simultaneous projection there is a choice of overall scale per subspace, these choices of scale, the α_i, are directly related to the scaling of the axes represented by such that the albedos of the bootstrap set span the albedo of the novel object. When (N=1), the minimum of $f(\hat{x})$ coincides with x iff the albedo of the novel object is equal (up to scale) to the albedo of bootstrap object. The more objects in the bootstrap set, the more freedom we have in representing novel objects. If the albedos of the class of objects are random signals, then at the limit a bootstrap set of m objects (3m images) would be required to represent all novel objects of the class. In practice, the difference in the albedo functions do not cover a large spectrum and instead occupy a relatively small subspace of m, therefore, a relatively small size (N<<m) is required. The algorithm for the synthesizing the image space of a novel object $\boldsymbol{y}$, given the bootstrap set and a single image $\boldsymbol{y_s}$ of $\boldsymbol{y}$ is as follows:

1. We are given N matrices, $A_1,...,A_N$, where each matrix contains three images (as its columns). This is the bootstrap set. We are also given a novel image y_s (represented as a vector of size m, where m is the number of pixels in the image). For good results, we make sure that the objects in the images are roughly aligned
2. Compute N vectors (of size 3) using the equation (3.11):

3. Solve the homogeneous system of linear equations in N described by (3.12) and scale the solution such that $\sum_i \alpha_i = N$.
4. Compute $x = \sum_i \alpha_i v_i$.
5. Compute the quotient image $Q_y = y_s / Ax$, where A is the average of $A_1, \ldots, A_N$.
6. The image space created by the novel object, under varying illumination, is spanned by the product of images Q_y and Az for all choices of z.

In order to test the successful results of this approach I conducted a wide range of experiments. The algorithm was tested using a bootstrap as in Fig. 2.5. The novel images were not part of that database but member of the staff in the Computer Engineering Department of ITU.

Fig 3.1 Example 1 of the image synthesis. The input image in upper leftmost position

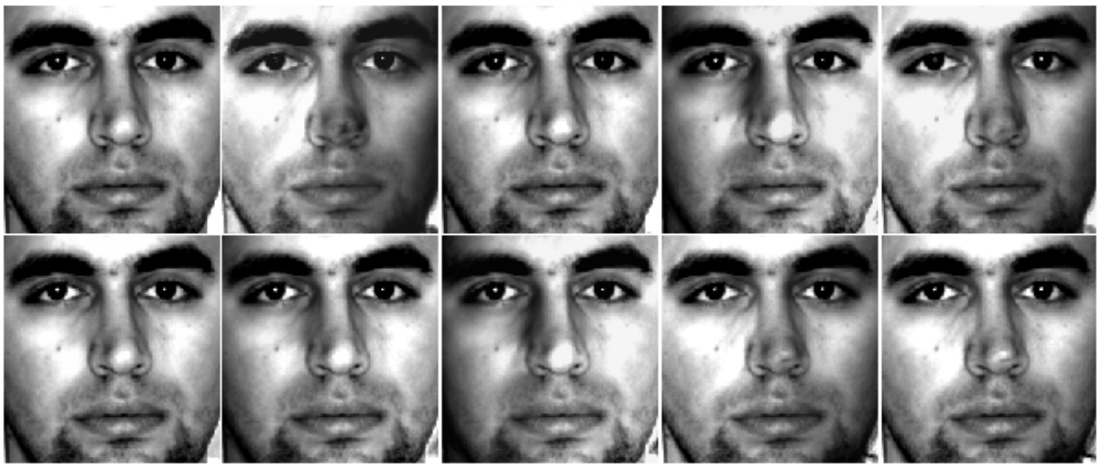

Fig 3.2 Example 2 of the image synthesis. The input image in upper leftmost position

Fig 3.3 Example 3 of the image synthesis. The input image in upper leftmost position

As we see the results of this algorithm are quite satisfactory creating the possibility to use LDA-based methods even when only one image per person is provided during the learning phase. An inherent assumption throughout the algorithm is that for a given pixel (x, y), n(x, y) is the same for all the images- the bootstrap set as well as the test images. The performance degrades when dominant features between the bootstrap set and the test are misaligned. This could arise in a variety of situations such as: 1) the class is of non-smooth objects like objects with sharp corners (chairs, for instance), 2) objects are seen from varying viewing positions , and 3) the class of objects is smooth (like human faces) but gross misalignment is caused by facial expressions, mustache, eye-glasses, etc. As this step will occur after the face detection it is supposed that dominant features will have been depicted by successful approaches such Active Appearance Model [6].

4. ILLUMINATION RESTORATION AND IMAGE RECONSTRUCTION

Illumination is a difficult problem in FRS. In order to overcome this obstacle up to now we have offered two methods for reducing the effects of illumination during the learning and also for dimensionality reduction. In front of us comes the problem of many input images under arbitrary illumination condition that, given without any preprocessing, are object to failure of the system. This is why illumination restoration becomes a prerequisite in creating a robust face recognition system toward illumination. In this chapter we will first introduce a conglomerate method [21, 22, 24] of some successful techniques related with illumination restoration and after that a reconstructionist method for illumination [23] will be introduced. Both methods aim to restore a face image illuminated with arbitrary illumination to having frontal illumination. Therefore these methods belong to the second category of illumination solution approaches, and they do not need calculating the estimation of face normal surface direction and albedos.

4.1 Light Fields

This approach is an appearance based algorithm for face recognition. The algorithm operates by creating a representation of light-fields of the subject's head. It has been firstly used in algorithms regarding pose variable face recognition, but can be extended also to illumination variation [21]. The plenoptic function or light-field is a function which specifies the radiance or light in free space. It is a 5D function of position (3D) and orientation (2D). In 2D, the light-field of a 2D object is also 2D.

Light field $L(\theta,\varphi)$, is a function of each pixel of an image that can be estimated by knowing the pixel intensity and the camera intrinsics as well as the relative orientation of the camera to the object. It has been proposed that we can use Light-Fields as a better space presentation in dealing with varying illumination. In this subsection we will give two different approaches that apply Light-Fields in face recognition.

