

**İSTANBUL TECHNICAL UNIVERSITY ★ INFORMATICS INSTITUTE**

**NEAREST NEIGHBOR DISCRIMINANT ANALYSIS  
BASED FACE RECOGNITION  
USING ENSEMBLED GABOR FEATURES**

**M.Sc. Thesis by  
Onur DOLU**

**Department : Advanced Technologies**

**Programme : Computer Sciences**

**JUNE 2009**



**İSTANBUL TECHNICAL UNIVERSITY ★ INFORMATICS INSTITUTE**

**NEAREST NEIGHBOR DISCRIMINANT ANALYSIS  
BASED FACE RECOGNITION  
USING ENSEMBLED GABOR FEATURES**

**M.Sc. Thesis by  
Onur DOLU**

**Date of submission : 4 May 2009  
Date of defence examination: 5 June 2009**

**Supervisor (Chairman) : Prof. Dr. Muhittin GÖKMEN (ITU)  
Members of the Examining Committee : Prof. Dr. Bilge GÜNSEL (ITU)  
Assoc. Prof. Dr. Zehra  
ÇATALTEPE(ITU)**

**JUNE 2009**



**İSTANBUL TEKNİK ÜNİVERSİTESİ ★ BİLİŞİM ENSTİTÜSÜ**

**PARÇALI GABOR ÖZNİTELİKLERİ KULLANARAK  
EN YAKIN KOMŞU AYRIŞIM ANALİZİ TABANLI  
YÜZ TANIMA**

**YÜKSEK LİSANS TEZİ  
Onur DOLU  
(704061007)**

**Tezin Enstitüye Verildiği Tarih : 4 Mayıs 2009  
Tezin Savunulduğu Tarih : 5 Haziran 2009**

**Tez Danışmanı : Prof. Dr. Muhittin GÖKMEN (ITU)  
Diğer Jüri Üyeleri : Prof. Dr. Bilge GÜNSEL (ITU)  
Doç. Dr. Zehra ÇATALTEPE (ITU)**

**Haziran 2009**



## **FOREWORD**

I would like to express my deep appreciation and thanks for my advisor, Prof. Muhittin GÖKMEN.

Special thanks to my family, for their patience and continuous support during my master study and during the preparation of this work.

And finally, I would like to thank Yu Su, Kadir Kırtaç, Fatih Kahraman and Abdulkerim Çapar for valuable discussions and for sharing their knowledge with me.

June 2009

Onur DOLU

Computer Engineer





## TABLE OF CONTENTS

	<u>Page</u>
<b>ABBREVIATIONS</b> .....	viii
<b>LIST OF TABLES</b> .....	ix
<b>LIST OF FIGURES</b> .....	x
<b>LIST OF SYMBOLS</b> .....	xi
<b>SUMMARY</b> .....	xiii
<b>ÖZET</b> .....	xv
<b>1. INTRODUCTION</b> .....	<b>1</b>
1.1 Image Based Face Recognition.....	4
1.2 Face Recognition Pitfalls.....	6
1.2.1 Illumination.....	7
1.2.2 Pose.....	8
1.2.3 Expressions.....	9
1.2.4 Occlusion and Facial Hair.....	10
1.2.5 Aging.....	12
1.3 Images As Vectors.....	14
<b>2. FACE RECOGNITION USING GABOR REPRESENTATION</b> .....	<b>15</b>
2.1 Preface.....	15
2.2 2D Gabor Filters.....	15
2.3 2D Gabor Filtered Representation of Images.....	18
<b>3. APPEARANCE BASED FACE RECOGNITION</b> .....	<b>21</b>
3.1 Linear Subspace Analysis.....	21
3.2 PCA.....	21
3.3 ICA.....	24
3.4 LDA.....	25
3.5 NNDA.....	27
3.5.1 The Criterion.....	30
3.5.2 Stepwise Dimensionality Reduction.....	33
3.5.3 NNDA to k-NNDA.....	33
3.5.4 Small Size Problem.....	34
3.6 Similarities, Distances and Correlations.....	35
<b>4. NNDA BASED FACE RECOGNITION USING ENSEMBLED GABOR FEATURES</b> .....	<b>39</b>
4.1 Preface.....	39
4.2 Previous Work On Gabor Features.....	39
4.3 EGNNC.....	41
4.3.1 Grouping the multiple Gabor feature segments.....	42
4.3.2 Designing a NNDA classifier based on each feature segment.....	43
4.3.3 Combining all these component with NNDA classifiers.....	45
<b>5. EXPERIMENTS AND RESULTS</b> .....	<b>47</b>
5.1 Results on Yale Database.....	47
5.2 Results on FERET Subset.....	48
<b>6. CONCLUSIONS AND FUTURE WORK</b> .....	<b>53</b>
<b>REFERENCES</b> .....	<b>55</b>
<b>CURRICULUM VITA</b> .....	<b>59</b>

## ABBREVIATIONS

<b>SVD</b>	: Singular Value Decomposition
<b>PPLS</b>	: Parametric Piecewise Linear Subspace
<b>PCA</b>	: Principal Component Analysis
<b>ICA</b>	: Independent Component Analysis
<b>LDA</b>	: Linear Discriminant Analysis
<b>GDA</b>	: Generalized Discriminant Analysis
<b>NNDA</b>	: Nearest Neighbor Discriminant Analysis
<b>NDA</b>	: Nonparametric Discriminant Analysis
<b>NLDA</b>	: Null Space Discriminant Analysis
<b>EBGM</b>	: Elastic Bunch Graph Matching
<b>KCCA</b>	: Kernel Canonical Correlation Analysis
<b>GPC</b>	: Gabor Principle Classifier
<b>GFC</b>	: Gabor Fisher Classifier
<b>GNNC</b>	: Gabor Nearest Neighbor Classifier
<b>EGNNC</b>	: Ensembled Gabor Nearest Neighbor Classifier

## LIST OF TABLES

	<b><u>Page</u></b>
<b>Table 1.1:</b> “The main information about the experimental results of most of the discussed methods” [5] .....	13
<b>Table 5.1:</b> Average recognition rates of GPC, GFC, GNNC and EGNNC on 200 class in subset .....	50
<b>Table 5.2:</b> Step size affect on EGNNC, actually vey small affect.....	51

## LIST OF FIGURES

	<u>Page</u>
<b>Figure 1.1:</b> Scattering of the most used biometrics in commercials .....	2
<b>Figure 1.2:</b> Comparison of various biometric features .....	2
<b>Figure 1.3:</b> General flow of a face recognition system .....	4
<b>Figure 1.4:</b> General layout of the face recognition methods.....	6
<b>Figure 2.1:</b> 3-D visualization of a Gabor kernel. ....	17
<b>Figure 2.2:</b> 2-D Gabor kernels of 5 scales and 8 orientations.....	18
<b>Figure 2.3:</b> Gabor filter representation(the real part and the magnitude) of a 64x64 sample image from ORL database. (a) The real part of the representation and (b) The magnitude part of the representation. ....	20
<b>Figure 2.4:</b> An example of the augmented Gabor feature vector (in image form rather than in vector form).....	20
<b>Figure 3.1:</b> Principal components (PC) of a 2D set of points .....	22
<b>Figure 3.2:</b> Face samples from the ORL face database .....	23
<b>Figure 3.3:</b> The mean face. ....	23
<b>Figure 3.4:</b> Eigenvectors corresponding to the 7 largest eigenvalues, shown as $p \times p$ images, where $p \times p = n$ .....	23
<b>Figure 3.5:</b> Eigenvectors corresponding to the 7 smallest eigenvalues shown as $p \times p$ images, where $p \times p = n$ .....	23
<b>Figure 3.6:</b> Properties between PCA and ICA .....	24
<b>Figure 3.7:</b> ICA basis vectors shown as $p \times p$ images.....	25
<b>Figure 3.8:</b> A comparison of PCA and FDA.....	25
<b>Figure 3.9:</b> First seven FDA basis vectors shown as $p \times p$ images .....	26
<b>Figure 4.1 :</b> Illustration of the basic idea of EGNNC	42
<b>Figure 5.1:</b> 2 subjects from Yale Database with expression, lighting variations and occlusion(glasses).	47
<b>Figure 5.2:</b> Performance of EGNNC on Yale as reduced dimension is 14 and step size 5.....	48
<b>Figure 5.3:</b> Example images used in the FERET experiments. ....	49
<b>Figure 5.4:</b> Comparison between GPC, GFC, GNNC, EGNNC.....	50

## LIST OF SYMBOLS

$\mathbf{T}(\mathbf{r})$	: Linear transformation function
$\mathbf{N}$	: Number of training images
$\mathbf{x}_i$	: $i$ -th training sample in a training set
$\mathbf{W}$	: Linear transformation matrix
$\mathbf{y}_k$	: $k$ -th projected training sample in a training set
$\mathbf{m}$	: Mean vector of the training images
$S_t$	: Total scatter matrix of the training set
$\lambda$	: Eigenvalues of the covariance matrix
$\mathbf{V}$	: Eigenvectors of the covariance matrix
$\beta_i$	: $i$ -th difference image in Eigenfaces approach
$\mathbf{c}$	: Number of classes in Fisherfaces approach
$\mathbf{N}_i$	: Number of training samples in $i$ -th class of the training set
$\mathbf{m}_i$	: Mean vector of training images belonging to $i$ -th class
$\mathbf{S}_b$	: Between-class scatter matrix of Fisherfaces approach
$\mathbf{S}_w$	: Within-class scatter matrix of Fisherfaces approach
$x_n^E$	: Extra-class nearest neighbor of the training sample $x_n$
$x_n^I$	: Intra-class nearest neighbor of the training sample $x_n$
$\Delta_n^E$	: Nonparametric extra-class distance of the training sample $x_n$
$\Delta_n^I$	: Nonparametric intra-class distance of the training sample $x_n$
$S_B$	: Nonparametric between-class scatter matrix of NNDA method
$S_W$	: Nonparametric within-class scatter matrix of NNDA method
$W_n$	: Weighting parameter for the training sample $x_n$ in NNDA
$\Theta_n$	: Accuracy of nearest neighbor classification for sample $x_n$ in NNDA
$\mathbf{k}$	: Nearest neighbor parameter of $k$ -nearest neighbor classification
$x_{([k/2]+1)}^I$	: The intra-class $([k/2]+1)$ -th nearest neighbor of the sample $x$
$x_{[k/2]}^E$	: The $[k/2]$ -th extra-class nearest neighbor of the sample $x$
$\Psi_{\mu,\nu}$	: Gabor kernel with scale $\mu$ and orientation $\nu$
$\sigma$	: Standard deviation of the Gaussian part of the Gabor kernel
$f$	: Spacing factor between Gabor kernels
$k_{\mu,\nu}$	: Wave vector of the Gabor kernel with scale $\mu$ and orientation $\nu$
$J^I$	: Gabor jet representation of image $I$
$X^{(\rho)}$	: Augmented Gabor feature vector with a downsampling factor $\rho$



# **NEAREST NEIGHBOR DISCRIMINANT ANALYSIS BASED FACE RECOGNITION USING ENSEMBLED GABOR FEATURES**

## **SUMMARY**

Our notably ability as recognizing people is a primary concern of human activities in everyday life. For recognition, some neurons in the temporal lobe of the brain respond to particular face features obtained from signals of eyes.

Having discriminative information of face individuality, face is the most challenging task for image analysis and object recognition. And face recognition process is stated as given still or video images of a scene, identify or verify one or more persons in the scene using a stored database of faces.

In last decades, Gabor features based face representation performed very promising results in face recognition area as its robust to variations due to illumination and facial expression changes. The properties of Gabor are, which makes it effective, it computes the local structure corresponding to spatial frequency (scale), spatial localization, and orientation selectivity and no need for manual annotations.

The contribution of this thesis, an Ensemble based Gabor Nearest Neighbor Classifier (EGNNC) method is proposed extending Gabor Nearest Neighbor Classifier (GNNC) where GNNC extracts important discriminant features both utilizing the power of Gabor filters and Nearest Neighbor Discriminant Analysis (NNDA). EGNNC is an ensemble classifier combining multiple NNDA based component classifiers designed respectively using different segments of the reduced Gabor feature. Since reduced dimension of the entire Gabor feature is extracted by one component NNDA classifier, EGNNC has better use of the discriminability implied in reduced Gabor features by the avoiding 3S (small sample size) problem as making minimum loss of discriminative information. The EGNNC is constructed by the following steps; dividing face image into  $M$  sub images according to their spatial locations, computing augmented Gabor features for each sub-image, grouping the augmented Gabor features to multiple feature segments, designing a NNDA classifier based on each feature segment, combining all these component NNDA classifiers by certain combination rules.

The accuracy of the EGNNC is shown by comparative performance work. Using a 200 class subset of FERET database covering illumination and expression variations, EGNNC achieved 100% recognition rate, outperforming its ancestor GNNC perform 98 percent and standard methods such as GFC and GPC for 65 features. Also for the YALE database, EGNNC outperformed GNNC on all  $(k, \alpha)$  tuples and EGNNC reaches 96 percent accuracy in 14 feature dimension, along with parameters step size = 5,  $k = 5$ ,  $\alpha = 3$ .





## PARÇALI GABOR ÖZNETELİKLERİ KULLANARAK EN YAKIN KOMŞU AYRIŞIM ANALİZİ TABANLI YÜZ TANIMA

### ÖZET

En önemli yeteneklerimizden biri olan kişileri tanımak, günlük yaşamdaki aktivitelerimizin temel hususudur. İnsanları tanımak için, beyindeki temporal lob adı verilen bölgedeki sinirler gözden gelen sinyalleri kullanarak belli yüz özelliklerine cevap verir.

Her yüzün kendine has ayırt edici özelliklerinden dolayı yüz tanıma, görüntü analizi ve obje tanıma alanlarında için en ilgi çekici olanıdır. Ve yüz tanıma işlemi şu şekilde açıklanır; durağan görüntü veya bir video görüntüsü verildiğinde bu görüntü içindeki kişi veya kişileri, yüz görüntüleri ile kaydedilmiş veri tabanı kullanarak tanıma veya doğrulama işlemidir.

Son yıllarda, ışık varyasyonlarına ve yüz ifade değişikliklerine karşı gürbüz olduğu üzere yüz tanıma alanında Gabor öznitelikleri tabanlı yüz temsil etme çok umut vaad edici sonuç vermiştir. Seçilen uzamsal frekans, uzamsal lokalizasyon ve yönelime göre yerel yapıyı hesaplaması, elle işaretlendirmeye ihtiyaç duymaması Gabor özniteliklerini efektif yapan özellikleridir.

Bu tez çalışmasındaki katkı, Gabor süzgeçleri ve En Yakın Komşu Ayrışım Analizi'nin (EYKAA) güçlerini birleştirerek önemli ayrışım öznitelikleri ortaya çıkaran Gabor En Yakın Komşu Sınıflandırıcısı (GEYKS) genişletip Parçalı Gabor En Yakın Komşu Sınıflandırıcısı (PGEYKS) metodunu ortaya koymaktır. PGEYKS; alçaltılmış gabor öznitelikleri barındıran farklı segmanları kullanarak, her biri ayrı dizayn edilen birçok EYKAA tabanlı bileşen sınıflandırıcılarını bir araya getiren grup sınıflandırıcısıdır. Tüm gabor özniteliklerinin alçaltılmış boyutu tek bir EYKAA bileşeninden çıkarıldığı gibi, PGEYKS; ayrışım bilgi kaybını minimum yapıp 3S (yetersiz örnek miktarı) problemini önleyerek alçaltılmış gabor öznitelikleri içindeki ayrıştırılabilirliği daha iyi kullanır. PGEYKS şu adımları izleyerek düzenlenir; yüz görüntüsü uzamsal konuma göre M alt görüntüye bölünür, her bir alt görüntünün ardı ardına eklenmiş Gabor öznitelikleri hesaplanır; eklenmiş Gabor öznitelikleri çoklu öznitelik segmanları şeklinde gruplanır; her bir segman için EYKAA sınıflandırıcısı tasarlanır; tüm bu EYKAA sınıflandırıcı bileşenleri belli bir kurala göre bir araya getirilir.

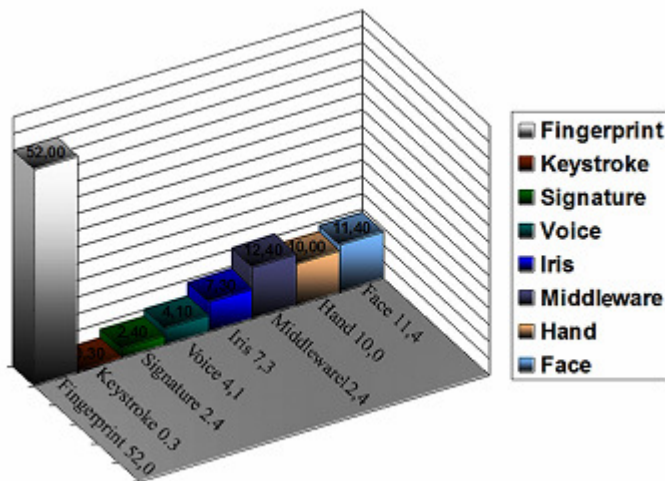
PGEYKS yönteminin tanıma başarımı karşılaştırmalı performans çalışması ile gösterilmiştir. Farklı ışıklandırma ve yüz ifadesi değişiklikleri barındıran 200 sınıflık FERET veritabanı alt kümesinde, 65 öznitelik için PGEYKS %100 başarımla elde ederek atası olan GEYKS'nin aldığı %98 başarımla ve diğer GFS (Gabor Fisher Sınıflandırıcı) ve GTS (Gabor Temel Sınıflandırıcı) gibi standard methodlardan daha iyi sonuçlar vermiştir. Ayrıca YALE veritabanı üzerindeki testlerde PGEYKS her türlü (k, alpha) çiftleri için GEYKS'ten daha başarılıdır ve 14 öznitelik için step size = 5, k = 5, alpha = 3 parametrelerinde %96 tanıma başarısına ulaşmıştır.



