**İSTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY**

# GABOR FEATURE BASED FACE RECOGNITION USING NEAREST NEIGHBOR DISCRIMINANT ANALYSIS

**M.Sc. Thesis by**
**Kadir KIRTAÇ, B.Sc.**

**Department : Computer Engineering**
**Programme : Computer Engineering**

**JUNE 2008**

**İSTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY**

# GABOR FEATURE BASED FACE RECOGNITION USING NEAREST NEIGHBOR DISCRIMINANT ANALYSIS

**M.Sc. Thesis by**
**Kadir KIRTAÇ, B.Sc.**
**504061517**

**Date of submission : 5 May 2008**
**Date of defence examination : 12 June 2008**

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

**JUNE 2008**

**İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ**

# EN YAKIN KOMŞU AYRIŞIM ANALİZİ KULLANARAK GABOR ÖZNİTELİKLERİ TABANLI YÜZ TANIMA

**YÜKSEK LİSANS TEZİ**

**Müh. Kadir KIRTAÇ**
**504061517**

**Tezin Enstitüye Verildiği Tarih : 5 Mayıs 2008**
**Tezin Savunulduğu Tarih : 12 Haziran 2008**

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

**HAZİRAN 2008**

**ACKNOWLEDMENTS**

I would like to thank to my supervisor, Prof. Muhittin Gökmen, for his guidance throughout my research process and during the preparation of this thesis.

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

I would also like to thank to TÜBİTAK(The Scientific and Technological Research Council of Turkey) for supporting me during my Master study under the grant "National Scholarship Programme for Master Science Students".

And finally, I would like to thank Xipeng Qiu, Onur Dolu, Fatih Kahraman and Abdulkerim Çapar for valuable discussions and for sharing their knowledge with me.

May, 2008 Kadir KIRTAÇ

**CONTENTS**

## ABBREVIATIONS

**LDA** : Linear Discriminant Analysis
**PPLS** : Parametric Piecewise Linear Subspace
**ICA** : Independent Component Analysis
**PCA** : Principal Component Analysis
**GDA** : Generalized Discriminant Analysis
**DLA** : Dynamic Link Architecture
**SVM** : Support Vector Machine
**NNDA** : Nearest Neighbor Discriminant Analysis
**NDA** : Nonparametric Discriminant Analysis
**NLDA** : Null-space Linear Discriminant Analysis
**EBGM** : Elastic Bunch Graph Matching

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF SYMBOLS

| | |
|---|---|
| **T(r)** | **:** Linear transformation function |
| **N** | **:** Number of training images |
| $\mathbf{x_i}$ | **:** i-th training sample in a training set |
| **W** | **:** Linear transformation matrix |
| $\mathbf{y_k}$ | **:** k-th projected training sample in a training set |
| **m** | **:** Mean vector of the training images |
| $S_T$ | **:** Total scatter matrix of the training set |
| $\lambda$ | **:** Eigenvalues of the covariance matrix |
| **V** | **:** Eigenvectors of the covariance matrix |
| $\beta_i$ | **:** i-th difference image in Eigenfaces approach |
| **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 the sample $x_n$ in NNDA |
| **k** | **:** Nearest neighbor parameter of *k*-nearest neighbor classification |
| $x^I_{([k/2]+1)}$ | **:** The intra-class $\left([k/2]+1\right)$-th nearest neighbor of the sample *x* |
| $x^E_{[k/2]}$ | **:** The $[k/2]$-th extra-class nearest neighbor of the sample *x* |
| $\psi_{\mu,v}$ | **:** Gabor kernel with scale $\mu$ and orientation *v* |
| $\sigma$ | **:** Standard deviation of the Gaussian part of the Gabor kernel |
| **f** | **:** Spacing factor between Gabor kernels |
| $k_{\mu,v}$ | **:** Wave vector of the Gabor kernel with scale $\mu$ and orientation *v* |
| $J^I$ | **:** Gabor jet representation of image *I* |
| $X^{(\rho)}$ | **:** Augmented Gabor feature vector with a downsampling factor $\rho$ |

## GABOR FEATURE BASED FACE RECOGNITION USING NEAREST NEAIGHBOR DISCRIMINANT ANALYSIS


## SUMMARY

Face recognition is one of the most frequent tasks man accomplish every day with no effort. Due to the highly informative and discriminative nature of the face stimuli, the brain uses the visual information formed with the help of the eyes as a biometric identifier.