4.1.1 Fisher Light-Fields

Suppose we are given a set of light-fields $L_{i,j}(\theta,\varphi)$, i = 1,…,N, j = 1,…,M where each of N faces is imaged under M different illumination conditions. We can use Fisher's Linear Discriminant to compute a projection that can be applied to the discriminant task. So we have to find the least square solution of the set of equations:

$$L(\theta,\phi) - \sum_{i=1}^{m} \lambda_i W_i(\theta,\phi) = 0 \tag{4.1}$$

where W_i, i=1,…,m are the generalized eigenvectors computed by LDA. After estimating the λ coefficients we do the recognition task of each incoming image by calculating it light-fields, projecting it in an F-LDA space and then applying a nearest neighbor's technique.

4.1.2 Illumination Invariant Bayesian Face Subregions

An image can be regarded as a product $I(x,y) = R(x,y)L(x,y)$ where R(x, y) is the reflectance map and L(x, y) is the illuminance at each point (x, y). Computing the reflectance and the illuminance fields from real images is, in general, an ill-posed problem. This approach uses two widely accepted assumptions about human vision to solve the problem: 1) human vision is mostly sensitive to scene reflectance and mostly insensitive to the illumination conditions, and 2) human vision responds to local changes in contrast rather than to global brightness levels. This algorithm computes an estimate of L(x, y) such that when it divides I(x, y) it produces R(x, y) in which the local contrast is appropriately enhanced. The solution for L(x, y) is found by minimizing:

$$J(L) = \iint_{\Omega} \rho(x,y)(L-I)^2\,dxdy + \lambda \iint_{\Omega} (L_x^2 + L_y^2)dxdy \tag{4.2}$$

Here Ω refers to the image. The parameter λ controls the relative importance of the two terms. The space varying permeability weight ρ(x; y) controls the anisotropic nature of the smoothing constraint. According to the authors this algorithm increases the rate of recognition over the CMU PIE database from 37.3% to 44%.

4.2 Block based Histogram Equalization Illumination Compensation [22]

A simple yet effective local contrast enhancement method, namely block-based histogram equalization (BHE), is firstly introduced. The resulting image processed using BHE is then compared with the original face image processed using histogram equalization (HE) to estimate the category of its light source. In this scheme, the light source for a human face is divided into 65 categories. Based on the category identified, a corresponding lighting compensation model is used to reconstruct an image that will visually be under normal illumination. In order to eliminate the influence of uneven illumination while retaining the shape information about a human face, a 2D face shape model is used.

4.2.1 Block based histogram equalization method

A light source has a different effect on different regions of a human face. Therefore, to determine the type of light source, one effective method is to compare the face images enhanced locally and globally. In this approach, an image is divided into a number of small blocks, and histogram equalization is performed within each of the image blocks. The pixel intensities in each image block are altered such that the resulting block has a histogram of constant intensity. Histogram equalization can increase the contrast in an image block, and the detailed information such as textures and edges weakened by varying illumination can be strengthened. However, this equalization process will increase the difference between the pixels at the borders of adjacent blocks. In order to avoid the discontinuity between adjacent blocks, they are overlapped by half with each other. Weighted averaging is then applied to smooth the boundaries, i.e.:

$$f(x,y) = \sum_{i=1}^{N} \omega_i(x,y) * f_i(x,y) \tag{4.3}$$

where $f_i(x, y)$ and $f(x, y)$ are the intensity values at (x, y) of block i and the smoothed image, respectively, N is the number of overlapping blocks involved in computing the value at (x, y), and $\omega_i(x, y)$, where $i = 1,\ldots, N$, is a weighting function for block i. The value of N depends on the position of the image block under consideration, which is 4 when the block is not at the border, and 2 or 1 when it is located at the

border or at one of the four corners of an image. The weighting function $\omega_i(x, y)$ is simply a product of individual weighting functions in the x and y directions:

$$\omega_i(x, y) = \omega'(x) * \omega'(y) \tag{4.4}$$

where $\omega'(\cdot)$ is a triangle function as shown below:

$$\omega'(x) = 1 - \left| \frac{x - S_B/2}{S_B/2} \right| \tag{4.5}$$

Where S_B is the length of a block, and x is the relative x-coordinate in the block. BHE is simple and computation required is much lower than that of AHE (Adaptive Histogram Equalization).

4.2.2 A varying illumination compensation algorithm

A face is supposed to be a Lambertian surface which can be described by the product of the albedo and the cosine angle between the point light source and the surface normal. Shashua [14] proposed that different people have the same surface normal but with different albedo. However, in most natural images, albedo change is the predominant factor that causes the gradient intensity and the geometric influence can not be neglected. In this approach is tried to divide the lighting conditions to 65 categories according to the YaleB database. Each of these categories has different azimuth and elevation angles of the lighting.

An image processed by BHE is considered as a reference image. This BHE-processed image is then compared to the same image processed by HE to obtain a pixelwise difference between the two images. This difference image, which is called an illumination map, reflects the effect of the light source on different locations on the face image, and can therefore be used to estimate the illumination category. To determine the illumination category of a query image, its illumination map is first computed. Then, linear discriminant analysis (LDA) is used to determine the illumination category of the image. In this approach, the training images are divided into 65 different categories, and each category includes 9 images that are under the same lighting condition and belong to different people in the YaleB database.

4.2.3 Illumination Compensation

For each point (x, y) in an image the effect of illumination can be written as follows:

$$f'(x,y) = A_i(x,y) \cdot f(x,y) + B_i(x,y), \qquad i = 1,...,65 \tag{4.6}$$

where f (x, y) and f'(x, y) represent the intensity values of the image under normal lighting condition and the image under a certain kind of illumination, respectively. $A_i(x, y)$ denotes the multiplication noise and $B_i(x, y)$ is the additive noise for the illumination mode *i*. These two functions depend on the lighting category, and we assume that they are more or less the same for images under the same illumination condition. Based on the training images in the YaleB database, we can estimate the optimal values for $A_i(x, y)$ and $B_i(x, y)$ for each illumination category by means of the least-squared method.