## 1. INTRODUCTION

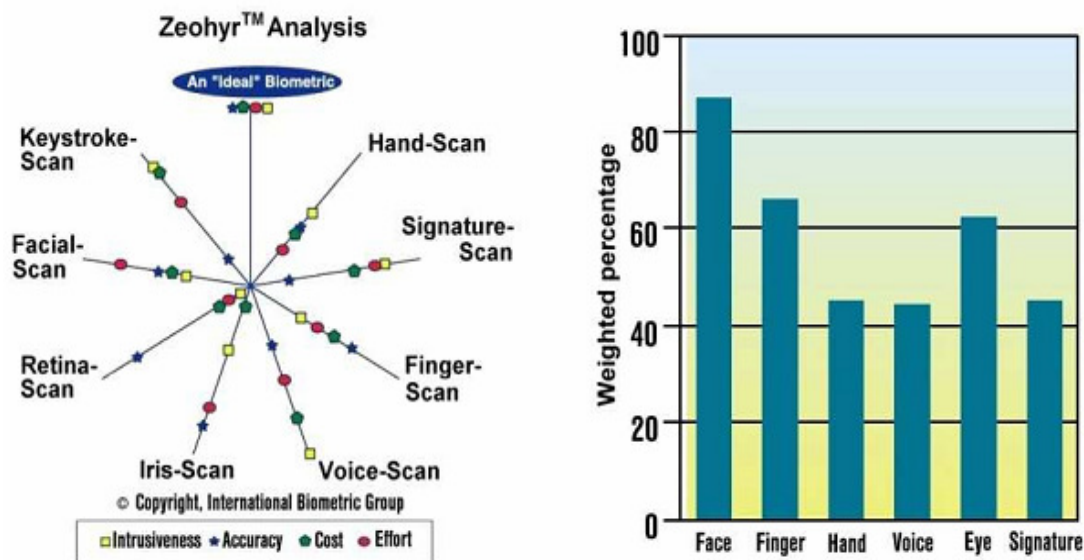
In past years, having discriminative information of face individuality, researchers in biometrics, pattern recognition, and computer vision communities have considerably got their attention on face recognition as a biometric identification application [1][2][3][4]. The communities of machine learning and computer graphics are also more and more concerned of face recognition. Our notably ability as recognizing people which is a primary concern of human activities in everyday life or in cyberspace is the motivation of this common interest among researchers working in different biometric fields. As well, face recognition technologies is required by the lots of of commercial, security, and forensic applications. Automated crowd surveillance, security access control, mugshot identification (e.g., for issuing driver licenses, airport visa control), face reconstruction, design of human interface devices for personal computers , multimedia communication and content-based image database management are the basic applications in face recognition technologies.

Finger prints, palm, voice, signature, face and many other human characteristics have been studied and possibly the most widespread biometrics are finger prints and iris, but Figure 1.1 [5] shows the scattering of the most used biometrics in commercials in the last decades. However, each biometric has some disadvantages. Iris recognition has precisely good results, but expensive to deploy and people do not accept much. Fingerprints has consistent results and are functional, but not appropriate for non-collaborative individuals. Therefore, between reliability and public approval, security and privacy, face recognition is good to be finding the middle ground. Its true that face recognition technology has a number of perils to social rights. At first, when false positives are analyzed, the identification system encroaches the privacy of people not to be recognized. Secondly, face template images could be stolen and could not be restored or could be restored with uncorrect one. So, using of face recognition technologies require lots of commercial and security applications. Face recognition has large benefits for places with large of unaware visitors where supplies a minor security level in unrestricted environment.



**Figure 1.1:** Scattering of the most used biometrics in commercials [5].

Also, face image is an effective biometric input. According to identification applications different kinds of biometric indicators are measured as intrusiveness, accuracy, cost, and ease of sensing [6] (see Figure 1.2 (a)). The six biometric indicators considered in [Url-1], facial features have the highest compatibility. Figure 1.2 (a) shows the scoring after a number of evaluation factors in a Machine Readable Travel Documents (MRTD) based system [Url-1].



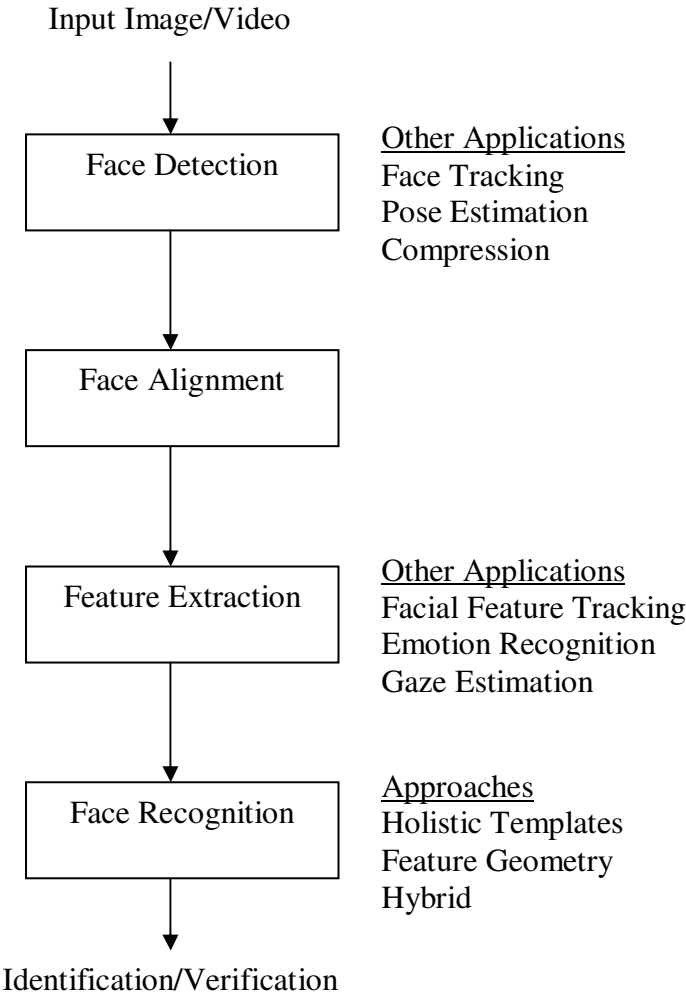
**Figure 1.2:** Comparison of various biometric features

Scenarios of face recognition technologies can be divided into two as face verification and face identification. And National Institute of Standards and Technology (NIST) added another scenario called the 'watch list' in the Face Recognition Vendor Test (FRVT) 2002 [7].

- Face verification is the 1:1 matching with the query of comparison between a face template and the face image whose identity is being processed. Performance of verification is measure with the ROC curve calculated by the true positive rate (valid users access allowed) over false positive rate (invalid users access allowed). The curve should have steadiness according to application needs.
- Face identification is the 1:N matching with the query of comparison between all the face templates in face database and the face image whose identity is being processed. Selecting the highest similarity/lowest distance between the image in the database and the test image is the identification of the test image. The identification process is a internal test and starts as sensor shots an example of an individual which is known as already in the database. All the other individuals face features in the system's database are compared with the test face images (normalized) features and a similarity/distance score is calculated for each comparison. These similarity/distance scores are sorted in descending/ascending order. "Top match score" of the test face image is the highest similarity score/lowest distance score for all individuals. "The cumulative match" is the correct match where the majority of the top r similarity/bottom r distance scores refers to the test subject.
- The watch list method does not require the test individual be in the system database. The other individuals in the system's database is compared with face image whose identity is being processed and a similarity/distance is reported for each individuals. These similarity/distance scores are sorted descending/ascending order so that the highest/lowest similarity/distance score is first. Alarm is raised, if a similarity/distance score is higher/lower than a predefined threshold. The system identifies the individual is stored in the system's database when an alarm is raised. Watch list applications have two main domains. The first is "Detection and Identification Rate", the percentage of times the system alarms and then recognize the face image correctly on the watchlist whose identity is being processed. The second is "False Alarm Rate", the percentage of times the system alarms for a person not in the watchlist.

**1.1 Image Based Face Recognition**

The general flow of a face recognition system can be listed as follows: taken a still image or a video of a scene, identify face or faces belongs to whom using a stored face database. To shrink the search, additional information can be used such as age, gender, facial expression or speech. The system has three main processes as segmentation of faces (face detection), feature extraction from the face regions, recognition as identification/verification (Figure 1.3). The problems of identification and verification are the response of system whether the determined individual is from known individual in the database, confirmation or rejection of the system of the face image whose identity is being processed.



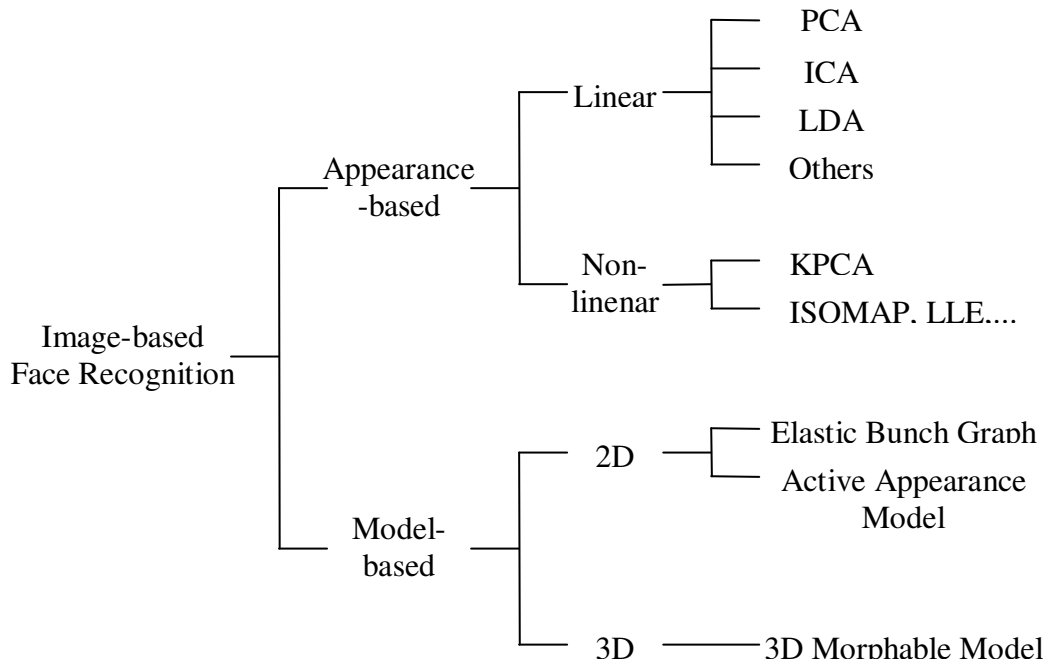
**Figure 1.3:** General flow of a face recognition system

Currently, there are three image-based face recognition techniques based on the face representation: (i) appearance-based which uses holistic texture features; (ii) model-based which utilize face shape and texture, (iii) hybrid methods which takes some parts of holistic methods and model based method into account for recognition.

(1) Appearance-based: Whole face region is used as raw input to the recognition system. Most of these techniques induce face region to a vector space structure and, in principle, necessitate dense correspondence in terms of alignment. An image is considered as a high-dimensional vector and a point in a high-dimensional vector space. Many view-based methods use statistical techniques to analyze the distribution of the image vectors in the vector space, and derive an efficient and effective representation in the feature space. This image vector representation also allows and for the synthesis of images. Eventually, face recognition can be threatened as a space-searching problem within machine-learning problem. And, eigenfaces is most generally used representation for face region [Kirby and Sirovich 1990; Sirovich and Kirby 1987], that is based on principal component analysis.

(2) Feature-based (structural) matching methods: Typically, these methods first extracts local features such as the eyes, nose, and mouth are and also their locations are extracted. The system constructs a model of the human face with extracted these facial features by using local statistics (geometric and/or appearance). The prior knowledge of human face is highly applied in designing the model. For example, distance and relative positions of internal facial element placements (e.g., the distance between two pupils is equal with mouth width, etc.). Kanade [8] developed one of the earliest face recognition algorithms based on automatic feature detection. Parameters for each face, computed as localizing the corners of the eyes, nostrils, etc. in frontal views, were compared against the parameters of known faces using Euclidean distance metric. Wiskott et al. [9] developed a more recent feature-based system based on elastic bunch graph matching which is an extension to graph matching [10]. Cootes et al. [11][12] combined both shape and texture which the face variations are learned, developed a 2D morphable face model. Also, 3D depth information could fed into model. Blanz et al. [13] proposed a method encoding the shape and texture as 3D morphable face model parameters, and recover these parameters from a single face image.

(3) Hybrid methods: Hybrid approaches use both holistic and local features. A machine recognition system should use both local features and the whole face region to recognize a face, just as the human perception system. One can argue that these methods could potentially put forwards better then the two types as appearance and feature methods.



**Figure 1.4:** General layout of the face recognition methods

Figure 1.4 shows general layout of the face recognition methods as leaves as examples methods.

## 1.2 Face Recognition Pitfalls

In last decade, many systems are able to reach greater than 90% recognition rates with proposed new methods in face recognition. However, face recognition remains as a challenging problem where face capture process could go through lots of variations in real-world environment due to;

- Illumination (including indoor/outdoor)
- Pose
- Expressions



- Occlusions (e.g., glasses, sunglasses) and Facial hair (e.g., beards, moustache, maybe exaggerated makeup)
- Aging

### **1.2.1 Illumination**

Ambient illumination becomes different significantly day to day and between indoor and outdoor environments. Because of the 3D formation of the face, a direct illumination source can spread strong shadows that highlight or cloud definite facial appearance. It has been revealed empirically and systems derived from Principal Component Analysis that differences in appearance caused by lighting are bigger than differences between persons. From the time when dealing with illumination variation is a common subject in computer vision, several approaches for enlightenment face recognition have been published. In (Adiniet al., 1997), look into the way, in which modifications in light can change effects of some face recognition techniques. They classify three different groups to rank the techniques: the shape of shadow, which pulls out the form of the face, from one or more of its sights, the demonstration techniques, which aim to get a portrayal of the face and the generative techniques, which generate an extensive set of mock metaphors including as much as possible alternates. The authors figured out that none of the tested method (edge map, 2D Gabor Filters, first and second derivatives of the gray level images) is able to solve the problem by itself and the outcomes they report looks like validating this suggestion.

In spite of this, numerous attempts have been made to reach improved performances in uncontrolled circumstances. In fact, Gao and Leung (2002) expanded the edge map method defining a new approach, called the Line Edge Map, in which the face curves are extracted and merged in fragments, which are then arranged in lines. The Hausdorff distance has also been revised to control these new characteristic vectors. Moreover, they also define a new percolating standard for monitoring the entire set of characters before presenting the actual test function. The technique has been tested on numerous circumstances for pretense and lighting and the results confirm that this approach surpasses other techniques, such as Linear Subspaces or Eigenfaces, presented in (Belhumeur et al., 1997).

On the other hand, the Fisherfaces continue to be better thanks to their facility to maximizing the between-individual changeability, minimizing the within-individual variations. This proposes that merging various linear techniques, presentations can be more developed. In fact, an in-depth revision on the presentations of the linear techniques when changes in lighting happen has been performed by Li et al. (2004). The inspected methods have been evaluated with respect to both recognition rate and time/memory complexity. The authors monitor that the LDA combined with a simplification of the SVD (Singular Value Decomposition), surpasses all the other techniques. However this synthesis is less adjustable to common face recognition issues, considering its computational outlay. For that reason, the authors recommend that to combine the LDA with the QR decay could stand for the best possible option in most cases, since it presents almost similar performances with the LDA/SVD approach with a minor outlay.

### **1.2.2 Pose**

The pose of the (probe) test and (gallery) train images is different in many face recognition scenarios. For example, the (probe) test image might be a 3/4 view recorded from a camera in the corner of a room where the (gallery) train image generally a frontal. According to the type of train (gallery) images used methods dealing with pose variation can be classified into two main categories.