Computer vision is inspired by this challenging function of the brain and aims to mimic this function to automatically identify people using their facial images. The face recognition problem can be stated as, given an input still image or a video sequence, identify or verify one or more individuals in the input, using a database containing face images of known individuals.

One of the successful approaches in face recognition is the Gabor feature based approach. The importance of the Gabor filters lie under the fact that the kernels are similar to the 2-D receptive field profiles of the mammalian cortical cells, offering spatial locality, spatial frequency and orientation selectivity. The Gabor filters representation of facial images were claimed to be robust to illumination and facial expression variations in many works. In this thesis, a brief overview on the state of the art Gabor feature based methods is presented, and a new combination of a Gabor feature based method is proposed, Gabor+NNDA. It applies the Nearest Neighbor Discriminant Analysis, to the augmented Gabor feature vectors obtained by the Gabor filters representation of facial images. To make use of all the features provided by different Gabor kernels, each kernel output is concatenated to form an augmented Gabor feature vector. Instead of applying NNDA in the original dimensionality, Principal Component Analysis(PCA) is first applied to the augmented Gabor feature matrix, and then NNDA is applied in the resulting Gabor+PCA feature space. As PCA is an optimal data decorrelation method in the minimum mean square error sense, no discriminative information is lost by applying PCA first, and the training time complexity is significantly reduced by applying NNDA in the reduced Gabor+PCA feature space. The feasibility of the proposed method has been succesfully tested on Yale database by giving a comparison with its predecessor NNDA. The effectiveness of the proposed method is shown by a comparative performance study against standard face recognition methods such as the combination of Gabor and Eigenfaces method and the combination of the Gabor and Fisherfaces method, using a subset of the FERET database, containing a total of 600 facial images of 200 subjects exhibiting both illumination and facial expression variations. The achieved 98 percent recognition rate in the FERET test shows the efficiency of the proposed method.

## EN YAKIN KOMŞU AYRIŞIM ANALİZİ KULLANARAK GABOR ÖZNİTELİKLERİ TABANLI YÜZ TANIMA

## ÖZET

Yüz tanıma insanların günlük yaşamlarında zorlanmadan ve sıklıkla gerçekleştirdikleri görevlerden biridir. Yüz uyaranının oldukça bilgilendirici ve ayırt edici özelliğinden dolayı, beyin gözlerin oluşturduğu görsel bilgiyi biometrik tanımlayıcı olarak kullanır.

Bilgisayarlı görü, insanları tanımlamak için yüz görüntülerini kullanarak beynin bu kompleks fonksiyonunu taklit etmeyi amaçlar. Yüz tanıma problemi, bir yüz görüntüsü veya yüz görüntüleri içeren bir video kaydı girdi olarak verildiğinde, bilinen kişilerin yüz görüntülerini içeren bir veritabanı kullanılarak girdideki bir ya da daha fazla yüz görüntüsünün tanımlanması veya doğrulanması olarak ifade edilebilir.