Based on the A-map and B-map of a category, the corresponding image f(x, y) which is under normal lighting can be computed from f'(x, y):

$$f(x,y) = \frac{f'(x,y) - B_i(x,y)}{A_i(x,y)}, \quad i = 1,...,65 \tag{4.7}$$

In order to avoid overflowing, all the intensity values of f(x, y) are restricted to the range of [0,255], so Eq. 5 can be rewritten as follows:

$$f(x,y) = \begin{cases} 0, & f(x,y) < 0 \\ 255, & f(x,y) > 255, \quad i = 1,...,65 \\ \dfrac{f'(x,y) - B_i(x,y)}{A_i(x,y)} & \text{otherwise} \end{cases} \tag{4.8}$$

The block size for the BHE process will affect the performance in determining the lighting category, as well as so the performance in compensating for the illumination effect and the rate for face recognition. Table 1 shows the recognition rates with different block sizes used in BHE. The PCA was used in the experiment. The block size to be used should be proportional to the size of the face under consideration. In this scheme, the block size is set based on the distance between the two eyes of a

face. The block size S_B is therefore set at ($\alpha * DisEye$), where α is a coefficient and DisEye denotes the distance between the two eyes. If the block size increases to the width of the whole image, BHE will be the same as HE. This result is shown in the last column of Table 4.1. From the experimental results, a block size of ($0.5 * DisEye$) will give the best performance level. Therefore, in the rest of the experiments, the block size for BHE is also set at ($0.5 * DisEye$).

Table 4.1: BHE with different block sizes

α	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1.0**	**HE**
Recognition Rate (%)	**52.7**	**62.7**	**66.7**	**60.7**	**62.0**	**60.0**	**60.7**	**56.7**	**54.0**

We understand from the result of this paper HE and BHE is an important preprocessing step which influences greatly the recognition performance. We also noted that a better sophisticated HE method, called Adaptive Histogram Equalization (AHE), in which the interpolation between methods is more complex has even better effect. The influence of these two methods on the face images (Fig 4.1) can be seen in Fig 4.2(HE) and in Fig 4.3(AHE).

Fig 4.1: Face examples under different lighting conditions

Fig 4.2: Face examples under different lighting conditions after HE preprocessing

As we will see in the results section the HE or AHE are necessary steps and they improve the recognition rate. However, these methods can often lead to noise amplification in “flat” regions, and “ring” artifacts at strong edges.

Fig 4.3: Face examples under different lighting conditions after AHE preprocessing

4.3 New Approach: Iterative Method in Illumination Restoration

The major advantages of this algorithm [23] are that no facial feature extraction is needed and the generated face images will be visually natural. This method is a based on a general idea that the ratio of two images of the same person is simpler to deal with than directly comparing the different images of persons. It uses a ratio-image, which is the quotient between a face image whose lighting condition is to be normalized and a reference face image. The two images are blurred using a Gaussian

filter, and the reference image is then updated by an iterative strategy in order to further improve the quality of the restored face image. With this algorithm, a face image with arbitrary illumination can be restored to having frontal illumination. This algorithm has several advantages over those previous algorithms:

- In the training, only a single face image under frontal lighting is required.
- It does not need either to estimate the face surface normals and light source directions, or to know in advance the face surface normals and albedos.
- It does not need to perform image warping, as in the relighting method. Generally, it is difficult to detect the accurate positions of the feature points, especially when a face is under a poor lighting condition.
- The reconstructed face image with frontal illumination that is restored from an arbitrary illumination can be used by all the successful face recognition algorithms developed previously, while some approaches in the third category are possibly limited to a few specific methods.
- From human perception, the restored face image with frontal illumination has a good visual effect so that we can further inspect and discriminate different faces.

4.3.1 Algorithm Outline

In this approach all the assumptions are done by considering objects (faces) with Lambertian reflection:

$$I(x, y) = \alpha(x, y)n(x, y)^T s \tag{4.9}$$

where as stated before, where I(x, y) is the image intensity of the pixel at (x, y), α(x,y) the albedo of the surface at (x, y), n(x, y) the unit inward normal vector to the surface at (x, y), and s a column vector signifying the product of the light source intensity with the unit vector for the light source direction. Let I_{ik} denote a face image of the ith person captured under the s_kth light source direction, where a light source is classified according to its direction. I_{r0} represents a face image of another person captured under the frontal light source (s_0) and is used as a reference image. Let F be a Gaussian low-pass filter with $\sigma_x = \sigma_y = \sigma$ given by the following formula:

$$F(x, y) = \frac{1}{2\pi\sigma^2} e^{(x^2+y^2)^2/2\sigma^2} \tag{4.10}$$

Then, we give the two blurred images of I_{ik} and I_{r0} denoted as B_{ik} and B_{r0} respectively as follows:

$$B_{ik} = F * I_{ik} = F * (\alpha_i n_i^T s_k) = (F * \alpha_i n_i^T) s_k \tag{4.11}$$

$$B_{ro} = F * I_{ro} = F * (\alpha_r n_r^T s_0) = (F * \alpha_r n_r^T) s_0 \tag{4.12}$$

where (*) the convolution operation. As the shape and albedos of all faces are similar, if the size of F in (4.10) is big enough we have that $B_{i0} \approx B_{r0}$. This means that two faces under the same lighting conditions will be similar to each other after blurring. Therefore, we have:

$$F * a_i n_i^T \approx F * a_r n_r^T \tag{4.13}$$

By using the formulas (4.11) – (4.13), we can obtain the face image under frontal illumination for the *i*th person from I_{ik} captured under an arbitrary lighting direction s_k as follows:

$$H_{io} = \alpha_i n_i^T s_0 \approx \alpha_i n_i^T s_k \frac{\left(F * \alpha_r n_r^T\right) s_0}{\left(F * \alpha_i n_i^T\right) s_k} = I_{ik} \frac{B_{ro}}{B_{ik}} \tag{4.14}$$

H_{i0} is an estimation of I_{i0} or a restored image of I_{ik}. Instead of selecting randomly a face image as the initial reference image, the initial reference image is selected as the mean image of a number of face images captured under the frontal light source (i.e. the set of training images). By formulas (4.11) and (4.12), the larger the value of r (the size of the Gaussian filter) is, the truer appearance the restored face image will have. However, the edges caused by the shading in the face image will also be magnified. Moreover, as the filter size increases, the noise in the shading is also magnified and appears more in the restored image. A smaller initial filter size will require more iterations, while a larger initial filter size leads to a worse restored image if the face images have shading.