Multi-view face recognition, necessitate train (gallery) images of every individual at every pose, is a direct extension of frontal face recognition. In face recognition the problem is to build algorithms to recognize a face from a novel pose that is has not previously been observed. Dealing with the problem of pose changes, linear subspaces have been extended. A method for recognizing faces with large 3D pose variations is presented by Okada and von der Malsburg (2002), taking means of parametric linear subspace model for representing each known person in the train. The authors study two different linear models:

(1) the LPCMAP model, which is a parametric linear subspace model, relates projection coefficients of training samples onto the subspaces and their corresponding 3D head pose, by integrating the linear subspaces spanned by principal components (PCs) of training samples and the linear transfer matrices;

(2) the PPLS model, a set of local linear models, each providing continuous analysis and synthesis mappings, generalizing to unknown poses by interpolation, extends the LPCMAP by using the piecewise linear approach.

If gallery covers a large 3D head pose variations as 50 degree rotation along each axis, the experimental results show that the recognition system is robust enough. The PPLS system performed better than the LPCMAP system also compressing the data size. However, some artificial samples and the small number of known people might accidentally increase the performance. Another drawback is that the representing facial shape and deriving head pose information requires pixel-wise landmark locations, but locating landmark in static facial images with arbitrary head pose is an ill-posed problem.

Then, to achieve a more robust and stable results solving the problem of pose variation in face recognition, Gross et al. (2002) proposed to the light-field approach. The light-field specifies the radiance of light in free space with 5D function of position (3D) and orientation (2D). Especially, the researchers apply the PCA to get a set of eigen light-fields with a collection of light-fields of different subjects' faces, also the mean light-field could also be calculated from all of the light-fields. Hence, a curve in the light-field represents an image of the object. For this highly occluded light-field curve that the objects' eigen coefficients can be calculated, especially for with simple reflectance properties, such as Lambertian. In training and testing, the system uses the vectorized in light-field vectors of input face images.

They test the eigen light-field method on the CMU (PIE) database and the FERET database, and it out performs both the standard Eigenfaces algorithm and the commercial FaceIt system. And also, they observed that the performance is more significant on the PIE database than on the FERET database using eigen light-fields over the other two algorithms, where the method accomplishes more pose variations.

### **1.2.3 Expressions**

Analysis of facial expressions is accomplished in parallel to face recognition based on neurophysiological studies. Some patients who have difficulties in identifying familiar faces, also seem to recognize expressions due to emotions.

Patients who has “organic brain syndrome” suffer from poor expression analysis but recognize faces quite well. From a machine recognition point of view, dramatic facial expressions may affect face recognition performance if only one photograph is available.

Donato recommends [14] that Gabor filters based and Independent Component Analysis(ICA) methods that outperforms Local Feature Analysis(LCA), LDA and Local PCA while investigating numerous methods for identifying twelve facial expressions.

Tian et al. reported [15] recognition rates of 96.4 percent for upper face parts and 96.7 percent for lower face parts using a almost frontal-view face images in their Automatic Face Analysis system. The systems analyzes stable facial features (brows, eyes, mouth) and active facial features (facial groove changings)

Kernel Canonical Correlation Analysis(KCCA) is proposed in [16] for facial expression recognition. Gabor filters are used transform manually annotated facial landmarks into a labeled graph. Then six-dimensional semantic vector constructed describing basic expressions for each training image. In KCCA, correlation between the semantic vector and labeled graph vector trains the facial expression machine learning. Results are better than conventional methods as LDA and GDA that is guaranteed according to their tests.

#### **1.2.4 Occlusion and Facial Hair**

Another downside of the face recognition system is the partially occluded objects, like glasses and facial hair, which results not a success recognition. At present, none of 3D face recognition methods results better than 2D special representation although the 3D segmentation process applied. The facial hair detection is more simply on 2D images but the segmentation process is more powerful when 3D facial surface extracion is done. So, an elegant combination between the texture image and facial surface is essential. The most of 2D + 3D approaches use 2D and 3D data separately and finally merge rates be related the same information taken from two difference sources. Therefore, an appropriate data sources is necessary, selecting the correct representation for facial hair and segmentation.

To cope with partially occluded objects using local approaches is a one method. In the main, these methods splits the face into different parts and then use a voting schema to find the best match. In (Martinez, 2002) proposed a method as dividing each face image into  $k$  different local parts to deal with how good a local match and not to misclassify as a result of voting test image. Using a Gaussian distribution or a mixture of Gaussians reduces Localization error with modeling each of these  $k$  local parts. For every local subspace, the probability of a given match can be directly associated with the sum of all  $k$  Mahalanobis distances by the mean feature vector and the covariance matrix. The difference from previous local PCA methods is the using a probabilistic approach rather than a voting space. The quantity of the occlusion that can be handled and the minimum number of local areas needed to successfully identify a partially occluded face is investigated by the author. Martinez verified experimentally that the 1/6 occlusion of the face does not decrease accuracy, while even for those cases where 1/3 of the face is occluded, the identification results are very close to those obtained without occlusions. When the eye area is occluded rather than the mouth area results worse, also shown in his work.

Kurita et al. (2003) proposed a method using means of an auto-associative neural network, reconstructs the occluded part of the face and detects the occluded regions in the input image in addition to the probabilistic approach proposed by Martinez that is only able to identify a partially occluded face. At first, the network is trained on the non-occluded images in normal conditions. While testing the replacing occluded regions with the recalled pixels original face can be reconstructed. The trained network has been tested using three types of test data: pixel-wise, rectangular, and sunglass. Even if 20–30% of the face images is occluded, the classification performance is not decreased is claimed by the the authors. On the other hand, the system need retraining in case of new individuals and the small number of avail training samples.

Moreover, Sahbi and Boujemaa (2002) proposed a method trying to handle both occlusions and illumination changes. For the challenging conditions, their work builds a complete scheme for face recognition based on salient feature extraction. Each feature in the query image align is aligned with its corresponding feature in the gallery set, if possible.

With this dynamic space warping alignment, these features are used in the matching process that get the better of occlusion effects and facial expressions. A region subdivided binary image is constructed describing shape variation between different faces after the features extracted. With this process, each feature corresponding matched features in each nominee face of the gallery set is modeled as the statistical deviation. And also, for each extracted and matched feature from the face model describes a matching class. This matching class expresses the possible deviation of this feature (modelled using a Gaussian distribution) with respect to the gallery images. The matching process succeeds with the tests that have been performed using the Olivetti and ARF public databases, noting that for little occlusion and rotation. So the accuracy of recognition is assured respect to small occlusions and rotations.

### **1.2.5 Aging**

Many of the measured techniques fall off in performances, as the time slip between the training and testing representations is significant. This clarifies that none of the described methods consider for problems caused by the age variations. Some approaches conquer this problem orderly by upgrading the gallery or retraining the system. On the other hand, this not very appropriate solution only applies to those systems yielding services which do the authentication, task regularly. It is unusable in other circumstances, such as low enforcement. Otherwise the age of the subject could be simulated trying to make the system stronger with regard to this kind of variation.

Numerous techniques about the age simulation are presented in literature: Coordinate Transformations, Facial Composites, Exaggeration of 3D Distinctive Characteristics, but none of these methods has been considered in the face recognition framework. In a latest work Lanitis and Taylor (2000) and Lanitis et al. (2002) offered a new method for age functions. Every image in the face database is illustrated by a group of parameters  $b$ , and for each subject the best age function is drawn conditional on his/her  $b$ . Different subject-based age functions allow considering for external factors which supply towards the age variations and this is the biggest improvement of this approach.

The authors analyzed this approach on a database including 12 people with 80 images in the gallery and 85 in the research. They reported an enhancement of about 4–8% and 12–15% by swapping the research and gallery set. In both conditions, the tests the signify age of the subjects has been pretended, before doing the recognition task. The number of the subject in the database is very small highlighting the lack of a standard FERET-like database, which analytically formulates the age variations. However, this is an attractive and still unknown portion in low enforcement applications, to improve the strength of the face recognition systems regarding modifications in age, mainly for the forecast of facial appearance of wanted/missing persons.

As a conclusion, lots of researcher try to deal with these face recognition pitfalls as proposing new methods. Table 1.1 shows the results of various tests have been applied by the authors using their methods againts on different databases.

**Table 1.1:** “The main information about the experimental results of most of the discussed methods” [5]

Authors	Method	Databases	Image	Max Gl	Time	rate	Expr	Illu	Pose	Occl	Age
	Name		size	–Max Pl	lapse	(%)					
Martinez and Kak	PCA	AR-Faces	85x60	100–250	No	70		No	No	No	No
Martinez and Kak	LDA	AR-Faces	85x60	100–250	No	88		No	No	No	No
Belhumeur et al.	Fisherfaces	YALE		144–16	No	99,6	Yes	Yes	No	No	No
Yu and Yang	Direct LDA	ORL	112x92	200–200	No	90,8	Yes	Yes	Yes	No	No
Lu et al.	DF-LDA	ORL	112x92	200–200	Yes	96		Yes	No	No	No
		UMIST	112x92	160–415	No	98		No	No	No	No
Cevikalp et al.	DCV	Yale	126x152	15–150	No	97,33		Yes	No	No	No
		AR-Faces	229x299	350–350	Yes	99,35					
Bartlett et al.	ICA	FERET	60x50	425–421	Yes	89	Yes	No	No	No	No
Lin et al.	PDBNN	SCR	80x20	320–1280	No	100	Yes	Yes	Yes	No	No
		FERET		200–200	No	99	Yes	Yes	No	No	No
		ORL			No	96		Yes	Yes	No	No
Joo Er et al.	RBF	? ORL	160x120	300–300		98,1	Yes		Yes	No	No
Meng et al.		PropertyDB				100					
Perronnin											
and Dugelay	HMM	FERET	128x128	500–500	No	97	Yes	No	No	No	No
Lades et al.	DLA	PropertyDB	128x128	88–88	No	90,3	Yes		Yes	No	No
Liu	Gabor EFM	FERET	128x128	200–100	No	99	Yes	No	No	No	No
		ORL	128x128	200–200	No	100	Yes	No	Yes	No	No
Wiskott et al.	EGM	FERET	256x384	250–250	No	80	Yes		Yes	No	No
		PropertyDB		108–//		90	Yes		Yes	No	No
Garcia et al.	WPA	MIT	480x640	155–155		80,5	Yes	Yes		No	no
		FERET	256x384	200–400		89					
Kouzani et al.	IFS	PropertyDB	64x64	100–100		100		No	No	No	No
Tan and Tan	IFS	ORL	92x112	200–//	No	95				No	No
Ebrahimpour et al.	IFS	MIT	480x640	90–90		90			Yes	No	No

Authors	Method	Databases	Image	MaxlGl	Time	rate	Expr	Illu	Pose	Occl	Age
Chen et al.	Th-Infrared	PropertyDB		166–166	No	98	Yes	Yes	No	No	No
Socolinsky and Selinger	Thermal	PropertyDB	99x132	770–2310	Yes	93	Yes	Yes	No	No	No
Buddharaju et al.	Th-Spectrum	Equinox		225–2500		86,8	Yes		Yes	No	No
Pan et al.	Hyperspectral	PropertyDB		200–1200	Yes	92	No	Yes	No	No	No

### 1.3 Images As Vectors

An image is considered as a high-dimensional vector and a point in a high-dimensional vector space. For example, a  $p \times q$  2D image can be represented to a vector  $x \in \mathbb{R}^{pq}$ , by concatenating each row of the image as lexicographic ordering of the pixel elements. Even though this high-dimensional mapping, data lies in a lower-dimensional manifold because of the natural constraints of the physical world and the imaging process. The main objective of the subspace analysis is to identify, represent, and parameterize this manifold with some optimality criteria.

Let  $X = (x_1, x_2, \dots, x_i, \dots, x_N)$  represent the  $n \times N$  data matrix, where each  $x_i$  is a  $n$  dimensional face vector concatenated from a  $p \times q$  face image, where  $n = p \times q$ . Here  $n$  is the total number of pixels in the face image and  $N$  is the number of different face images in the training set. The mean vector of the training images  $m = \sum_{i=1}^N x_i$  is subtracted from each image vector.



## 2. FACE RECOGNITION USING GABOR REPRESENTATION

### 2.1 Preface

Among the others, understanding how people process and recognize each other's face is the most challenging task for image analysis and object recognition. And the development of corresponding computational models for automated face recognition. As well as classification issues, representation should be cared for a good face recognition methodology, i.e. minimum manual annotated representation. The Gabor filters reveal intended characteristics of spatial locality and orientation selectivity whose kernels are similar to the 2-D receptive field profiles of the mammalian cortical simple cells. No need for manual annotations where the Gabor filter representation implements recognition without correspondence. It computes the local structure corresponding to spatial frequency (scale), spatial localization, and orientation selectivity. As a result, the Gabor filter representation of face images should be robust to variations due to illumination and facial expression changes.

### 2.2 2D Gabor Filters

Gabor Filters optimally localized in the space and frequency domains and are a set of filters (kernels)  $\psi_{\vec{k}}$ , where  $\vec{k} = k_{\mu, \nu}$ . Each kernel is a product of a Gaussian envelope function and a complex plane wave. Gabor kernels, in image coordinates  $z = (x, y)$ , are defined as follows,

$$\psi_{\mu, \nu}(z) = \frac{\|k_{\mu, \nu}\|^2}{\sigma^2} e^{(-\|k_{\mu, \nu}\|^2 \cdot \|z\|^2 / 2 \cdot \sigma^2)} \left[ e^{i \cdot z \cdot k_{\mu, \nu}} - e^{-\sigma^2 / 2} \right]. \quad (2.1)$$

where  $\mu$  indicates the orientation and  $\nu$  indicates the scale of the Gabor kernels,  $\| \cdot \|$  denotes the norm operator, and the wave vector is defined as follows:

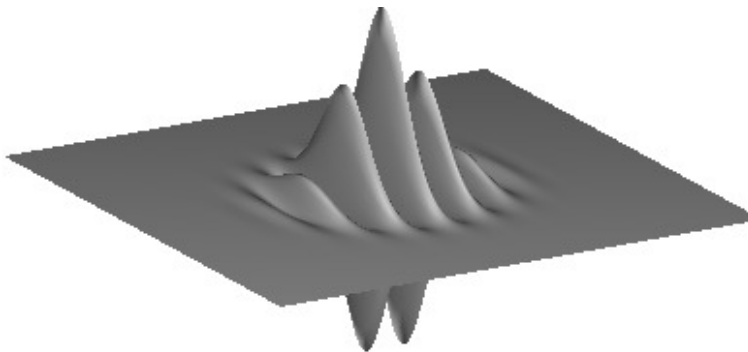
$$k_{\mu,v} = k_v e^{i\phi_\mu}, \quad (2.2)$$

where  $k_v = k_{\max} / f^v$  and  $\phi_\mu = \pi\mu/8$ .  $k_{\max}$  is defined as maximum frequency and  $f$  is the spacing factor between kernels in the frequency domain.

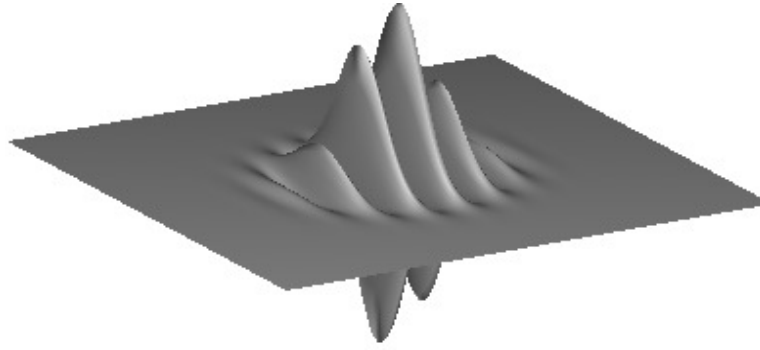
The first exponential term in the square brackets in equation (2.1) indicates the oscillatory part while the second exponential term compensates for the DC value of the kernel, to make the filter independent from the absolute intensity of the image. The kernel, exhibiting complex response, combines a real part(cosine part) and an imaginary part(sine part). The filters are parameterized by  $k_{\mu,v}$ , which controls the width of the Gaussian window and scale and orientation of the oscillatory part. The  $\sigma$  parameter determines the ratio of window width to scale, in other words, the number of the oscillations under the envelope equation (2.2).

Lades et al. investigated  $\sigma = 2\pi$ ,  $f = \sqrt{2}$  and  $k_{\max} = \pi/2$  yielding with optimal results along with 5 scales,  $v \in \{0, \dots, 4\}$ , and 8 orientations,  $\mu \in \{0, \dots, 7\}$ . In [17], Shen et al. also discussed tuning the Gabor kernel parameters and after two experiments they showed that 5 scales and 8 orientations yielded with optimal recognition performance.

Figure 2.1 illustrates the 3-D shape of real and imaginary part of the kernel. And Figure 2.2 illustrates the 2-D representations of real part of gabor filters with 5 scales and 8 orientations and their magnitudes, along with parameters  $\sigma = 2\pi$ ,  $f = \sqrt{2}$  and  $k_{\max} = \pi/2$ .