Yüz tanımada başarılı yaklaşımlardan biri Gabor öznitelikleri tabanlı yaklaşımdır. Gabor süzgeçlerinin önemi, Gabor çekirdeklerinin memelilerin görme sinirlerindeki iki boyutlu kortikal hücre profillerine oldukça benzemesi, ve önemli ölçüde uzaysal lokalite, uzaysal frekans ve yönelim seçilimi sunmasıdır. İki boyutlu görüntülerin Gabor süzgeçleri temsillerinin aydınlanma ve yüz ifadeleri değişimlerine karşı dayanıklı oldukları bir çok çalışmada gösterilmiştir. Bu tez çalışmasında, literatürdeki bazı önemli Gabor öznitelikleri tabanlı yöntemler incelenmekte ve Gabor+NNDA adında yeni bir Gabor öznitelikleri tabanlı yöntem önerilmektedir. Önerilen yöntemde, yüz görüntülerine bütünsel olarak Gabor süzgeçlerinin uygulanması ile elde edilen artırılmış Gabor öznitelik vektörlerine En Yakın Komşu Ayrışım Analizi(EYKAA) uygulanmaktadır. Artırılmış Gabor öznitelik vektörü, yüz görüntüsüne farklı ölçek ve yönelimlerde Gabor süzgeçleri uygulanmasıyla elde edilen konvolusyon çıktılarının birleştirilmesiyle elde edilmektedir. Artırılmış Gabor öznitelik matrisine EYKAA uygulanmadan önce Temel Bileşenler Analizi(TBA) uygulanır ve sonuçlanan Gabor+TBA öznitelik uzayında EYKAA uygulanır. TBA yöntemi en düşük ortalama karesel hata açısından en uygun veri dekorelasyon yöntemi olduğundan, başlangıçta TBA uygulanması veri kümesindeki ayrıştırıcı bilginin kaybına yol açmaz. Sonuçlanan Gabor+TBA öznitelik uzayında EYKAA eğitimi uygulanmasıyla da orijinal boyutta eğitim uygulanmasının gerektireceği yüksek zaman maliyetinden kaçınılmaktadır. Önerilen yöntemin uygunluğu, yöntemin çıkış noktası olan En Yakın Komşu Ayrışım Analizi yöntemi ile Yale veri kümesi üzerindeki karşılaştırması ile gösterilmektedir. Ayrıca yöntemin etkinliği, yöntemin standart yöntemler olan "Gabor ve Fisher yüzleri birleşimi" ve "Gabor ve Özyüzler birleşimi" yöntemleri ile, FERET veri kümesinin aydınlanma ve yüz ifadesi değişimleri içeren 200 sınıflı bir altkümesi üzerinde performans karşılaştırmaları yapılarak gösterilmektedir. FERET veri kümesi üzerinde elde edilen yüzde 98'lik tanıma oranı önerilen yöntemin etkinliğini göstermektedir.

# 1. INTRODUCTION

Due to the highly informative and discriminative nature of the face stimuli, face recognition has been considered as a biometric identification application in computer vision community. Face recognition has attracted the attention of researchers from broad areas including computer vision, image processing and computational neuroscience fields. Significant amount of attention is paid to face recognition in the last two decades, however automatic face recognition still has many problems. The reason of growing attention is the need of identity verification in digital world, the expanding interest for security applications in public, and to make use of facial modelling and analysis techniques in human-computer interaction.

As hardware prices got cheaper and with the available technology after thirty years of research, many commercial applications have been developed and the hope for full-automatic face recognition became strong. Face recognition can be considered as a user-friendly biometric identification system when compared to other reliable biometric identification systems, such as iris recognition and fingerprint analysis. For example, in iris recognition or fingerprint analysis, one has to spend effort to input his biometric information to the system, whereas in face recognition, no user cooperation or effort is needed. Other examples of user-cooperative identification and verification systems are atm of a bank, a computer requiring password many of today's websites inquiring user password.

Face recognition problem can be stated as: given an input still image or a video sequence, identify or verify one or more individuals in the input, using a database containing face images of known individuals. The solution of the problem is to extract the face image(s) from the scene first, the so-called face detection, then the normalization or alignment of the extracted facial image(s), feature extraction from normalized face(s) in the next step, and verification or identification in the final phase. Figure 1.1 shows a sketch of the given solution [1,2].

Face recognition systems can be classified into two groups: recognition from still images and recognition from video sequences. These two group of applications differ with the quality of images they use, the segmentation algorithms used for background removal, and the recognition or matching criteria.