The eigenspace Φ_S introduced in section 2.2 is computed according and is used to remove some noise points in the initial restored image, as computed by Eq. (4.14). This initial restored image is projected onto the eigenspace using Eq. (2.5), and then a corresponding reconstructed image can be obtained using Eq. (2.6). Obviously, the

adverse effect from the noise points and edges from shading in the reconstructed image can be largely reduced.

Altogether this approach has been developed according to the following algorithm:

1. A mean face image and an eigenspace are computed by PCA based on a set of training images, which are all captured under frontal light source
2. An initial restored image can be calculated using Eq. (4.9)-(4.14), where the mean face image is used as the initial reference image and the size of the initial Gaussian filter F is 5. A reconstructed image is obtained from the initial restored image based on the eigenspace computed in step (1), and should have a smaller number of noise points.
3. An iterative procedure is used to obtain a final restored image with frontal illumination. During the iterative procedure, the reference image and the raw image are adaptively filtered based on the edge image obtained in step (4), and the reference image is also updated with the new reconstructed image so as to obtain a visually better restored image.
4. The iterative procedure continues until a stopping criterion is met. In this approach the stopping criteria is $|D(t)-D(t-1)|<\varepsilon$, or the exceeding of the maximum iterations number. The measure D(t) is calculated as follows:

$$D(t)=\frac{\left\|H_{io}(t)-H_{io}(t-1)\right\|}{\sqrt{\left\|H_{io}(t)\right\|\cdot\left\|H_{io}(t-1)\right\|}},\ t\geq 2 \tag{4.13}$$

Some examples of the outputs of this iterative are given in the Fig 4.4, Fig 4.5 and Fig 4.5. As we see the results are quite satisfactory and the images have been restored to having frontal illumination. This will be also examined further during the experiments done on different illumination angles.