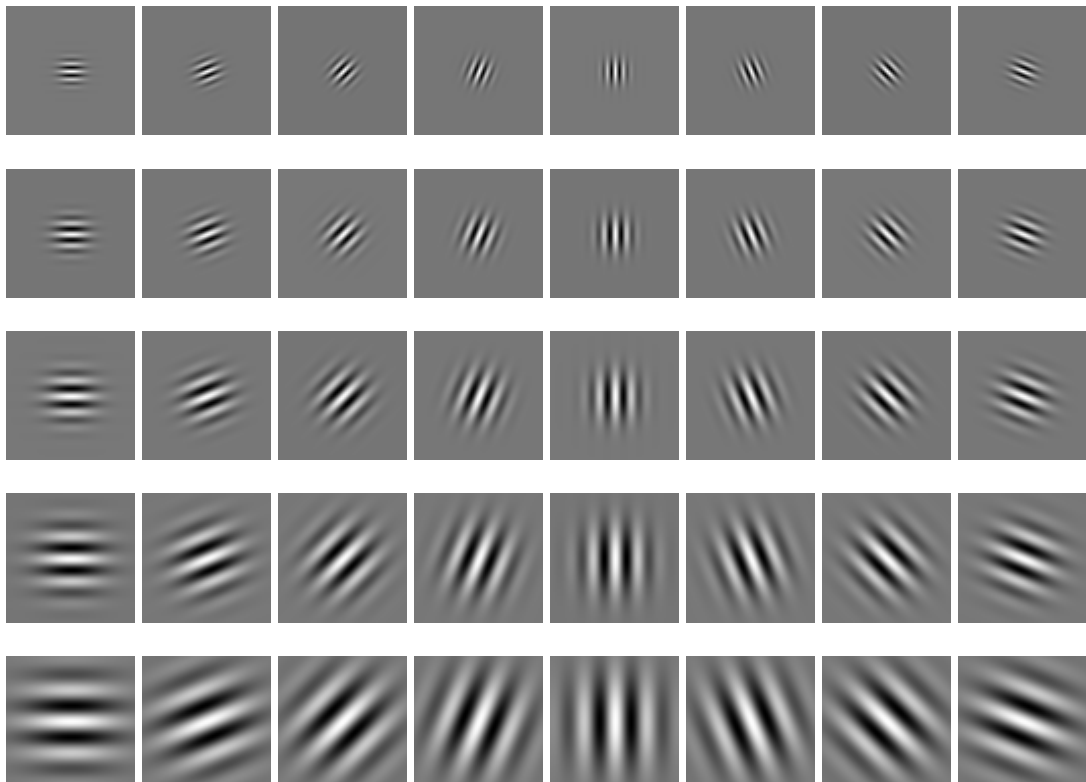
(a)



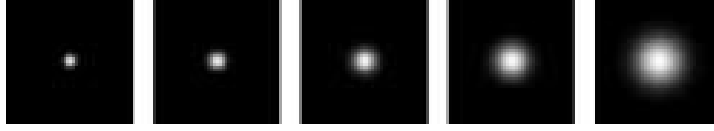
(b)

**Figure 2.1:** 3-D visualization of a Gabor kernel.

In Figure 2.1 (a) the cosine part(real part) of the kernel with  $\mu = 0.72, \nu = 45^\circ$  and in (b) the sine part(imaginary part) is shown . The kernel has a size of 128 units in each of the first two dimensions.



(a)



(b)

**Figure 2.2:** 2-D Gabor kernels of 5 scales and 8 orientations.

In Figure 2.2 (a) The real part of the Gabor kernels at five different scales and eight orientations with the parameters;  $\sigma = 2\pi$  ,  $f = \sqrt{2}$  ,  $k_{\max} = \pi/2$  and in (b) the magnitude of the Gabor kernels at five different scales is illustrated.

### 2.3 2D Gabor Filtered Representation of Images

Gabor filters have been widely used in many image processing and computer vision applications such as texture segmentation, face detection, head pose estimation, vehicle detection, character recognition, fingerprint recognition, face identification, tracking and verification, after the ground-breaking work of extending 1-D Gabor filters to 2-D by Daugman [18].

The 2-D Gabor filter representation of an image is the convolution of the image with a family of kernels  $\psi_{\mu,\nu}$ , where  $\mu$  is the orientation and  $\nu$  is the spatial scale of the kernel. The convolution result of an image  $I(x,y)$  with a Gabor kernel  $\psi_{\mu,\nu}$  is defined as follows,

$$G_{\mu,\nu}(z) = I(x, y) * \psi_{\mu,\nu}, \quad (2.3)$$

where  $z = (x,y)$  and  $*$  denotes the convolution operator.  $G_{\mu,\nu}(z)$  is the convolution result corresponding to the Gabor kernel at orientation  $\mu$  and scale  $\nu$ . The convolution results of Gabor kernels with five scales and eight orientations forms the set of Gabor filter representation of an image  $I(z)$ . This set can be defined as,

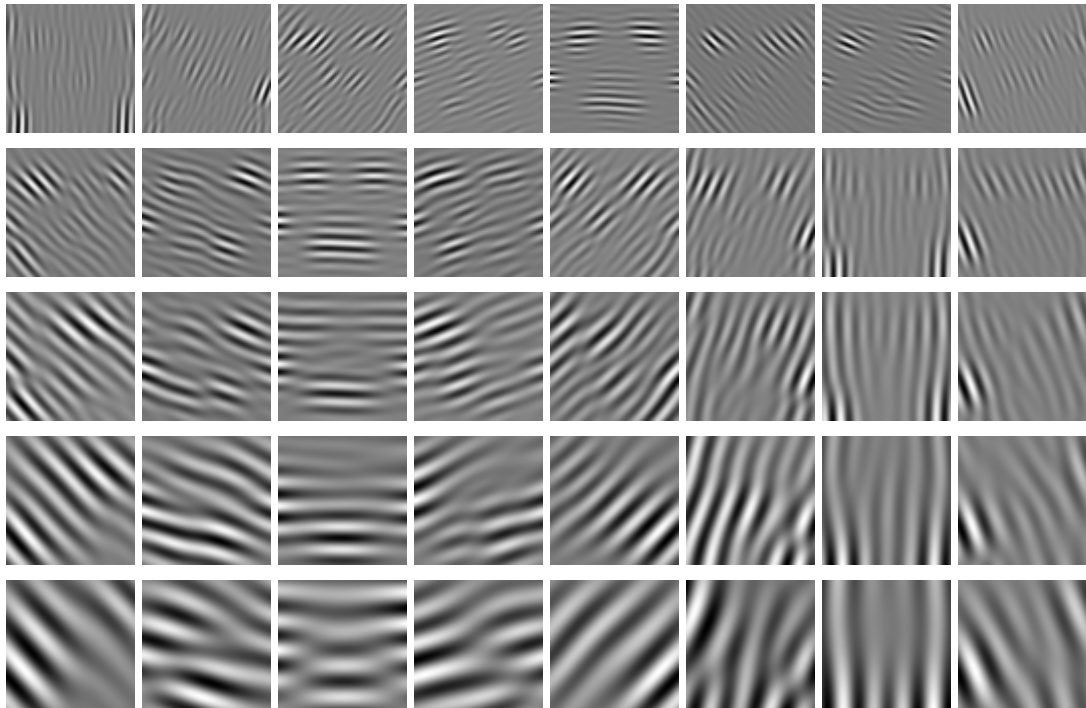
$$S = \{G_{\mu,\nu}(z) : \mu \in \{0, \dots, 7\}, \nu \in \{0, \dots, 4\}\}. \quad (2.4)$$

This feature set is called as Gabor jets in [9, 10]. To perform the time-consuming convolution operation efficiently in Fourier domain, this set can be derived using

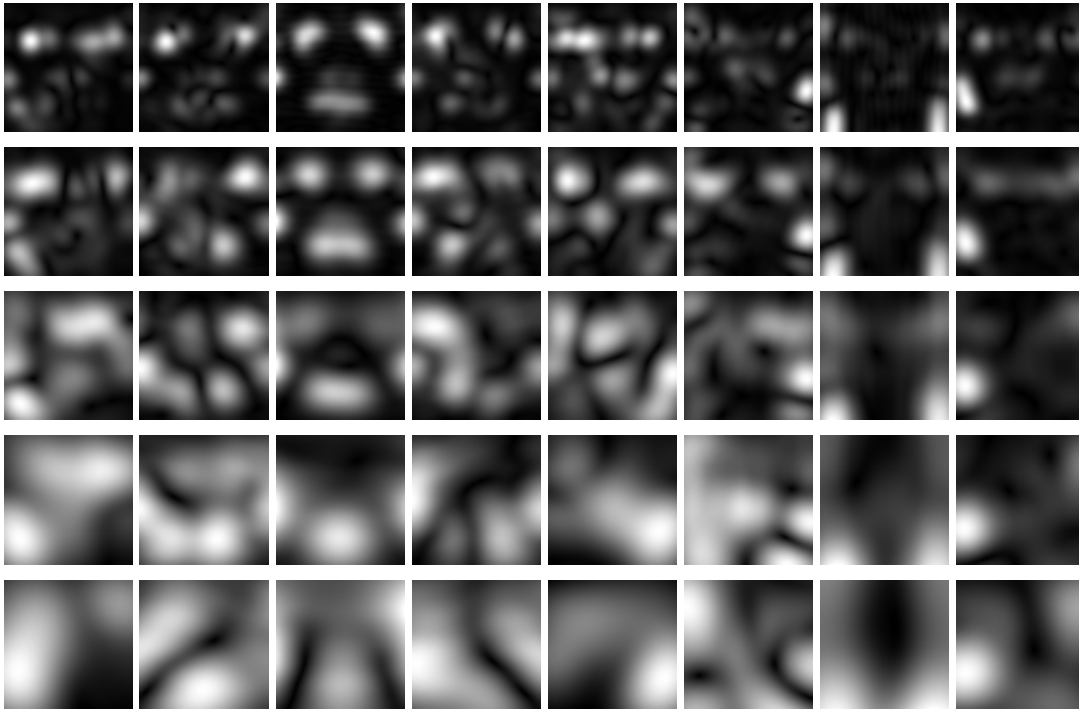
Fast Fourier Transform(FFT) and then taking the inverse Fourier transform back. This operation is defined as,

$$G_{\mu,\nu}(z) = F^{-1}\{F\{I(z)\}F\{\psi_{\mu,\nu}(z)\}\}, \quad (2.5)$$

where  $F^{-1}$  and  $F$  denote the inverse Fourier transform and Fourier transform, respectively. In face recognition, the magnitude of the Gabor jets extracted at some fiducial points [10], or the augmented form of the magnitude of the convolution results from full images are used as feature vectors [19]. Magnitude of the filter response is the square root of the sum of the squares of real and imaginary parts of the filter output. Figure 2.3 shows the Gabor filter representation(the real part and the magnitude) of a 64x64 sample image from the ORL database. And Figure 2.4 shows (in image form rather than in vector form) an example of the augmented Gabor feature vector.

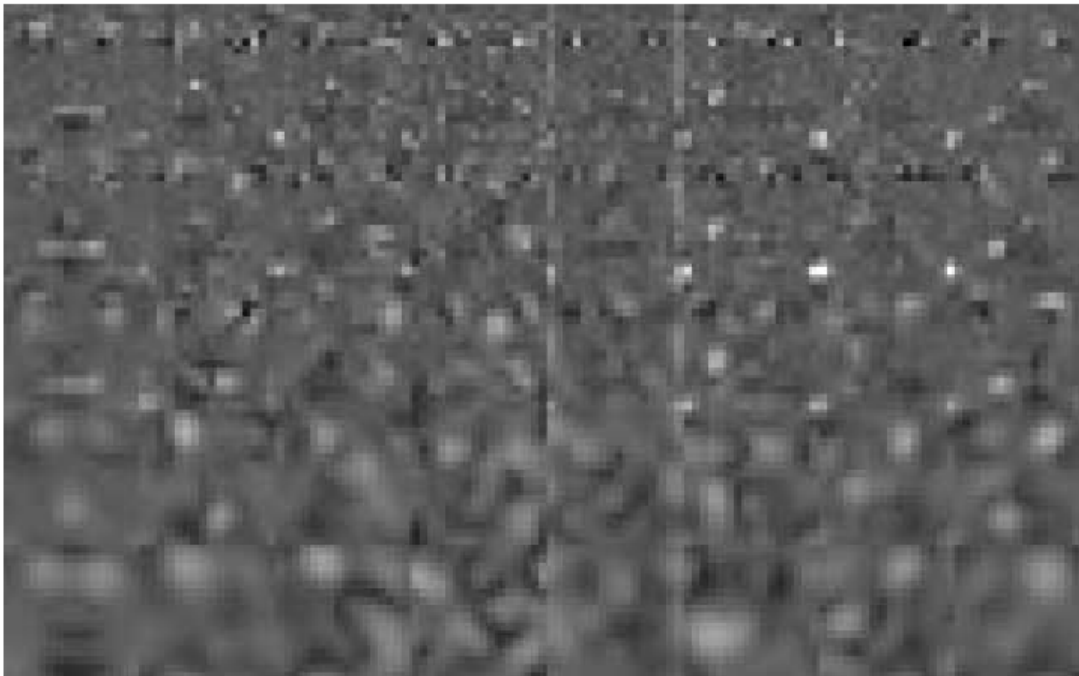


(a)



(b)

**Figure 2.3:** Gabor filter representation (the real part and the magnitude) of a 64x64 sample image from ORL database. (a) The real part of the representation and (b) The magnitude part of the representation.



**Figure 2.4:** An example of the augmented Gabor feature vector (in image form rather than in vector form)

### 3. APPEARANCE BASED FACE RECOGNITION

#### 3.1 Linear Subspace Analysis

PCA [20], ICA [21] and LDA [22][23]; three classical linear appearance-based classifiers, are introduced in the this section. Each classifier has its own basis vector representation extracted from high dimensional face vector space using different statistical analysis. The projection coefficients are used as the feature representation of each face image by projecting the face vector to the basis vectors. Using similarity or distance metrics the matching score between the test face image and the training prototype is calculated between their coefficients vectors. The higher the matching score, the better the match.

All the three analysis can be considered as a linear transformation from the original image vector to a projection feature vector, i.e.

$$Y = W^T X \quad (3.6)$$

where  $Y$  is the  $d \times N$  feature vector matrix,  $d$  is the dimension of the feature vector, and  $W$  is the transformation matrix. Note that  $d \ll n$ .

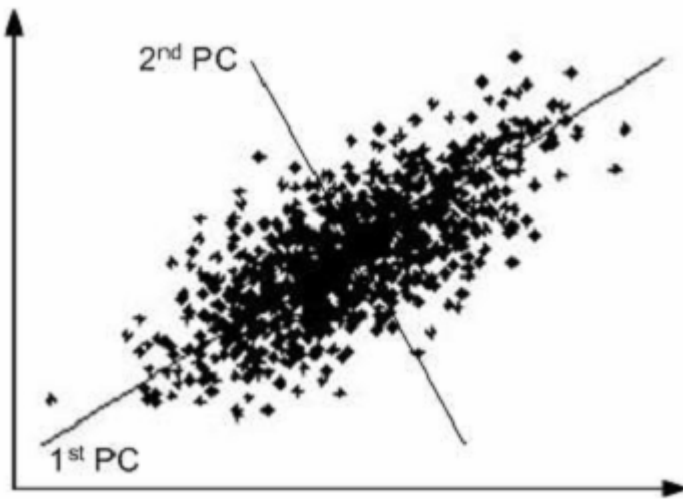
#### 3.2 PCA

To find the vectors which best representation for the distribution of face images within the entire image space the Eigenface method uses the Principal Component Analysis (PCA) for dimensionality reduction [20]. These vectors characterize the subspace of face images and the subspace is called face space. All faces in the training set are projected onto the face space to find a set of eigen values that describes the donation of each vector in the face space. To identify a test image, the test image is projected onto the face space to obtain the corresponding set of eigen values. The test image can be classified by comparing the eigens of the test image with the set of eigens of the faces in the training set.

Karhunen-Loeve transformation is the key procedure of PCA [25]. The image may be thought as a sample of a stochastic process since the image elements are considered to be random variables. The Principal Component Analysis basis vectors are defined as the eigenvectors of the scatter matrix  $S_T$ ,

$$S_T = \sum_{k=1}^N (x_k - m)(x_k - m)^T, \quad (3.7)$$

The eigenvectors related to the  $d$  largest eigenvalues constructs the transformation matrix  $W_{PCA}$ . A 2D example of PCA is shown in Figure. 3.1.