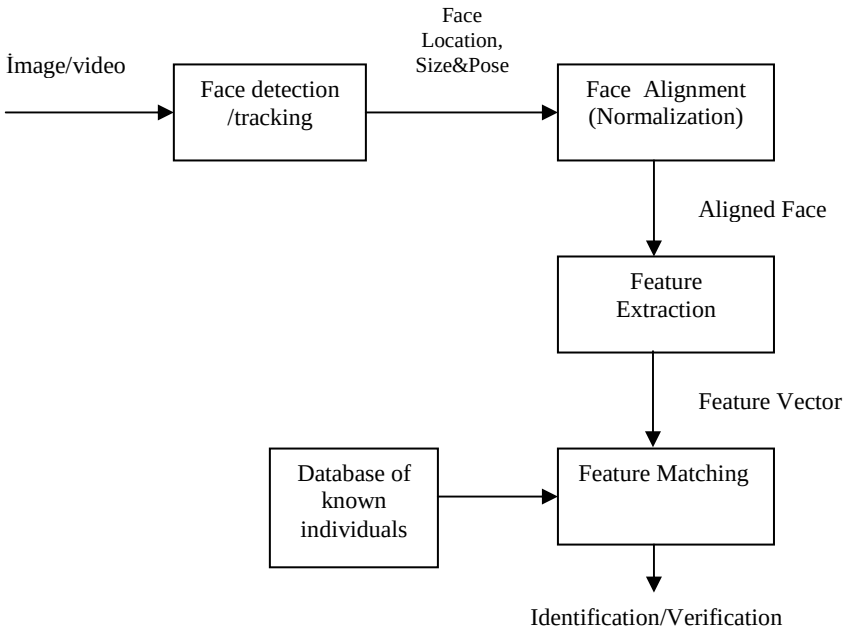



**Figure 1.1:** Block diagram of a generic face recognition system

Identification involves identifying an unknown query person using a database of known individuals, whereas verification is to decide whether the input face is the claimed identity, and then approve or reject the face.

Face detection is the segmentation of the facial image from the background. In recognition from video, the detected face is tracked using a face tracking algorithm. Face alignment is the localization of the facial image, using specified features such as eyes, nose and mouth. After the localization, using the predefined locations, the facial images are normalized with respect to properties such as pose and size, using geometrical affine transformations. Facial images can be further normalized with respect to illumination effects, using histogram equalization. A further statistical normalization can be applied so that each pixel value will have zero mean and unit variance, resulting with values in a specified range, such as [0,1] range. After the normalization procedure, feature extraction is applied so that the identification will be performed using the resulting important features, the so-called feature vector.

The feature vector is matched with a known individual against all individuals of the database with a specific confidence, if the confidence is enough for identification, the system reports the matched identity, else the system reports an unknown face.

Face recognition performance depends on the features extracted from facial images and pattern classification methods that use the extracted features to classify the faces, so the feature extraction and classification problems are the key concepts in subspace face recognition.

### 1.1. Face Recognition in Subspaces

Subspace analysis of images for face recognition is based on the fact that face image resides in a subset of the input image space. For example, a 100x100 image has 10,000 pixel values, and many classes of objects can be represented with a $256^{10,000}$ total of combinations of these pixel values. Thus, face can be regarded as one of the classes of these objects and is said to reside in a subset of the image space, called as face subspace.

Two of the most popular subspace methods are Eigenfaces [3], which is based on principal component analysis, and Fisherfaces [4], which is based on linear disriminant analysis. Both of the methods will be investigated in Section 2.

The distribution or manifold of all faces explains the variations in facial appearances of individuals, whereas the nonface manifold explains everything but faces. These manifolds are highly nonconvex and nonlinear. Face recognition can be considered as the task of distinguishing between faces in the face manifold, whereas face detection is the task of distunguishing between the face and nonface manifold in image space [1]. Several subspace methods will be discussed later in the next sections.

### 1.2. Challenges in Face Recognition

Automatic face recognition can be considered as a complicated pattern classification problem. The problem is even more difficult to solve when the search is the done among the individuals belonging to the same class. Moreover, in most practical applications, no more than one training image per class is available and the images are acquired under uncontrolled conditions. The researchers have focused on developing robust classifiers to tackle with illumination and pose variations images have, whereas less effort has been made to deal with occlusion and aging effects.