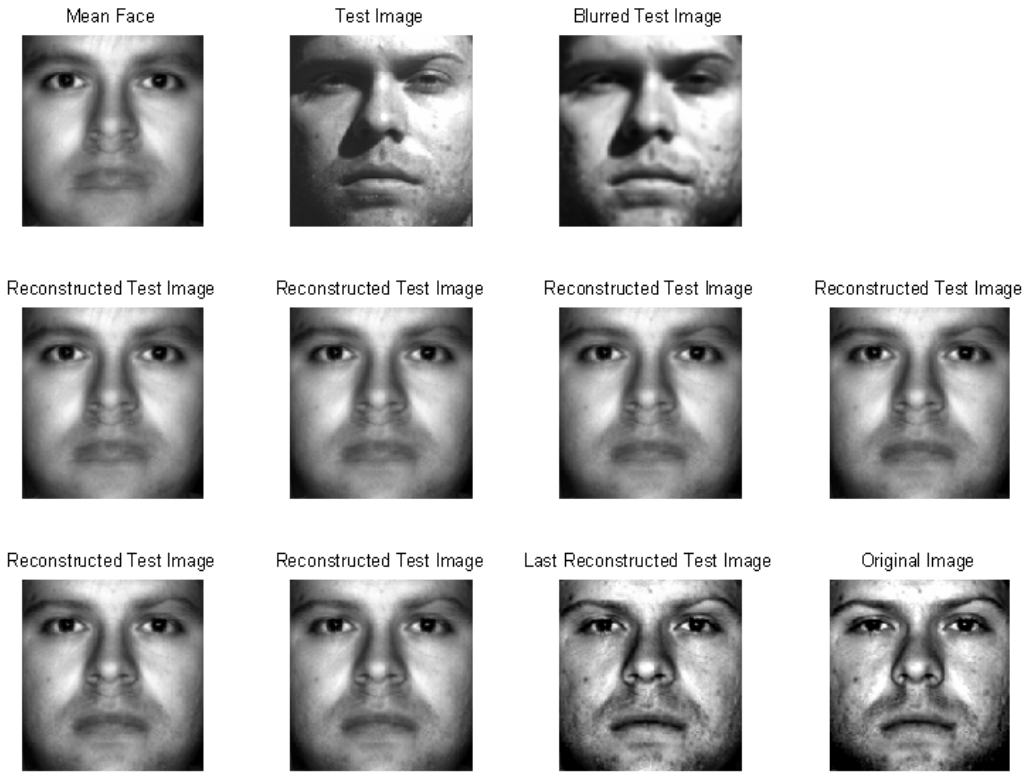

Fig 4.4: Illumination restoration steps with iterative procedure for person #1

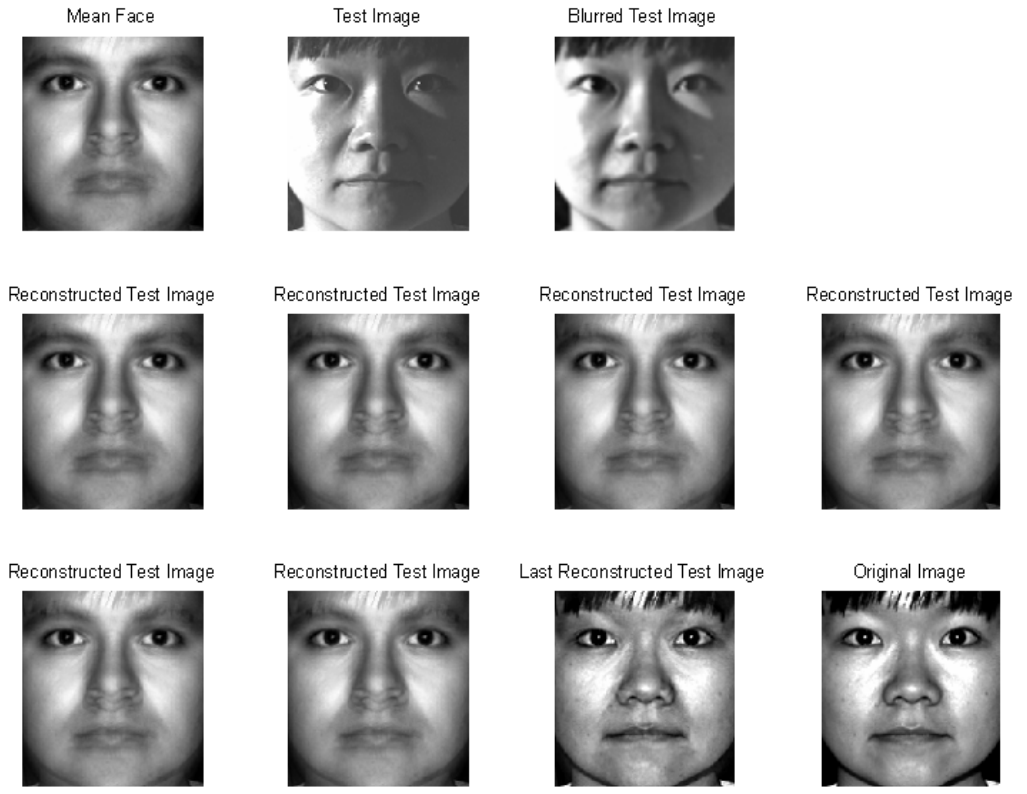

Fig 4.5: Illumination restoration steps with iterative procedure for person #2

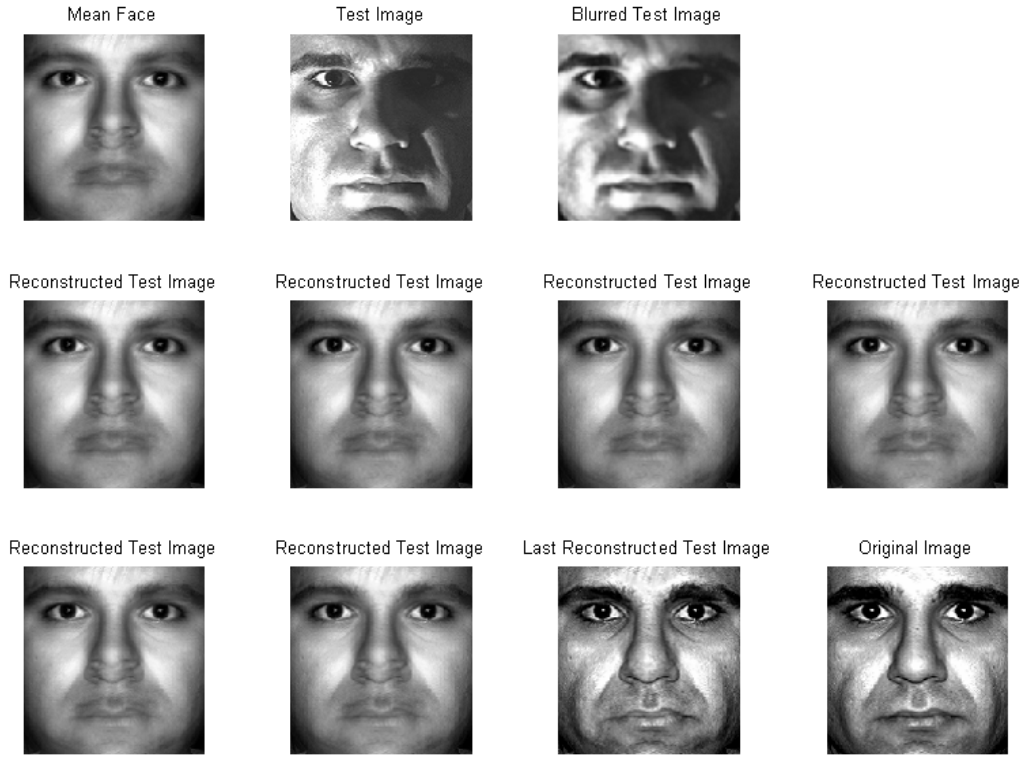

Fig 4.6: Illumination restoration steps with iterative procedure for person #3

In all these examples before the illumination restoration process, image normalization has been done in order to change the images to zero mean and unit standard deviation. This preprocessing step has a visually effect on the images and in the same time improves the performance of the PCA projection.

5. STUDY OVERVIEW AND EXPERIMENTAL RESULTS

In the previous sections we give the approaches that all standing alone provide performance improvement when integrated in FRS. Now we will make these parts interact altogether in order to create a robust system against varying illumination. All the experiments have been applied using the YaleB database and also a part of database of the Computer Engineering Department at Istanbul Technical University.

5.1 System Overview

As we said in the introduction section, the main aim of this work is to guarantee a successful recognition system when only one image per person is available during the training (learning) phase. For these PCA is a natural approach, but it has an important drawback: recognition of an object (or face) under a particular lighting can be performed reliably provided that object has been previously seen under similar circumstance. In other words, this method in its original form has no way of extrapolating to novel viewing conditions.

In order to surpass this problem, it has been shown that LDA-based approaches offer more sensitivity to recognizing of novel viewing conditions. But they create a new problem: it is needed more than one image per person during the training phase. To solve this problem we have introduced the Quotient Image for image synthesis. By using this method we can create the image space of any input image making possible the application of LDA even if there are offered “de facto” only one image per person. Furthermore we used a variant between F-LDA and D-LDA that exploits the best part of these two approaches and in the same time gives solution to the SSS problem without eliminating the null space of the image database.

As the last step of this we do the recognition process to any new input image by projecting to a dimension presented by the algorithm in section 2 and Euclidian distance is then used as distance classifier. In order to improve the performance any input image is preprocessed in order to restore the illumination to a frontal one.