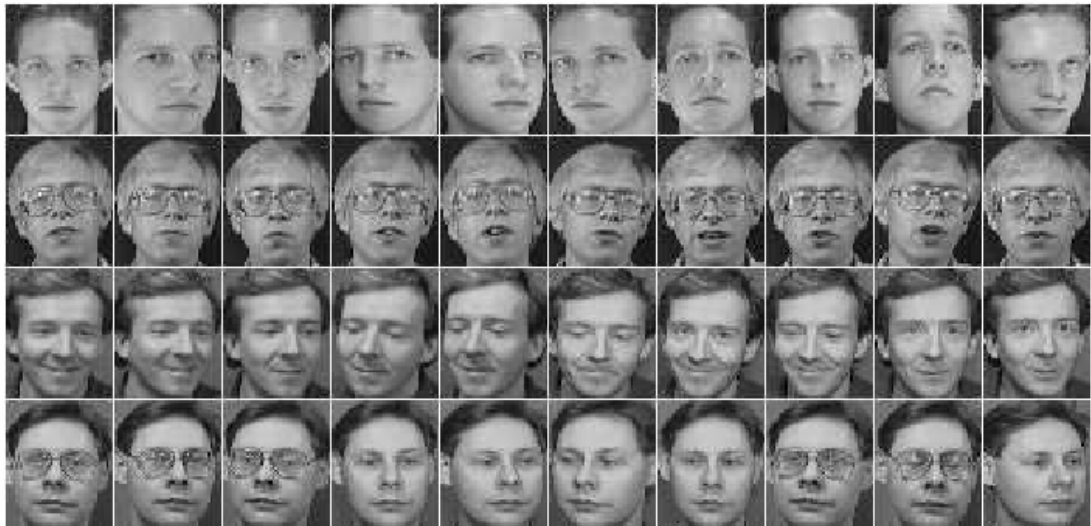


**Figure 3.1:** Principal components (PC) of a 2D set of points

The first principal component provides an optimal linear dimension reduction from 2D to 1D, but also results in the mean square error.

Face samples from ORL face database [Url-2] are shown in Figure 3.2. The corresponding mean face is given in Figure 3.3. The eigenvectors (a.k.a. eigenface) corresponding to the 7 largest eigenvalues are shown in Figure 3.4. The input vector (face) in an  $n$ -dimensional space is reduced to a feature vector in a  $d$ -dimensional subspace after applying the projection. Also the eigenvectors corresponding to the 7 smallest eigenvalues are provided in Figure 3.5. For most applications, these very small eigenvalued eigenvectors to are considered as noise, and not taken into account during identification. Several extensions of PCA are developed, such as modular eigenspaces [26] and probabilistic subspaces [27].





**Figure 3.2:** Face samples from the ORL face database



**Figure 3.3:** The mean face.



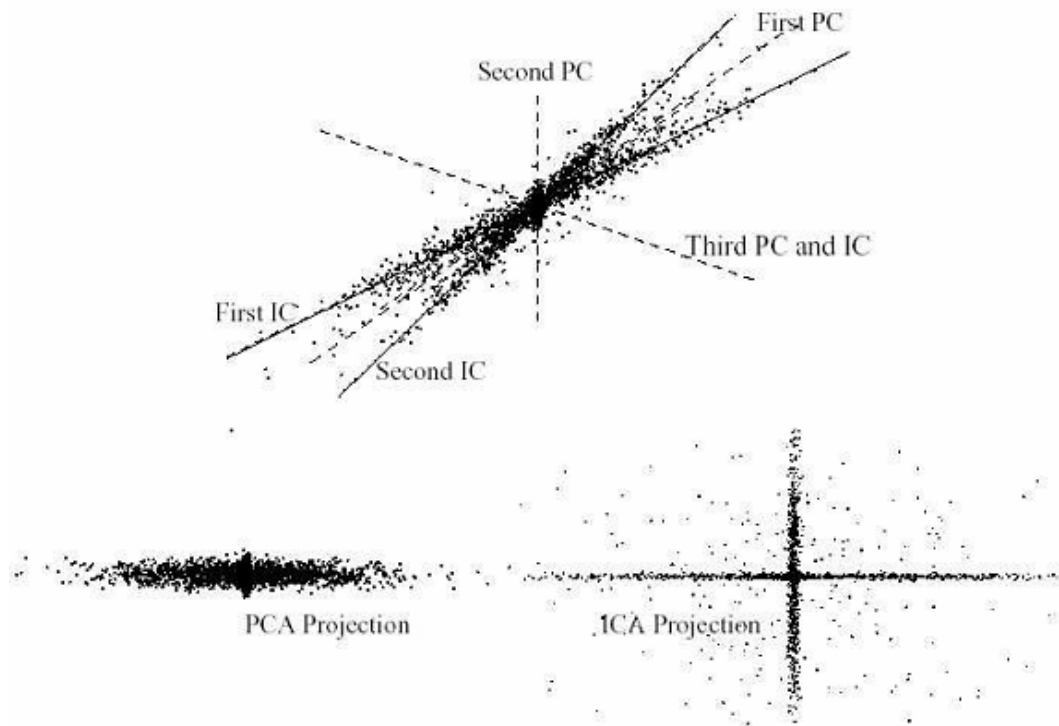
**Figure 3.4:** Eigenvectors corresponding to the 7 largest eigenvalues, shown as  $p \times p$  images, where  $p \times p = n$



**Figure 3.5:** Eigenvectors corresponding to the 7 smallest eigenvalues shown as  $p \times p$  images, where  $p \times p = n$

### 3.3 ICA

Except the components distribution are extracted to be non-Gaussian, Independent Component Analysis (ICA) [28] is similar to PCA. Maximizing non-Gaussianity props up statistical independence. Figure 3.6 presents the different feature extraction properties between PCA and ICA.



**Figure 3.6:** Properties between PCA and ICA

In Figure 3.6, the top graph shows the example 3D data distribution and the corresponding principal component and independent component axes. Each axis is a direction of PCA or ICA. Note the PCA axes are orthogonal while the ICA axes are not. ICA chooses a different subspace than PCA, if only 2 components are allowed. The bottom left one shows distribution of the first PCA axis of the data. The bottom right shows distribution of the first ICA axis of the data [29]. To sum up for this example, ICA leans to extract more intrinsic structure of the original data clusters.

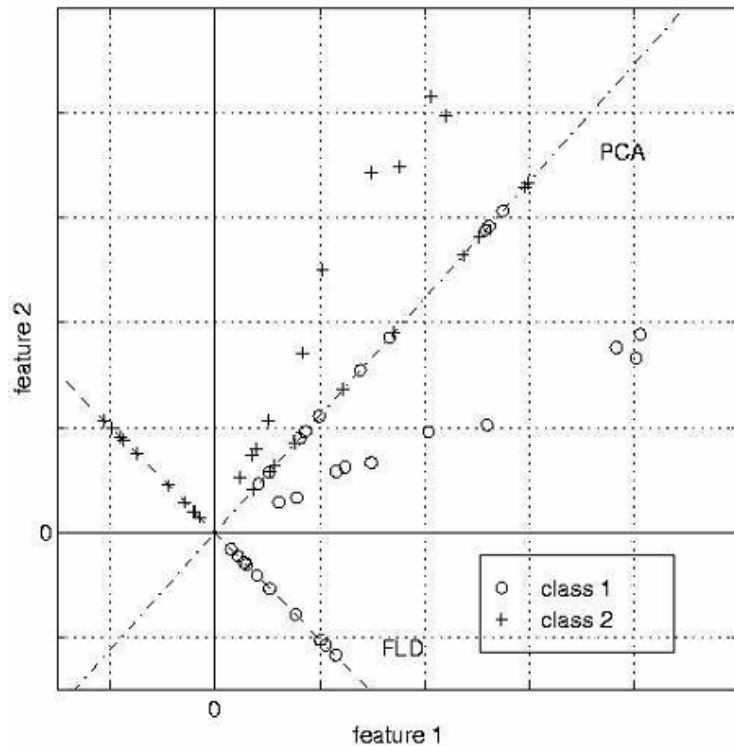
Bartlett et al. [21] built two architectures based on Independent Component Analysis, statistically independent basis images and a factorial code representation. The ICA splits the high-order moments of the input in addition to the second-order moments taken into account in PCA. Both methods results in similar performance. The obtained basis vectors based on fast fixed-point algorithm of the ICA factorial code representation [29] are illustrated in Figure 3.7. There is no special order applied on the ICA basis vectors.



**Figure 3.7:** ICA basis vectors shown as  $p \times p$  images

### 3.4 LDA

Both PCA and ICA construct the face space with no prior information as the face class information. The whole face training data is calculated. In Linear Discriminant Analysis, the aim is to find an efficient way to span the face vector space. Using the class information could be helpful to the identification processes.



**Figure 3.8:** A comparison of PCA and FDA

Figure 3.8 demonstrates a comparison of principal component analysis (PCA) and Fisher's linear discriminant (FDA) for a two class problem and FDA is better than PCA in the sense of discriminating the two classes where data for each class lies near a linear subspace [22].

The Fisherface algorithm [22] is derived from the Fisher Linear Discriminant (FDA), which uses class specific information. Images in the learning set are divided into the corresponding classes via defining different classes with different statistics. Then, techniques similar to those used in Eigenface algorithm are applied. The Fisherface algorithm results in a higher accuracy rate when compared with Eigenface algorithm.

The Linear Discriminant Analysis finds a transform  $W_{OPT}$ , such that,

$$W_{opt} = \arg \max_w \frac{|W^T S_b W|}{|W^T S_w W|} = [w_1 | w_2 | \dots | w_d]. \quad (3.8)$$

where  $S_b$  is the between-class scatter matrix and  $S_w$  is the within-class scatter matrix, defined as,

$$S_b = \sum_{i=1}^c N_i (m_i - m)(m_i - m)^T, \quad (3.9)$$

$$S_w = \sum_{i=1}^c \sum_{k=1}^{N_i} (x_k - m_i)(x_k - m_i)^T. \quad (3.10)$$

In the above expression,  $N_i$  is the number of training samples in class  $i$ ,  $c$  is the number of distinct classes,  $m_i$  is the mean vector of samples belonging to class  $i$  and  $x_i$  represents the set of samples belonging to class  $i$ . The FDA basis vectors using a subset of ORL face database [Url-2] are shown in Figure 3.9.



**Figure 3.9:** First seven FDA basis vectors shown as  $p \times p$  images

Although the FDA is one of the best and most common used method, it has also restrictions such as when coping with a high dimensional data, FDA usually faces with the small sample size problem. Furthermore it can be unsuccessful if the class distributions are more common where as it promises to find the optimal directions if each class has a Gaussian distribution with a common covariance matrix.

### 3.5 NNDA

The Nearest Neighbor Discriminant Analysis is the abbreviation of NNDA and it is a nonparametric linear feature extraction method that is based on the nearest neighbor classification (NN). NNDA calculates the optimal discriminant directions without assuming the class distributions be a part of any parametric distributions. When dealing with nonparametric classification; nearest neighbor classification is an effective technique, which is widely used in the pattern recognition area. Even more, it does not rely on nonsingularity of the within class scatter matrix and it is closed to Bayes classifier.

LDA maximizes the between scatter matrix at same time minimizing the within class scatter matrix.

$$S_b = \sum_{i=1}^c p_i (m_i - m)(m_i - m)^T \quad (3.11)$$

$$S_w = \sum_{i=1}^c p_i S_i \quad (3.12)$$

$S_b$  is the between class scatter matrix and  $S_w$  is the within class scatter matrix where  $c$  is number of classes,  $m_i$  is mean vector and  $p_i$  is priori probability of class  $i$ .  $S_i$  is the covariance matrix and  $m$  is the global mean vector as given below,

$$m = \sum_{i=1}^c p_i m_i \quad (3.13)$$

LDA maximizes the ratio of determinant of  $S_b$  to  $S_w$  while trying to determine a set of projection vectors  $W \in R^{D \times d}$  defined as,

$$W = \arg \max_w \frac{|W^T S_b W|}{|W^T S_w W|} \quad (3.14)$$

D is the data dimensionality before the projection and d is the data dimensionality after the projection.

The transformation matrix W consists of d eigenvectors matched to first d largest eigenvalues of  $S_w^{-1} S_b$  from equation (3.9).

However,  $S_w^{-1}$  does not exist where  $S_w$  is singular when sample size is small. And also, LDA could not determine the discriminant direction if class densities are multimodal and uses the same mean. That's why a lot of works have been introduced to solve these problems.

### **Approaches on solving Singularity of $S_w$ ;**

Singularity of the  $S_w$  has been pointed out by lots of researchers in recent years and they also try to handle the computation problem of the LDA.

Before applying LDA, PCA is applied to keep away from the singularity of  $S_w$ . Principal Component Analysis (PCA) projects high dimensional data onto a low dimensional feature space. PCA is very effective for eliminating the noise in the data by maximizing the covariance of all data without label knowledge. A set of mutually orthogonal basis functions where their coefficients are pair wise uncorrelated is determined by the PCA that functions find directions of the data for the maximum variance. After PCA, LDA is applied on low dimensional PCA space where  $S_w$  is non singular. But some discriminative information is lost that's why this technique is not optimal.

Instead of dividing the equation (3.9) by  $S_w$ , Liu et al [31] suggested using the total scatter matrix as  $S_t = S_b + S_w$  divider. But it is exactly same as the Fisher's method. Nevertheless, for any transformation matrix W in the null space of  $S_w$  the suggested method gets maximum value, 1, when  $S_w$  is singular. Therefore, the class separability cannot be maximized at all time for the transformation matrix W where  $|W^T S_w W|$  maximized.

And also it takes time to estimate an inverse matrix. Then, there is a method suggested by Chen et al [32] which contains most discriminative information. According to their approach, using eigenvectors of  $S_w$  null space is spanned with zero eigenvalues.

NLDA is introduced which is an LDA variant in the null space of  $S_w$ . In NLDA, when  $S_w$  is zero, the projection vectors are selected that maximize the  $S_b$ . But, discriminative information out of  $S_w$ 's null space is not considered. However, almost no discriminative information has within null space of  $S_w$  that resulted in over fitting. Therefore, maximizing the between-class scatter in the null space of  $S_w$  instead of original data space is not optimal. And also, when  $N - c$  is near to the dimension  $D$ , the NLDA performance decreases meaningfully where  $N$  is the number of samples and  $c$  is the number of classes. Thus, significant information is wasted when the dimensionality of the null space is small. For this problem, there is an algorithm is introduced by Yu et al [33] where the null space of  $S_b$  is removed first. There is no discriminative information in the null space according to their method. The optimal discriminant vectors are not bound to be on the subspace which class centers divide this subspace. Thus, this method is not correct.

### **Approaches for problems of $S_b$ ;**

Class separability is not clearly defined by  $S_b$  when the class densities are multimodal. Moreover, LDA could not determine the discriminant direction when some classes have a shared mean where class scatter means not exists.

Not more than  $c - 1$  feature can be extracted where  $S_b$ 's rank is  $c - 1$ . It is not much optimal for  $c - 1$  feature from the view of Bayes not until posteriori probability functions are chosen, despite the fact that it is optimal according to Fisher's method.

Actually, calculating the class distributions close to the decision boundary is the main purpose of the classification.

A nonparametric discriminant analysis is introduced by Fukunaga and Mantock [34] that tries to handle the problems of LDA. The between scatter matrix  $S_b$  is assumed that has a nonparametric behavior in nonparametric analysis. In general  $S_b$  is full rank where makes feature dimensionality that is extracted less tight. Moreover, extracted features that are related to the classification are conserved because of  $S_b$ 's nonparametric structure.

Bressan et al [35] discovered the connection between nearest neighbor (NN) classifiers and nonparametric discriminant analysis (NDA). They extend the two-class NDA to multi class NDA by giving a small adjustment.

However, all the nonparametric techniques try to solve the problem of  $S_b$ , singularity of  $S_b$  still exists where  $S_b$ 's rank must be at most  $N - c$ .

### 3.5.1 The Criterion

Let  $w_i$  ( $i = 1, 2, 3, \dots, c$ ) are the classes for a multi-class problem.

For each sample, assume that there is a nearest neighbor in different class. It can be named as extra-class nearest neighbor and defined below,

$$x_n^E = \arg \min_z \|z - x_n\|, \forall z \notin w_i \quad (3.15)$$

And also for each sample, assume that there is a nearest neighbor in the same class of the sample. It can be named as intra-class nearest neighbor and defined below,

$$x_n^I = \arg \min_z \|z - x_n\|, \forall z \in w_i, z \neq x_n \quad (3.16)$$

And, extra-class and intra-class differences are,

$$\Delta_n^E = x_n - x_n^E \quad (3.17)$$

$$\Delta_n^I = x_n - x_n^I \quad (3.18)$$

Then between-class and within-class scatter matrix are,

$$\hat{S}_b = \sum_{n=1}^N w_n (\Delta_n^E)(\Delta_n^E)^T \quad (3.19)$$

$$\hat{S}_w = \sum_{n=1}^N w_n (\Delta_n^I)(\Delta_n^I)^T \quad (3.20)$$

Extra-class, intra-class differences and between-class and within-class scatter matrices are nonparametric.

Here,  $w_n$  is the sample weight defined as,



$$w_n = \frac{\|\Delta_n^I\|^\alpha}{\|\Delta_n^I\|^\alpha + \|\Delta_n^E\|^\alpha} \quad (3.21)$$

where  $\|\Delta_n^E\|$  is the distance between the sample  $x_n$  and its nearest neighbor in the other class,  $\|\Delta_n^I\|$  is the distance between the sample  $x_n$  and its nearest neighbor in the same class and  $\alpha$  is a control parameter between zero and plus infinity.

The sample weight  $w_n$  is used for giving more importance to the samples near to the other class than the samples in the class center. In other words, the samples in the class center are more distant than the samples in the boundary of the class to their nearest neighbor in the other classes. Therefore, multiplying extra-class and intra-class differences with  $w_n$  makes slight effect on the scatter matrix for the samples in the class center. In general, the sample weight,  $w_n$ , is 0.5 in the boundary of the class and decreases to zero as closing to the class centers. And, this changing can be controlled by the parameter  $\alpha$ .

The accuracy of the nearest neighbor classification can be calculated as

$$\theta_n = \|\Delta_n^E\|^2 - \|\Delta_n^I\|^2 \quad (3.22)$$

$x_n$  is classified correctly if  $\theta_n$  is greater than zero else  $x_n$  is classified as incorrectly. Here, accuracy means that how sure the  $x_n$  is classified. For example, if the difference of extra-class and intra-class norms is notably large the classifier is more confident for the  $x_n$ .

After the features extracted by linear projection matrix  $W_{D \times d}$ , the projected samples are  $x_n^{new} = W^T x_n$  and  $\delta_n^E = W^T \Delta^E$  is the nonparametric extra-class difference,  $\delta_n^I = W^T \Delta^I$  is the non parametric intra-class difference. To be more confident on the classification the accuracy must be as large as possible in the projected space by determining optimal  $W$  as below.

$$\hat{W} = \arg \max_W \sum_{n=1}^N w_n (\|\delta_n^E\|^2 - \|\delta_n^I\|^2) \quad (3.23)$$

In other words, the optimization problem is while minimizing the distances of the samples in the same class determine the linear transform that maximize the distances between classes.

If the  $\delta_n^E = W^T \Delta^E$  and  $\delta_n^I = W^T \Delta^I$  is put into equation (3.18) the formula becomes,

$$\begin{aligned}
& \sum_{n=1}^N w_n (\|\delta_n^E\|^2 - \|\delta_n^I\|^2) \\
&= \sum_{n=1}^N w_n (W^T \Delta^E)^T (W^T \Delta^E) - \sum_{n=1}^N w_n (W^T \Delta^I)^T (W^T \Delta^I) \\
&= \text{tr} \left( \sum_{n=1}^N w_n (W^T \Delta^E) (W^T \Delta^E)^T \right) - \text{tr} \left( \sum_{n=1}^N w_n (W^T \Delta^I) (W^T \Delta^I)^T \right) \tag{3.24} \\
&= \tau \rho \left( \Omega^T \left( \sum_{v=1}^N \omega_v (\Delta^E) (\Delta^E)^T \right) \Omega \right) - \tau \rho \left( \Omega^T \left( \sum_{v=1}^N \omega_v (\Delta^I) (\Delta^I)^T \right) \Omega \right) \\
&= \text{tr} (W^T \hat{S}_b W) - \text{tr} (W^T \hat{S}_w W) \\
&= \text{tr} (W^T (\hat{S}_b - \hat{S}_w) W)
\end{aligned}$$

where  $\hat{S}_b$  is between-class scatter matrix,  $\hat{S}_w$  is within-class scatter matrix and  $\text{tr}()$  operator is the trace of a matrix. The scatter matrices  $\hat{S}_b$  and  $\hat{S}_w$  are defined in the equation (3.14) and (3.15).

Thus, equation (3.18) becomes,

$$\hat{W} = \arg \max_W \left( \text{tr} (W^T (\hat{S}_b - \hat{S}_w) W) \right) \text{ subject to } W^T W = I \tag{3.25}$$

The equation (3.20) is called as nearest neighbor discriminant analysis (NNDA) criterion.

Finally, the transformation matrix  $\hat{W}$  consists of  $d$  eigenvectors matched to first  $d$  largest eigenvalues of  $\hat{S}_b - \hat{S}_w$  from equation (3.20).

### 3.5.2 Stepwise Dimensionality Reduction

In the nearest neighbor discriminant analysis nonparametric extra-class and intra-class ( $\Delta^E$  and  $\Delta^I$ ) differences are calculated in the original high dimensional space then projected onto low dimensional space ( $\delta_n^E = W^T \Delta^E$  and  $\delta_n^I = W^T \Delta^I$ ). The method does not assure  $\delta^E$  and  $\delta^I$  agree exactly with the non-parametric extra-class and intra-class differences in projection subspace under orthonormal transformation case exception. Calculating the projection matrix  $\hat{W}$  by stepwise dimension reduction is a solution for this problem. The nonparametric extra-class and intra-class differences are recalculated in its current dimensionality for each step. Hence, nonparametric extra-class and intra-class differences consistency are preserved in the execution of dimensionality reduction.

Given D dimensional samples  $\{x_1 \dots x_N\}$ , the aim is to find d-dimensional discriminant subspace. These are steps of the stepwise dimensionality reduction;

- Reduce the dimensionality of samples to  $d_t$  in step t, and  $d_t$  meet the conditions:  
 $d_{t-1} > d_t > d_{t+1}$ ,  $d_0 = D$  and  $d_T = d$ .
- For  $t = 1 \dots N$ 
  1. Calculate the nonparametric between-class  $\hat{S}_b^t$  and within within-class scatter matrix  $\hat{S}_w^t$  in the current  $d_{t-1}$  dimensional space;
  2. Calculate the projection matrix  $\hat{W}_t$ ;  $\hat{W}_t$  is  $d_{t-1} \times d_t$  matrix.
  3. Project the samples by the projection matrix  $\hat{W}_t$ ,  $x = \hat{W}_t^T x$ .
- The final transformation matrix  $\hat{W} = \prod_{t=1}^T \hat{W}_t$

### 3.5.3 NNDA to k-NNDA

Given a training sample  $x \in w_i$  and its k-nearest neighbors, k-NN classify the x correctly if the number of neighbors belong to  $w_i$  bigger than other classes' number of the neighbors.

According to the method, k-nn classifier is restricted by the criterion which is  $x$  will be classified correctly if no less than  $[k/2] + 1$  neighbors belong to  $w_i$ . 1-NNDA is extended to k-NNDA in this criterion.

$x_{[k/2]}^E$  is defined as extra-class  $[k/2]^{\text{th}}$  nearest neighbor and  $x_{[k/2]+1}^I$  is defined as intra-class  $[k/2]+1^{\text{th}}$  nearest neighbor for the sample  $x$ . Now, k-NN will classify correctly if the distance between  $x$  and  $x_{[k/2]+1}^I$  is less than the distance between  $x$  and  $x_{[k/2]}^E$  where the majority ( $[k/2] + 1$ ) of its  $k$  nearest neighbors are belong to  $w_i$ .

Then the nonparametric extra-class and intra-class differences in equation (3.12) and (3.13) becomes;

$$\Delta^E = x - x_{[k/2]}^E \quad (3.26)$$

$$\Delta^I = x - x_{[k/2]+1}^I \quad (3.27)$$

Then the accuracy of the k-NN is

$$\theta_n = \|\Delta_n^E\|^2 - \|\Delta_n^I\|^2 \quad (3.28)$$

where  $\Delta^E$  and  $\Delta^I$  are nonparametric extra-class and intra-class differences and defined in equation (3.21) and (3.22). k-NN classifies the sample  $x$  more accurately if the difference  $\theta_n$  is large.

The remaining part of the calculations is same.

### 3.5.4 Small Size Problem

Because of the small number of available training face images and complex facial variations, present appearance-based face recognition systems come across difficulty. Human face appearances have many variations resulting from varying lighting conditions, different head poses and facial expressions. In real-world scenarios, only a small number of samples for each subject are available for training. Martinez and Kak [30] have shown that the trying from nondiscriminant techniques (e.g., PCA) to discriminant approaches (e.g., LDA) is not always warranted and may sometimes lead to poor results when small and nonrepresentative training data sets are used.

The main problem of LDA is that it usually faces with the small sample size problem when coping with high dimensional data.  $S_w$  can turn into singular when data does not contain enough training samples. Therefore, LDA can get stuck while computing the projection vectors. For example, LDA needs 10000 training data to guarantee that  $S_w$  is nonsingular for an image dataset consists of 100 width x 100 height sized images where an image vector is 10000 dimensional. Many approaches have been recommended to handle with this problem. But, in the preprocessing phase they all lose some information more or less about the discriminative information that is the common problem of these submitted variant LDA approaches.

LDA supposes each class has a Gaussian distribution with a common covariance matrix that is another disadvantage of the LDA. When each class has a unimodal distribution and separated by scatter of class means, LDA calculates the optimal projection directions. However, if the class distributions are multimodal and share same mean, LDA cannot calculate the optimal projection direction that  $S_b$  has rank  $c - 1$  where  $c$  is the number of class. So that at most  $c - 1$  feature can be extracted. However,  $c - 1$  features are optimal sub solution from the view of Bayes and Fisher's criterion not until selecting the posteriori probability functions.

NNDA does not rely on nonsingularity of the within class scatter matrix where it can be defined as an extension of a nonparametric discriminant analysis. And also, NNDA does not suppose to be true that class distributions are not any parametric distributions where it tries to locate the important directions.

### **3.6 Similarities, Distances and Correlations**

While in the identification process, projected query face image must be decided to the closest or most similar to the projected training face image that reveal the most similarity or the closest distance, identifies the query image. Several questions about several of these measures have been discussed. For original dimensionality in the subspace, L2 norm and Cosine angle measures presented similar results, however L1 norm, Mahalanobis distance and correlation measures resulted different. Several distance and similarity measures are briefly discussed below considering  $x$  and  $y$  as  $N$ -dimensional column vectors for each;

$L_1$  norm is the most simple distance measure and is also called as city block distance.

$L_1$  norm is defined as follows,

$$L_1(x, y) = \sum_{i=1}^N |x_i - y_i|. \quad (3.29)$$

$L_2$  norm is the Euclidean distance measure. It is the sum of squared distances of two vectors.  $L_2$  norm is defined as follows,

$$L_2(x, y) = \sum_{i=1}^N (x_i - y_i)^2 = (x - y)^T (x - y). \quad (3.30)$$

The Mahalanobis distance takes into account the covariance among the variables in calculating eigen space distances. For each vector dimension, the eigen value of that dimension is produced and the results are summed up. The mathematical definition of mahalanobis distance is as follows,

$$Mah(X, Y) = -\sum_{i=1}^m X_i Y_i C_i, \quad \text{where } C_i = \frac{1}{\sqrt{\lambda_i}}. \quad (3.31)$$

Cosine measure is a similarity measure in which cosine angle between two vectors in the subspace is calculated. It is the dot product of the two normalized vectors. Cosine similarity measure is defined as follows,

$$\cos(x, y) = \frac{x \cdot y}{\|x\| \|y\|}. \quad (3.32)$$

Normalized cross correlation is a similarity measure in which correlation is calculated for between two vectors in the subspace via subtracting the mean for every feature and dividing by their standard deviation. The formulation is defined as below,

$$ncc(x, y) = \frac{1}{n-1} \sum_{i=1}^N \frac{(x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y} \quad (3.33)$$

where  $N$  is the number of features in  $x$  and  $y$ . In functional analysis terms, this can be thought of as the dot product of two normalized vectors. That is

$$X = \sum_{i=1}^N (x_i - \bar{x}), Y = \sum_{i=1}^N (y_i - \bar{y}) \quad (3.34)$$

And then the above sum in equation (3.28) equal to

$$ncc(x, y) = \left\langle \frac{X}{\|X\|}, \frac{Y}{\|Y\|} \right\rangle \quad (3.35)$$

where  $\langle \cdot, \cdot \rangle$  is the inner product and  $\|\cdot\|$  is the  $L_2$  norm.





## **4. NNDA BASED FACE RECOGNITION USING ENSEMBLED GABOR FEATURES**

### **4.1 Preface**

Among the different face recognition methodologies, as well as the three most well-known face recognition methods, namely, Eigenface [20], Fisherface [22], and Bayesian Inference [24], the appearance-based approaches have been dominant for years. Especially, fisher discriminant analysis based approaches (FDA) have results promising [22, 36].

The representation of face plays an important role for a successful face recognition system besides the feature extraction and classification algorithms. In last decades, Gabor filter based face descriptors have been accepted as one of the most successful face representation methods. With the various advantages of the Gabor filters, Gabor features based face representation performed very promising results in face recognition area as used in methods the Elastic Bunch Graph Matching (EBGM) [9], Gabor-Fisher Classifier (GFC) [19], and AdaBoosted GFC [37].

### **4.2 Previous Work On Gabor Features**

Labeled graph is constructed on EBGM [9]. Each vertex on the graph corresponds to a predefined facial landmark with fixed high-level semantics, and labeled by the multi-scale, multi-orientation Gabor features calculated from the image area centered at the vertex landmark. And the connection between the two vertices landmarks are the edges and they labeled by the distance between these two landmarks. Identification is done by the elastic matching between the reference graph and the probe one, after the construction of the graph. Elastic graph can model the local facial features well by choosing facial landmarks carefully. Therefore, the most of the local features is defined as well as the overall facial configuration. However, there are two drawbacks of the graph. The high complexity of graph construction,

matching and imprecise landmarks annotation may also decrease recognition performance.

Liu [19] proposed a straightforward method to utilize Gabor features for face recognition. Firstly, for each pixel in the normalized face images (alignment is done via eyes locations) the multi-scale and multi-orientation Gabor features are computed and concatenated to form a high-dimensional Gabor features vector. Then it is uniformly down-sampled to a low-dimensional feature vector. Secondly, for the dimension reduction Principle Component Analysis (PCA) is applied, and for feature extraction is discriminated by enhanced Fisher Discriminant Analysis. Finally, in face identification, these extracted Gabor features are used [19]. Liu's method is simple and does not need to know more facial landmarks except the two eyes. The method named as Gabor Fisher Classifier (GFC), and Liu has experimentally shown its good performance. However, the uniform down-sampling process harms the final classification as rejecting lots of informative Gabor features and reserving many redundant ones while reducing the dimension of the high-dimensional Gabor features to low dimension.

AdaBoosted GFC (AGFC) was proposed to select informative Gabor features optimally to avoid the loss of discriminant Gabor features and the capturing redundant ones via simple down-sampling process. One would understand that AdaBoost is a feature selection tool to reduce the dimension of Gabor features. Fisher discriminant as the final classifier that is applied on the Gabor features selected by AdaBoost. Both GFC and AGFC reduce the feature dimension much, but reducing operation results inevitably as loss of lots of discriminative Gabor features. For instance, for the 64x64 image and 40 Gabor filters as 5 scales and 8 orientations, down-sampling in GFC reduces the dimension from 163,840 to 9,000, and in AGFC is reduced to less than 3,000. Even after down sampling or feature selection, both GFC and AGFC still have to cope with the "small sample size" (3S) problem [38, 39] by using PCA or Null-space analysis carefully as they exploit FDA for feature extraction. No matter which method is used for dimension reduction, losing discriminative information as features is inescapable.

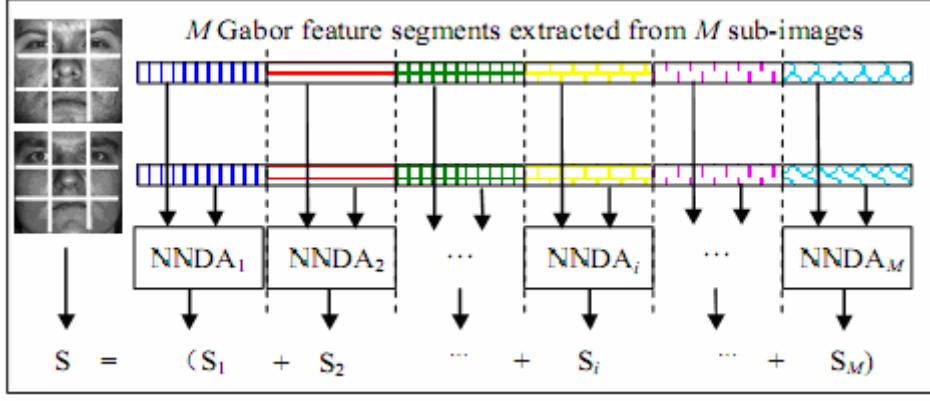
### 4.3 EGNNC

As is well known in face recognition area, Fisher Discriminant Analysis using Gabor feature can achieve promising recognition accuracy where the discriminability is implied with the multi-scale and multi-orientation extracted Gabor features. However, because of 3S problem, Gabor feature has too high dimension for normal FDA. The singularity of within-class scatter matrix is inevitable where the dimension of Gabor feature is always far larger than the number of training samples. Generally, dimension reduction methods and Null-Space methods are the two solutions to this problem. Reducing the original feature dimension to less than the number of the training set is the basic idea for the solution. Common approaches for this problem are down sampling [19], feature selection (e.g. AdaBoost [37]), and PCA [20] or their combination. Null-Space [38, 39] based methods are widely concerned in recent years. However, similar performance is reported methods only based on dimension reduction [38, 39].

In this thesis, the main purpose is increasing the accuracy of the recognition by decreasing the loss of discriminative information as avoiding the above mentioned 3S problem. The proposed method is inspired from [40] and named as Ensemble based Gabor Nearest Neighbor Classifier (EGNNC). Nearest Neighbor Discriminant Analysis (NNDA) was shown to be an efficient nonparametric feature extraction tool from the point of view of nearest neighbor classification. It does not suffer from the small sample size problem and it does not need to estimate any parametric distribution because of its nonparametric nature. Moreover, it does not suffer from the singularity of the within-class scatter matrix as no matrix inversion is required in eigenvector computation. Thus, minimum discriminative information is lost in the feature preparation stage.