5.2 Experimental Results

Most of the experiments have been done using the YaleB database. The YaleB database consists of 10 distinct persons. 45 images of each of these people were divided into four subsets. The first subset includes 70 images captured under light source directions within a span of 12◦ (only the first image is captured under a frontal light source). The second and third subsets have respectively 120 images and 140 images each, which were captured under light source directions within 25◦ and within 50◦, respectively. The fourth subset contains 140 images captured under light source directions within 75◦.

Fig. 5.1: Some cropped images from the Yale B database. The images are divided into four subsets according to the angle between the light source direction and the camera axis: (a) Subset1 (up to 12◦); (b) Subset2 (up to 25◦);(c) Subset3 (up to 50◦);(d) and Subset4 (up to 75◦)

In all the images the position of the two eyes of each face were located and translated to the same position. The images were cropped to a size of 180x190. Some images of these four subsets are illustrated in Fig. 5.1(images belonging to test person #5). In order to improve the performance of dimensionality reduction steps and recognition all the images were normalized to zero mean and unit variance. After that the algorithms were tested for both HE and AHE as a further preprocessing, with the second having a best performance. The images after normalization and AHE are shown in Fig 5.2.

Fig. 5.2: Some cropped images from the Yale B database after being normalized and AHE-processed

We continued our experiments to check whether these two preprocessing steps influences or not the recognition rate and we verified that they improve in a significant way the recognition when PCA was used for initially testing purpose. The results are given in Table 5.1.

Table 5.1: Results over the YaleB database (PCA used)

	No Preprocessing	HE	AHE
Recognition Rate (%)	43.4	74	81.5

In order to understand the way class based methods are affected by the size of the databases used during training we have tested various combinations under different circumstances. In all the cases the images used in training were left outside during recognition. We checked the results of both Fisher LDA and D-LDA with the second method having slightly better performance. As we from Table 5.2 the size of the database has a crucial impact on the performance of these methods. Despite this relation we have to admit that it is very uncasual that do many images would be available during the training phase. So image synthesis comes in front of us as a necessary step.

Table 5.2: D-LDA results for different size databases

	Subset 1	Subset 1	Subset 1	Subset 1	Total
30 images(%)	100	100	78.57	47.5	79.33
70 images(%)	100	100	90.71	51.66	84.22
270 images(%)	100	100	92.88	93.33	96.0

5.2.1 Quotient Image Results

In order to better understand the effectiveness of the algorithm a wide range of experiments have been conducted on the Quotient Image algorithm. Firstly it has been dealt with synthesizing new images from any arbitrary illuminated input image outside the YaleB database where the bootstrap consisted of 30 pictures of 10 persons of YaleB. Some of these results were previously shown in chapter 3. Furthermore we were interested in checking the algorithm performance when the bootstrap size was reduced. We checked for the performance when only 15 images

were available in the bootstrap (5 persons) with two different bootstrap combinations (Fig 5.3 (a) and (b)). In addition we also checked for the cases when 9 (3 person) images, 6 images (2 persons) and finally 3 images (1 person) were provided.

Table 5.3: Coefficient results for different bootstrap combinations

Coeff / Person #	5	5	3	2	1
X_1	0.11302	0.036265	0.23729	0.099189	0.15915
X_2	0.386478	0.42984	0.31587	0.39592	0.45989
X_3	0.41526	0.47669	0.353122	0.44	0.29723

In all the cases after calculating the Quotient Image by the formula $Q_y=y_s/Ax$ we than create the image space of the novel image by the product of product of Q_y and Az for all choices of z. The values of the calculated X variable are given in Table 5.3.

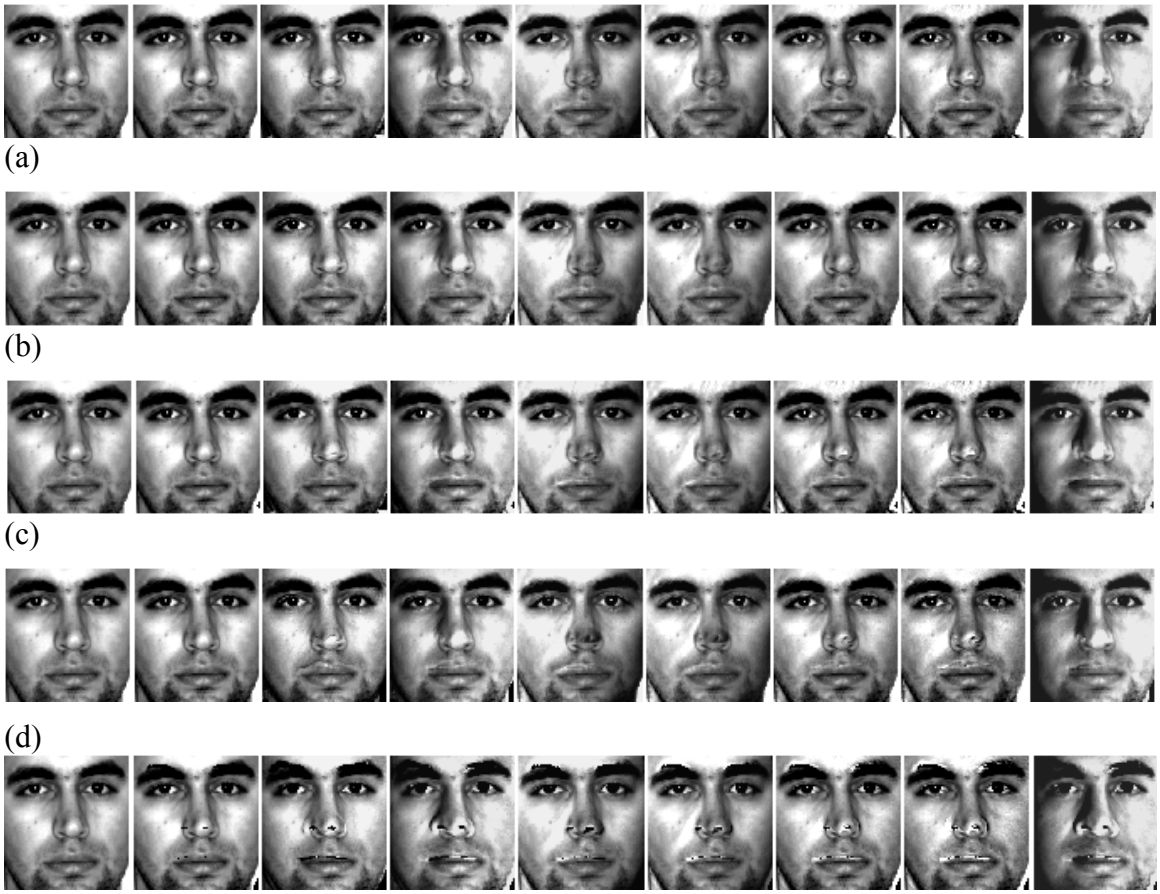

Fig 5.3: Image Re-rendering results for different bootstrap combinations: (a) 5 person; (b) 5 person; (c) 3 person; (d)2 person (e)1 person. The shown nine pictures have been reconstructed under the following conditions from left to right, original, original reconstructed(Table 5.2), [1 0 0], [0 1 0], [0 0 1], [-0.1 0.9 0.1], [-0.4 0.7 0.6], [-0.8 0.8 0.9], [0.5 -0.6 1]

According to these values we can create the image space of the input image by given values the z variable randomly or by some normal distribution around them. From Fig. 5.3 we can understand that a bootstrap consisting of 10 persons is quite consistent. Even when we reduce it by a half, the results are quite satisfactory. This is because as we reported before the albedos of possible faces occupy only a small part of the dimension in which they are spanned. Of course the larger the bootstrap size the more accurate will be the recovery of X and the Quotient. For small bootstrap sizes some noise points and changes in appearance features become visible. In order to eliminate this problem surface reconstruction can be thought as a possible solution.

In order to prepare the training set for the LDA process we create the image space of all 10 persons of the YaleB database, where 10 images per person were synthesized. For these we have used 15 images bootstraps where the object being reconstructed has been left outside. The results of using D-LDA with this artificially created database are given in Table 5.4.

Table 5.4: LDA Results with synthesized faces

	LDA	HE + LDA	AHE + LDA
Recognition Rate (%)	78.66	81.77	85.75

5.2.2 Reconstruction and System Results

The final step of this work is that any incoming image will be reconstructed in order to restore frontal illumination. Some output examples were shown previously in section 4. In the experiments over the YaleB database subsets, the results were almost identical to the frontal illuminated image for the first and second subset. The next two subsets, where the illumination conditions are worse, have also high performance but it has to be noted that some noise points and feature corruption became visible. This can be visible seen in Fig. 5.4 and Fig 5.5 where input and output has been given for the third and forth subset belonging to person no. 5. We see the upper-leftmost image shows signs of “aging” which is considered a feature corruption and is related with bad illumination conditions. In order to improve the visual appearance smaller values were given to controller of the iteration loop. This improved the appearance but in the same time increased greatly the iterations

number. One further point to observe is that even though some of the examples are corrupted an important factor is the convergence of the pictures to their belonging identity in the PCA space (also LDA) during the iteration steps.

(a)

(b)

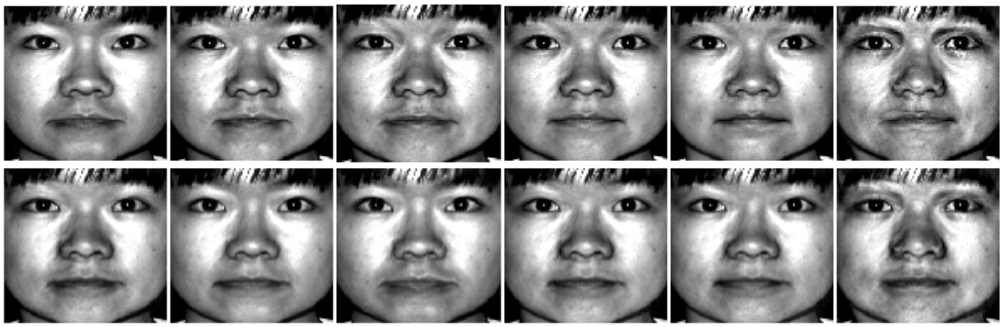

Fig 5.4: Illumination restoration for images of subset 3(up to 50°) (a) Before preprocessing and illumination restoration (b) After illumination restoration

(a)

(b)

Fig 5.5: Illumination restoration for images of subset 4(up to 70°) (a) Before preprocessing and illumination restoration (b) After illumination restoration

After the illumination restoration process the results of the whole approach were tested with 450 input images. Different distance measurements were experimented:

- Manhattan Distance (L1- norm)
- Cosine angle between two vector representations
- Euclidian Distance (L2 - norm)

Euclidian Distance gives the best results for this classification purpose. The general results of the new method are given in Table 55 (HE as preprocessing step) and Table 5.6 (AHE as preprocessing step).

Table 5.5: System Results with Histogram Equalization preprocessing

	Subset1	Subset2	Subset3	Subset4	Subset5
HE + PCA (%)	100	97.5	66.42	44.16	74
HE + New (%)	100	100	92.8	91.66	95.56

Table 5.6: System Results with Adaptive Histogram Equalization preprocessing

	Subset1	Subset2	Subset3	Subset4	Subset5
AHE +PCA (%)	100	100	83.57	50	81.55
AHE +New (%)	100	100	100	95	98.66

In order to further increase the recognition rate several combinations during the training phase have been applied. For the LDA step the best performance was achieved when 10 synthesized images per object were available during the training and all the discriminatory feature vectors of the LDA - projection matrix were used (9 features). In all the experiments the results were compared with PCA because it is the most important discriminatory technique used when only one image per person is available. Obviously this proposed work can significantly improve the recognition rates. Other methods that have acclaimed high recognition rates on the YaleB database, but either their computation burden is large or they require a large number of images per person.