EGNNC is an ensemble classifier combining multiple component GNNC classifiers designed as using segments of the Gabor features. In the method, no need to control the Gabor features of a right dimension for each Gabor feature segment for component NNDA contrast to [40], and also no 3S problem occurs contrarily in the FDA.

Figure 4.1 shows how similarity of two faces compute the in the proposed EGNNC method.



**Figure 4.1 :** Illustration of the basic idea of EGNNC

All the sub-images have the same size without overlapping. Therefore, the larger  $M$  is, the fewer features contained in each feature segment.

The Ensemble based Gabor Nearest Neighbor Classifier is formed by the following steps.

### 4.3.1 Grouping the multiple Gabor feature segments

First, face image is divided into  $M$  sub-images according to their spatial locations. Then,  $M$  feature segments are obtained by computing the Gabor features for each sub-image as below;

A set of 40 Gabor kernels are used with the following parameters:  $\sigma = 2\pi$ ,  $k_{\max} = \pi/2$ , and  $f = \sqrt{2}$ . Each sub-image is normalized to zero-mean and unit variance and convolved with the set of 40 Gabor kernels (5 scales and 8 orientations), and each convolution results with a magnitude response of size  $n \times n$ , where  $n$  is the size of the all sub-image. Thus, a total of 40  $n \times n$  magnitude sub-image is obtained. Let the kernel with orientation  $\mu$  and scale  $\nu$  be denoted by  $\psi_{\mu,\nu}(z)$ , with  $z = (x, y)$  denoting a sub-image coordinate. The convolution of the sub-image  $I_i(z)$  with  $\psi_{\mu,\nu}(z)$  is defined as,

$$G_{\mu,\nu}(z) = I_i(z) * \psi_{\mu,\nu}(z) \quad (4.1)$$

Therefore the set  $S = \{G_{\mu,\nu}(z) : \mu \in \{0, \dots, 7\}, \nu \in \{0, \dots, 4\}\}$  forms the Gabor filter representation of the sub-image  $I_i(z)$ .

To benefit from different spatial frequencies (scales), spatial localities and orientation selectivities, these representation results are concatenated and an augmented Gabor feature segment  $X$  is derived. Before the concatenation, each  $G_{\mu,v}(z)$  is downsampled by a factor of  $\rho$  and normalized to zero-mean and unit variance. Downsampling is performed after  $G_{\mu,v}(z)$  is smoothed by a 5x5 Gaussian window. After smoothing  $G_{\mu,v}(z)$ , downsampling is applied by picking up smoothed values by  $\rho/8$  steps from each column and row. Then, a feature vector is formed by concatenating each row (or column) of the final  $G_{\mu,v}(z)$ . Let  $G_{\mu,v}^{(\rho)}$  denote the normalized vector constructed from  $G_{\mu,v}(z)$  (downsampled by  $\rho$  and normalized to zero-mean and unit variance), the augmented Gabor feature segment is defined as follows,

$$x^{(\rho)} = \left\{ G_{0,0}^{(\rho)t} \mid G_{0,1}^{(\rho)t} \mid \dots \mid G_{7,4}^{(\rho)t} \right\}^t, \quad (4.2)$$

where  $t$  is the transpose operator. The augmented Gabor feature segment benefits from different spatial frequencies(scales), spatial localities and orientation selectivities, thus, yielding with a highly discriminating capability.

#### 4.3.2 Designing a NNDA classifier based on each feature segment

The augmented Gabor feature segment defined in equation (4.2) resides in a high dimensionality of  $\mathfrak{R}^{\mathfrak{N}}$ , where  $\mathfrak{N}$  is the dimensionality of vector space. In a typical application,  $\mathfrak{N}$  is as high as 10.240, after downsampling is applied with a factor of 64. However, “perceptual tasks such as similarity judgment tend to be performed on a low-dimensional representation of the sensory data. Low dimensionality is especially important for learning, as the number of examples required for attaining a given level of performance grows exponentially with the dimensionality of the underlying representation space” [41]. PCA is the optimal dimensionality reduction technique in the sense of mean-square error.

Let

$$\Sigma_{x^{(\rho)}} = \mathcal{E} \left\{ \left[ x^{(\rho)} - \mathcal{E} \left( x^{(\rho)} \right) \right] \left[ x^{(\rho)} - \mathcal{E} \left( x^{(\rho)} \right) \right]^t \right\}. \quad (4.3)$$

where  $\mathcal{E}(\cdot)$  is the expectation operator.

The PCA of a random vector  $x^{(\rho)}$  factorizes its covariance matrix  $\Sigma_{x^{(\rho)}}$  into the following form,

$$\Sigma_{x^{(\rho)}} = \Phi \Lambda \Phi^t \text{ with } \Phi = [\phi_1 \phi_2 \dots \phi_D], \Lambda = \text{diag} \{ \lambda_1, \lambda_2, \dots, \lambda_D \}, \quad (4.4)$$

where  $\Phi \in \mathfrak{R}^{D \times D}$  is an orthogonal eigenvector matrix and  $\Lambda \in \mathfrak{R}^{D \times D}$  a diagonal eigenvalue matrix with diagonal elements in decreasing order ( $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_D$ ). An important property of PCA is the optimal signal reconstruction ability in the sense of mean-square error, when a few amount of high order eigenvectors corresponding to largest eigenvalues are used. Therefore, the dimensionality reduction with PCA on the Gabor features of the segment is defined as,

$$y^{(\rho)} = T^t x^{(\rho)}, \text{ where, } T = [\phi_1 \phi_2 \dots \phi_d], d < D \text{ and } T \in \mathfrak{R}^{d \times D} \quad (4.5)$$

The lower dimensional feature vector  $y^{(\rho)} \in \mathfrak{R}^d$  captures the most expressive features of the original data  $x^{(\rho)}$ . However, PCA driven schemes are shown to be useful only with respect to data compression and decorrelation of second order statistics. PCA does not take into account the discrimination aspect of the an done should not expect optimal performance for tasks such as face recognition. One solution has been proposed by Belhumeur et al., the so-called Fisher Linear Discriminant (FDA) approach.

FDA is a popular discriminant analysis tool that maximizes the ratio of the between-class scatter to the within-class scatter. In the Gabor+Fisherfaces scheme, FDA approach discussed in section 3.4 is applied on the augmented Gabor feature vector  $X^{(\rho)}$  and the linear projection matrix  $T$  is constructed by the the first  $d$  leading eigenvectors of  $S_b S_w^{-1}$ , where  $S_b$  is the between-class scatter and  $S_w$  is the within-class scatter matrix of the augmented feature vector  $X^{(\rho)}$ .

After PCA is applied to the augmented feature segment and dimensionality is reduced from  $D$  to  $d$ , NNDA is applied on the resulting Gabor+PCA features to yield with highly discriminating features.

### 4.3.3 Combining all these component with NNDA classifiers

Sum rule used for combination and as NNDA approach is detailed in section 3.5 the training algorithm is given below.

For each segment  $S_j$ , according to same spatial location of images;  $x_i$  is the sub-image of  $i^{\text{th}}$  image in the  $j^{\text{th}}$  segment.

1. Given  $D$  dimensional samples  $\{x_1 | x_2 | \dots | x_N\}$ ,  $d$ -dimensional discriminant subspace is searched.
2. Normalize each sample  $x_i$  to zero-mean and unit variance.
3. Apply a set of 40 Gabor kernels(5 scales and 8 orientations) to each sample  $x_i$ , resulting with  $G_{i,\mu,\nu}(z)$ ;  $\mu \in \{0, \dots, 7\}$ ,  $\nu \in \{0, \dots, 4\}$ ,  $i \in \{1, \dots, N\}$ .
4. Downsample each filter output  $G_{i,\mu,\nu}(z)$  with a factor of  $\rho$  to achieve  $G_{i,\mu,\nu}^{(\rho)}(z)$ , and normalize the final  $G_{i,\mu,\nu}^{(\rho)}(z)$  to zero-mean and unit variance.
5. Concatenate rows(or columns) of each resultant  $G_{i,\mu,\nu}^{(\rho)}(z)$  to form an augmented feature segment  $x_i^{(\rho)}$ .

$$x_i^{(\rho)} = \{G_{i,0,0}^{(\rho)} | G_{i,0,1}^{(\rho)} | \dots | G_{i,7,4}^{(\rho)}\}^t$$

6. Form the final Gabor feature segments matrix  $X^{(\rho)}$  by assembling each  $x_i^{(\rho)}$  in columns, side by side.

$$X^{(\rho)} = \{x_1^{(\rho)} | x_2^{(\rho)} | \dots | x_N^{(\rho)}\}$$

7. Apply PCA on Gabor feature segments matrix  $X^{(\rho)}$  to learn the PCA projection matrix  $T_{pca}$ .

$$T_{pca} = [\phi_1 \phi_2 \dots \phi_{D-1}], T \in \mathfrak{R}^{D-1 \times D}$$

8. Project feature matrix  $X^{(\rho)}$  with the learned PCA model.

$$Y_{pca} = T_{pca}^t X^{(\rho)}$$

9. Apply NNDA on  $Y_{pca}$  to learn the NNDA projection matrix.

$$T_{nda} = [\phi_1 \phi_2 \dots \phi_d], T \in \mathfrak{R}^{d \times D-1}$$

10. Project Gabor+PCA features  $Y_{pca}$  with the learned NNDA model.

$$Y = T_{nda}^t Y_{pca}$$

11. Obtain the final transformation matrix for segment  $j$ .

$$S_j = Y$$

In classification, a new test image  $x'$  divided into  $M$  parts and  $M$  feature segments are obtained by computing the Gabor features for each sub-image in the test image. Then each segment is projected onto to the each Gabor+NNDA feature space with the linear projection matrix  $S_j$  as follows,

$$y'_j = S_j x'_j \quad (4.6)$$

Then, a normalized cross correlation is applied between,  $y'_j$  and for all  $y_j$ , where  $y'_j$  is the  $j^{\text{th}}$  segment of the test image in the reduced gabor space and  $y_j = S_j x_j$  is the  $j^{\text{th}}$  segment of a training image in the reduced gabor space. After adding all correlation results for each segment by sum rule, the maximum correlated training image's label is selected as test image's label.

In the next section, the accuracy of EGNNC will be presented as comparing it with GNNC and NNC on Yale database [22], and as comparing it with GPC, GFC and GNNC methods on a 200 subject subset of FERET database.



## 5. EXPERIMENTS AND RESULTS

Tests has been performed on two databases as Yale database [22], and a subset of FERET Database [42] which consists of 600 images of 200 subjects.

OpenCV library is used [Url-3] for Gabor filters extraction, NNDA, PCA and FDA are implemented on Matlab 7.0.1 R14 platform. A Intel Pentium 4 CPU 3.20 GHz , 1.0 GB RAM PC is used for tests.

For each experiment of EGNNC, normalized cross correlation is used as a metric and nearest neighbor classifier is used for labeling. Normalized cross correlation is meaningfull as metric in the labeling phase where the EGNNC is an ensembled approach and correlation must be considered within the segments.

### 5.1 Results on Yale Database

Yale database [22] was prepared in Yale University CS department.. In database there are 165 grayscale 64x64 face images for 15 people where each person has 11 images in the dataset. One for each subject, there are the following facial expressions or configurations as center-light, w/glasses, happy, left-light, w/no glasses, normal, right-light, sad, sleepy, surprised, and wink. Figure 5.1 displays 8 Images of 2 subject. The 64x64 scaled images aligned by Deng Cai [Url-4] were used in the tests actually original Yale images are of size 320x243.



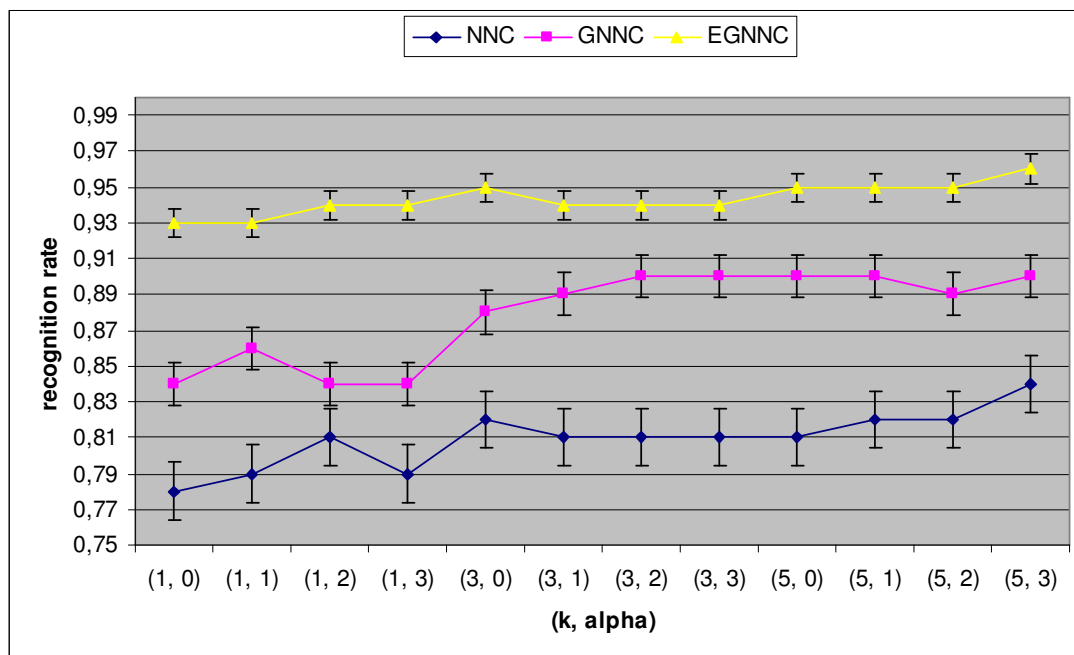
**Figure 5.1:** 2 subjects from Yale Database with expression, lighting variations and occlusion(glasses).

The proposed EGNNC method outperforms its ancestor, GNNC (Gabor Nearest Neighbor Classifier) [43], and NNC (Nearest Neighbor Classifier aka NNDA). In the

experiment, 5 random partitions are formed. The formations are selected randomly as 5 images in the training and the remaining 6 images int testing in each attempt. Average results of the 5 experiments is given.

Parameters of the experiment are the step size of EGNNC is set to 5, and the reduced subspace is of 14 dimensions noting the original dimensionality is  $64 \times 64 = 4,096$ . NNDA parameter  $(k, \alpha) : k \in \{1,3,5\}, \alpha \in \{0,1,2,3\}$  tuples are changed where step size is constant. Figure 5.2 shows the results.

It can be clearly observed that EGNNC outperformed GNNC on all  $(k, \alpha)$  tuples . EGNNC reaches 96 percent accuracy in 14 feature dimension, along with parameters step size = 5,  $k = 5, \alpha = 3$ .



**Figure 5.2:** Performance of EGNNC on Yale as reduced dimensionality is 14 and step size 5

## 5.2 Results on FERET Subset

The FERET database contains 1564 sets of images for a total of 14,126 images that consist of 1199 individuals and 365 duplicate sets of images where collected in 15 sessions between August 1993 and July 1996. A duplicate set is a second set of images of a person already in the database and was usually taken on a different day. The image sets were acquired without any restrictions as having subjects made facial

expressions, rotation in the large interval with various illumination conditions. Images taken on different times during the same photo session at least two frontal image.

Tests are performed on a subset of FERET database. There are 600 face images of 200 subjects such that each subject has three images. The image are size of 256x384 with 256 gray scale levels. First, images are aligned according to manually annotated eye pupils via rotation and scaling transformations. Then, face region is extracted with the size of 128x128, which is further normalized to zero-mean and unit variance in the gabor filtering. Some example images are shown in Figure 5.3 that are used in the experiments.



**Figure 5.3:** Example images used in the FERET experiments.

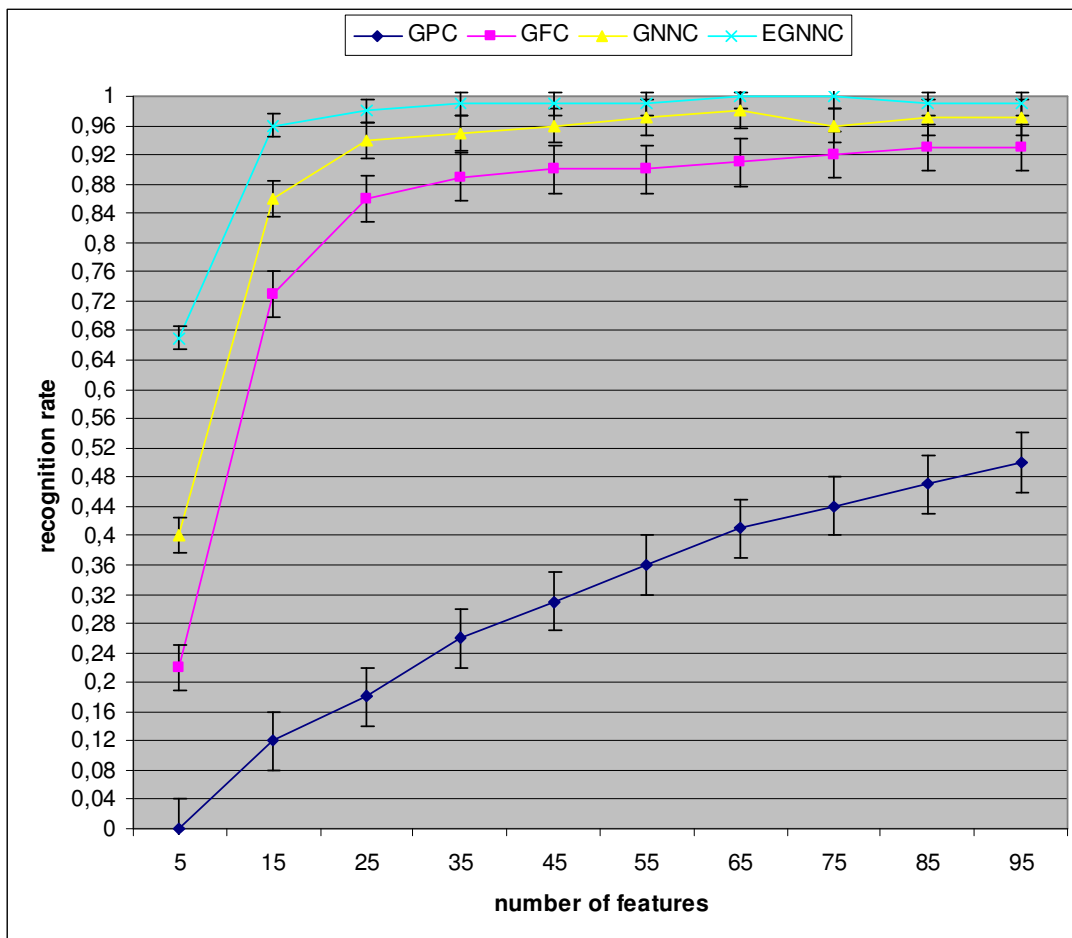
In the subset, images have various lighting conditions and facial expressions. As in the Yale experiment face images are selected randomly for training and testing. The top two rows are training images and the bottom row are test images show in the Figure 5.3.

Experiment 1: The performance of the proposed method EGNNC is compared with its ancestors GPC (Gabor + Eigenface), GFC (Gabor + Fisherfaces) and GNNC (Gabor + NNDA). The parameters are set same with the previous experiment as *step size* = 13, *alpha* = 1, *k* = 1. Results are shown in the Figure 5.4. EGNNC outperforms as recognition rate 100 where GNNC achieves the its highest recognition rate with 98 percent in 65 feature dimensions. Previous methods has already beaten by GNNC as

GFC achieves a 92.6 percent accuracy and GPC achieves 40.6 in the same feature dimension. The average recognition rates of the figure is tabulated in Table 5.1.

**Table 5.1:** Average recognition rates of GPC, GFC, GNNC and EGNNC on 200 class in subset

	Mean	Standard Deviation
GPC	0.3051	0.1610
GFC	0.8270	0.2194
GNNC	0.9021	0.1813
EGNNC	0.9433	0.0088



**Figure 5.4:** Comparison between GPC, GFC, GNNC, EGNNC

Experiment 2: The step size is investigated in the second experiment. Table 5.2 shows the results for constant  $\alpha = 1$  and  $k = 1$  as changing step size.

**Table 5.2:** Step size affect on EGNNC, actually vey small affect

Number of Features	Step size 1	Step Size 5	Step Size 13
5	57	67	67
15	93	96	96
25	97	99	98
35	98	99	99
45	99	99	98
55	100	99	100
65	99	99	100
75	99	99	99
85	100	99	100
95	99	100	99

As it seems on the table, step size has very little effect on results. Therefore, reducing the dimension as extracting features with NNDA on one step is enough where ensembling process prevents loss of discriminative information.



## 6. CONCLUSIONS AND FUTURE WORK

The contribution of this thesis, an Ensemble based Gabor Nearest Neighbor Classifier (EGNNC) method is proposed extending Gabor Nearest Neighbor Classifier (GNNC) [43] where GNNC extracts important discriminant features both utilizing the power of Gabor filters and NNDA as explained below.

2-D Gabor filters have been used in both image processing and computer vision, after the ground-breaking work of extending 1-D Gabor filters to 2-D by Daugman [18]. Gabor filters give the optimized resolution in space-frequency localization and result with illumination, expression and pose invariant image features.

Nearest Neighbor Discriminant Analysis (NNDA) was shown to be an efficient nonparametric feature extraction tool from the point of view of nearest neighbor classification. It does not suffer from the small sample size problem and it does not need to estimate any parametric distribution because of its nonparametric nature. Moreover, it does not suffer from the singularity of the within-class scatter matrix as no matrix inversion is required in eigenvector computation.

EGNNC is an ensemble classifier combining multiple NNDA based component classifiers designed respectively using different segments of the reduced Gabor feature. Since reduced dimension of the entire Gabor feature is extracted by one component NNDA classifier, EGNNC has better use of the discriminability implied in reduced Gabor features by the avoiding 3S problem as making minimum loss of discriminative information. In the method, no need to control the Gabor features of a right dimension for each Gabor feature segment for component NNDA contrast to [40] where using all gabor features leads to redundant information, and also no 3S problem occurs contrarily in the FDA.

For both relative and absolute performance indices, the accuracy of the method is shown. Using a 200 class subset of FERET database covering illumination and expression variations, EGNNC achieved 100% recognition rate, outperforming its

ancestor GNNC [43] perform 98 percent and standard methods such as GFC and GPC [19] for 65 features. The effects of NNDA's training parameters  $k$ ,  $\alpha$  and step size that is previously investigated for GNNC [43] also tested for EGNNC. Theoretical interpret for  $\alpha$  value is consistent with the results as recognition rate is higher for small  $\alpha$  values. The denominator of the equation (3.21) grows more rapidly than the nominator for greater  $\alpha$  values. Therefore, the sample weight  $w_n$  which is used for giving more importance to the samples near to the other class than the samples in the class center, get smaller and accuracy also reduces.

It is stated in [43] that to get higher accuracy, step size parameter of NNDA that controls the stepwise dimensionality reduction process, must be bigger. However, in EGNNC the recognition rates for one step dimension reduction (step size = 1) are really close to the results when step size is 13. Therefore, there is no time consuming in EGNNC as getting similar results for different step sizes, contrast to GNNC [43].

For a more intelligent work of feature gaining, AdaBoost approach can be added before the feature extraction on NNDA as mixing feature selection and feature extraction. AdaBoost is a powerful feature selection tool to reduce the dimension of Gabor features and the new method could be named as AEGNNC. And also instead of applying a simple classification method as nearest neighbor for EGNNC, more sophisticated classification schemes such as Support Vector Machines or Neural Networks can be applied as future work.



## REFERENCES

- [1] **Chellappa R., Wilson C.L., and Sirohey S.**, 1995. "Human and machine recognition of faces: A survey," Proc. IEEE, vol. **83**, pp. 705–740.
- [2] **Wechsler H., Phillips P., Bruce V., Soulie F., and Huang T.**, 1996,. Face Recognition: From Theory to Applications, Springer-Verlag.
- [3] **Zhao W., Chellappa R., Rosenfeld A., and Phillips P.J.**, 2000. "Face recognition: A literature survey," CVL Technical Report, University of Maryland.
- [4] **Gong S., McKenna S.J., and Psarrou A.**, 2000. Dynamic Vision: from Images to Face Recognition, Imperial College Press and World Scientific Publishing.
- [5] **Abate, A. F., Nappi, M., Riccio, D., Sabatino, G.**, 2006. 2D and 3D Face Recognition: A Survey, Pattern Recognition Letters, vol. **28**, no. 14, pp.1885-1906.
- [6] **Hietmeyer R.**, 2000. "Biometric identification promises fast and secure processing of airline passengers," The Int'l Civil Aviation Organization Journal, vol. **55**, no. 9, pp. 10–11.
- [7] **Phillips P.J., Grother P., Micheals R.J, Blackburn D.M., Tabassi E, and Bone J.M.**, 2003. "FRVT 2002: Overview and summary,"
- [8] **Kanade T.**, 1973. Picture Processing by Computer Complex and Recognition of Human Faces, PhD thesis, Kyoto University.
- [9] **Wiskott L., Fellous J.M., Kruger N., and von der Malsburg C.**, 1997. "Face recognition by elastic bunch graph matching," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. **19**, no. 7, pp. 775–779.
- [10] **Lades M., J. Vorbruggen C., Buhmann J., Lange Jorg, Malsburg C., Wurtz R. P., and Konen W.**, 1993. "Distortion invariant object recognition in the dynamic link architecture," IEEE Trans. Computers, vol. **42**, no. 3, pp. 300–310, Jan.
- [11] **Cootes T.F., Edwards G.J., and Taylor C.J.**, 1998. "Active appearance models," in Proc. European Conference on Computer Vision, vol. **2**, pp. 484–498.
- [12] **Cootes T.F., Edwards G.J., and Taylor C.J.**, 1998. "Face recognition using active appearance models," in Proc. European Conference on Computer Vision, vol. **2**, pp. 581–695.
- [13] **Volker Blanz, Sami Romdhani, and Thomas Vetter**, 2002. "Face identification across different poses and illuminations with a 3D morphable model," in Proc. IEEE International Conference on Automatic Face and Gesture Recognition, pp. 202–207.

- [14] **Donato, G., Bartlett, M. S., Hager, J. C., Ekman, P., Sejnowski, T. J.**, 1999. Classifying Facial Actions, *IEEE Trans. PAMI*, vol. **21**, no. 10, pp.974-989.
- [15] **Tian, Y., Kanade, T., Cohn, J. F.**, 2001. Recognition Action Units for Facial Expression Analysis, *IEEE Trans. PAMI*, vol. **23**, no. 2, pp.97-115.
- [16] **Zheng, W., Zhou, X., Zou, C., Zhao, L.**, 2006. Facial Expression Recognition Using Kernel Canonical Correlatin Analysis(KCCA), *IEEE Trans. Neural Networks*, vol. **17**, no. 1, pp.233-238.
- [17] **Shen L., Bai L., and Fairhurst M. C.**, 2007. “Gabor wavelets and General Discriminant Analysis for face identification and verification,” *Image Vision Comput.*, vol. **25**, no. 5, pp. 553–563.
- [18] **Daugman, J. G.**, 1985. Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two-Dimensional Cortical Filters, *Journal of Opical. Society of America*, vol. **2**, no. 7, pp. 1160-1169.
- [19] **Liu, C., Weschler, H.**, 2002. Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition, *IEEE Trans. On Image Processing*, vol. **11**, no. 4, pp. 467-476.
- [20] **Turk M. and Pentland A.**, Mar. 1991, “Eigenfaces for recognition,” *Journal of Cognitive Neuroscience*, vol. **3**, no. 1, pp. 71–86.
- [21] **Bartlett M.S., Lades H.M., and Sejnowski T.J.**, 1998, “Independent component representations for face recognition,” in *Proceedings of SPIE*, vol. **3299**, pp. 528–539.
- [22] **Belhumeur P. N., Hespanha J. P., and Kriegman D. J.**, Jul. 1997, “Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. **19**, no. 7, pp. 711–720.
- [23] **Swets D. L. and Weng J.**, 1996, “Using discriminant eigenfeatures for image retrieval,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. **18**, no. 8, pp. 831–836.
- [24] **Moghaddam B., Pentland A.**, 1998. “Beyond Eigenfaces: Probabilistic Matching for Face Recognition”, *IEEE International Conference on Automatic Face and Gesture Recognition, AFGR’98*, pp. 30-35.
- [25] **Kirby M. and Sirovich L.**, Jan. 1990, “Application of the Karhunen-Loève procedure for the characterization of human faces,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. **12**, no. 1, pp. 103–108.
- [26] **Pentland A., Moghaddam B., and Starner T.**, Jun. 1994, “View-based and modular eigenspaces for face recognition,” in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 84–91.
- [27] **Moghaddam B.**, Feb. 2002, “Principal manifolds and probabilistic subspaces for visual recognition,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. **24**, no. 6, pp. 780–788.

- [28] **Hyvarinen A.**, 1999. "Survey on independent component analysis," *Neural Computing Surveys*, vol. **2**, pp. 94–128.
- [29] **Hyvarinen A.**, 1999. "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Networks*, vol. **10**, no. 3, pp. 626–634.
- [30] **Martinez A.M. and Kak A.C.**, 1999. "PCA versus LDA," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. **23**, no. 2, pp. 228–233.
- [31] **Liu X., Srivastava A. and Gallivan K.**, 2004. Optimal linear representations of images for object recognition. *IEEE Trans. Pattern Anal. Machine Intell.*, **26**(5):662–666.
- [32] **Zhang W. and Chen T.**, 2003. Classification based on symmetric maximized minimal distance in subspace (SMMS). In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [33] **Yu H. and Yang J.**, 2001. A direct LDA algorithm for high-dimensional data with application to face recognition. *Pattern Recogn.*, **34**:2067–2070.
- [34] **Fukunaga K. and Mantock J.**, 1983. Nonparametric discriminant analysis. *IEEE Trans. Pattern Anal. Machine Intell.*, **5**:671–678.
- [35] **Bressan M. and Vitria J.**, 2003. Nonparametric discriminant analysis and nearest neighbor classification. *Pattern Recogn. Lett.*, **24**:2743–2749.
- [36] **Zhao W., Chellappa R. and Krishnaswamy A.**, 1998. "Discriminant Analysis of Principal Components for Face Recognition", *IEEE International Conference on Automatic Face and Gesture Recognition, AFGR'98*, pp. 336-341.
- [37] **Shan S., Yang P., Chen X., Gao W.**, 2005. AdaBoost Gabor Fisher Classifier for Face Recognition, *Proceeding of IEEE International Workshop on Analysis and Modeling of Faces and Gestures, AMFG 2005, LNCS 3723*, pp.278-291.
- [38] **Cevikalp H., Neamtu M., Wilkes M. and Barkana A.**, 2005. "Discriminative Common Vectors for Face Recognition", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.**27**, no.1, pp.4-13.
- [39] **Huang R., Liu Q., Lu H., and Ma S.**, 2002. "Solving the small size problem of LDA", *International Conference on Pattern Recognition, ICPR'02*, vol.**3**, pp.29-32.
- [40] **Su Y., Shan S., Chen X., and Gao W.**, 2006. "Hierarchical Ensemble of Gabor Fisher Classifier for Face Recognition". In *FGR 2006*, pages 91–96.
- [41] **Edelman S.**, 1999. *Representation and Recognition in Vision*, MIT Pres, Cambridge, MA, USA.
- [42] **Phillips P. J., Moon H., Rizvi S. A., Rauss P. J.**, 2000. The FERET Evaluation Methodology for Face-Recognition Algorithms, *IEEE Trans. on PAMI*, vol. **22**, no. 10, 1090-1104.
- [43] **Kirtac K., Dolu O., Gokmen M.**, 2008. Face Recognition By Combining Gabor Wavelets and Nearest Neighbor Discriminant Analysis, *International Symposium on Computer and Information Sciences, ISCIS' 2008, Issue 27-29 Oct. 2008, Page(s):1 – 5*.

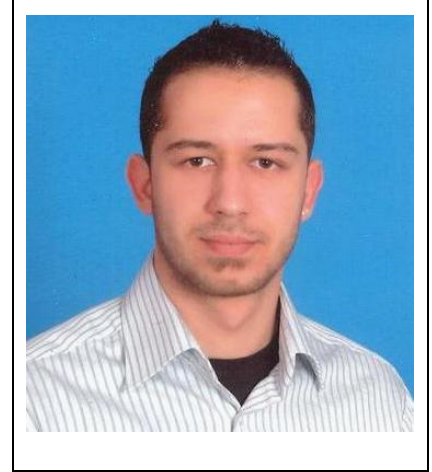
**Url-1 International Biometric Group**, <<http://www.biometricgroup.com>>, accessed at 17.04.2009.

**Url-2 ORL face database**, <<http://www.uk.research.att.com/facedatabase.html>>, accessed at 02.10.2009.

**Url -3 Zhou, M.**, <<http://www.personal.rdg.ac.uk/~sir02mz>>, accessed at 10.10.2009.

**Url-4 Cai, Deng.**, <<http://www.cs.uiuc.edu/homes/dengcai2/Data/FaceData.html>>, accessed at 21.09.2009.

## **CURRICULUM VITA**



**Candidate's full name:** Onur DOLU

**Place and date of birth:** 01.01.1984

**Permanent Address:** Saray Mah. Harput Sok.  
Antrium Rezidans Sitesi  
Blok:2 Kat:5 No:24 Umraniye / İST

**Universities and  
Colleges attended:** B.Sc., Computer Engineering,  
Istanbul Technical University

### **Publications:**

**Kirtac K., Dolu O., Gokmen M.**, 2008. Face Recognition By Combining Gabor Wavelets and Nearest Neighbor Discriminant Analysis, International Symposium on Computer and Information Sciences, ISCIS' 2008, Issue 27-29 Oct. 2008, Page(s):1 – 5.