6. CONCLUSION

A complex system for dealing with illumination was introduced. In this work was aimed the solution of SSS problem of class-based discrimination problems. For this a image space synthesis method was explained and the image space of the each image of the training set was created. In order to eliminate excessive noise points, application of surface reconstruction can be thought as a future work for the improvement of the system performance. After creating the image space of each training image a hybrid variant of F-LDA and D-LDA was applied in order to best use the discriminatory features of the system. The transformation projection matrix was applied to each image in the training set and 30 vectors (size 9x1 – all feature vectors used) were calculated.

Every incoming image was processed with illumination restoration algorithm and then projection was made in order to extract the discriminatory features. The general result over the YaleB database, consisting of 450 images, was 98.66% which can be considered very successful when comparing to existing methods. Another approach [16] achieved 100% recognition rates with all the data subsets, but seven images of each person in subset 1 has to be used to obtain the shape and albedo of a face, which in their own have computationally burden if to be considered the integration in a real time system.

As a conclusion, this three step study offers an innovative approach for creating a robust Face Recognition System under varying illumination. This study offers possibility of creating a real time system because it is not computationally complex. In future works, this study will be extended to dealing not only with upright frontal views but also with different poses. One possible solution based on this study is to apply multiple reference subspaces for different poses.

7. REFERENCES

[1] **Kelly, M. D.** 1970. “Visual identification of people by computer.” Tech. rep. AI-130, Stanford AI Project, Stanford, CA.

[2] **Kanade, T.** 1973. Computer recognition of human faces. Birkhauser, Basel, Switzerland, and Stuttgart, Germany.

[3] **Turk, M., Pentland,** A., 1991. “Eigenfaces for Recognition” Journal of Cognitive Neuroscience, Vol. 3, Num. 1, pp. 71-86.

[4] **P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman,** 1997."Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection" IEEE Trans. On PAMI, Vol. 19, No. 7.

[5] **B.Moghaddam and A.Pentland**, 1995. Probabilistic Visual Learning for Object Detection, Proc. Int'l Conf. Computer Vision, pp. 786-793.

[6] **T.F.Cootes, G.J.Edwards, C.J.Taylor**, 1998. “Active Appearance Models”, ECCV, vol.2, pp. 484-498.

[7] **G. Edwards, T. Cootes, and C. Taylor**, 1999. Advances in Active Appearance Models, Proc. Int'l Conf. Computer Vision, pp.137-142.

[8] **P.Penev and J.Atick**, 1996. “Local Feature Analysis: A General Statistical Theory for Object Representation,” Network: Computation in Neural Systems, vol.7, pp. 477-500.

[9] **L.Wiskott, J.M.Fellous, N.Kruger and C.V.D.Malsburg**, 1997. “Face Recogniton by Elastic Bunch Graph Matching”, IEEE Trans. On PAMI, 19(7), pp. 775-779.

[10] **P.J.Phillips, H.Moon**, 2000. “The FERET Evaluation Methodology for Face-Recognition Algorithms”, IEEE TPAMI, Vol.22, No.10, pp. 1090-1104.

[11] **G.Guo, S.Z.Li and K.Chan**, 2000. “Face Recognition by Support Vector Machines”, FG’02, pp. 196-201.

[12] **Y. Adini, Y. Moses, and S. Ullman**, 1997. Face recognition: The problem of compensating for changes in illumination direction. PAMI, 19(7), pp. 721–732.

[13] **W. Zhao, R. Chellappa**, 2000. SFS based view synthesis for robust face recognition, Proceedings of the 4th Conference on Automatic Face and Gesture Recognition.

[14] **A. Shashua, R.R. Tammy**, 2001. The quotient image: class-based rerendering and recognition with varying illuminations, IEEE Trans. Pattern Anal. Mach. Intell. 23 (2), pp. 129–139.

[15] **A. Shashua**, 1997. On Photometric Issues in 3D Visual Recognition from a Single 2D Image, Int. J. Computer Vision 21 (1–2), pp. 99–122.

[16] **A.S. Georghiades, P.N. Belhumeur, D.J. Kriegman**, 2001. From few to many: illumination cone models for face recognition under variable lighting and pose, IEEE Pattern Anal. Mach. Intell. 23 (2), pp. 643–660.

[17] **Yale University**, the YaleB database: http://cvc.yale.edu/projects/yalefacesB.

[18] **Juwei Lu, K.N. Plataniotis, A.N. Venetsanopoulos**, 2002. "Regularization Studies of Linear Discriminant Analysis in Small Sample Size Scenarios with Application to Face Recognition", Pattern Recognition Letter, in press.

[19] **A.S. Georghiades, P.N. Belhumeur, D.J. Kriegman**, 2003. Illumination-based Image Synthesis: Creating Novel Images of Human Faces Under Differing Pose and Lighting. Pattern Recognition.

[20] **P.N. Belhumeur, D.J. Kriegman**, 1997. What is the set of imagesof an object under all possible lighting conditions? In Proc. IEEE Conf. On Comp. Vision and Patt. Recog., pp 1040-1046.

[21] **R. Gross, S. Baker, I. Matthews, T. Kanade**, 2003. Face Recognition across pose and illumination.

[22] **Xie, K. M. Lam**, 2005. Face recognition under varying illumination based on a 2D face shape model , Pattern Recognition, Vol. 38, No 10, pp. 1705-1716.

[23] **D. H. Liu, K. M. Lam, L. S. Shen**, 2005. Illumination invariant face recognition, Pattern Recognition, Vol.38, No 2, pp. 221-230.

8. BIOGRAPHY

Erald VUÇINI has received his B.Sc in Computer Engineering from Yildiz Technical University , Istanbul, Turkey in June 2004 with highest honour and as the first of the departement of Computer Engineering. He has received his M.Sc. in Computer Engineering from Istanbul Technical University with high honours.

He has a continous interests in topics related with mathematics. He has winned for two years consecutively the Albanian National Mathematical Olympiad. He has participated in many international activities being also awarded the Bronze Medal in the Balkan Mathematical Olympiad. Due to his excellent results he has been awarded the Golden Medal of Albania and a scholarship for studies in Turkey.
Currently, he has been admitted as a Ph.D. student in Vienna Technical University, Vienna, Austria.