### 1.2.1. Varying Illumination

The ambient light can change dramatically during the day indoor and outdoor. Because of the 3D nature of the face, the part of the face that is exposed to light direction become highlighted and the other part of the face is diminished under the shadow. This poses a wide variation even between the face images of the same class. So, "the variations between the images of the same face due to illumination and viewing direction are almost always larger than the image variation due to change in face identity" [5].

To cope with illumination effects, a two-way solution can be suggested. The first one is to normalize input images, and second is to develop effective feature extraction methods to deal with the problem. Some of the popular normalization methods are mean value normalization, histogram equalization and illumination correction [1]. One of the basic normalization operation is contrast stretching [6]. It is a simple linear stretching in which the original range of the input image is linearly transformed to a specified full range of, say [0,255], using a linear transformation function *T(r)*. In histogram equalization(linearization), the contrast of the input image is increased so that the intensities of the resulting image will be better distributed on the histogram. This will result with a contrast increase in the lower contrast part without affecting the global contrast. The result is obtained by by spreading out the most frequent intensity values in the image. Figure 1.2 shows a sample of linearly streched and histogram equalized images [1].

Another simple illumination correction operation is a least–squares method in which the best fitting plane $I'(x, y)$ is desired. $I'(x, y)$ can be defined as follows,

$$I'(x, y) = a \times x + b \times y + c, \quad \textbf{(1.1)}$$

where coefficients *a*, *b* and *c* can be estimated using a least-squares method; and then the illumination is corrected in the resulting difference image $I''(x, y)$ as follows,

$$I''(x, y) = I(x, y) - I'(x, y). \quad \textbf{(1.2)}$$

(a) (b) (c)

**Figure1.2:** Example of histogram equalization and linear stretching (a) original input image *I(x,y)* (b) linearly strecthed image (c) histogram equalized image

A considerable amount of research has been conducted on face recognition under varying illumination.

In [7], Adini et al. have presented an empirical study in which 3 class of different image representations are evaluated against changes in illuminations. These 3 classes of representations are edge maps, one dimensional and two dimensional derivatives of images and the images convolved with 2-D Gabor filters. They reported that none of the representations could be succesful enough to cope with illumination variation and direction. They also reported similar results for the images exhibiting variation in viewpoint and expressions.

Extending the edge map respresentation of [7], Gao et al. presented Line Edge Map(LEM) representation [8] in which face contours are extracted and combined in segments and then organized in lines. In this study, they have evaluated the novel representation under controlled conditions, varying illumination conditions, varying expressions and varying pose. They have also proposed a new filtering technique that speeds up the searching process and demonstrated a modified Hausdorff distance to be used with their new feature representation. They reported superior results than Eigenfaces approach [3], but Fisherfaces [4], which can maximize the between-person variability minimizing the within-person differences, still remained superior.

In [9], Georghiades et al. proposed illumination cone model for face recognition under variable illumination. They have utilized the fact that the set of images of an object under arbitrary lighting forms a convex cone in the image space. They further showed how to construct the cone from as few as 3 images of each was under small lighting changes.

In the recognition phase, distances of each test image to each illumination cone was computed, and the face that gave the shortest distance was chosen as the corresponding identity. Distance to each cone was computed by solving a convex optimization problem, since each cone was convex.

In [10], Basri et al. proposed a method based on spherical harmonics. They showed that under any lighting conditions, the set of images of a convex, Lambertian object lies in a nine-dimensional subspace. First, they represented lighting functions using spherical harmonics. They modeled the reflectance functions in analogous with the convolution of each lighting function using a kernel that represents Lambert's reflectance. They showed that 99.2 percent of the kernel's energy lies in first nine component, the zero, first, second order harmonics and so on. They finally showed how to analytically derive this nine-dimensional subspace from a model of an object that includes 3-D structure and albedo. They also discussed how this analytically derived harmonic basis could be used in a linear subspace based object recognition algorithm, instead of using a basis derived by SVD.

In a recent work by Stan Z. Li et al. [11], a novel illumination invariant method based on near-infrared images was proposed. They extracted features from the near-infrared images using Local Binary Patterns. LBP gave a robust solution to the monotonic gray level transform caused by the NIF imaging. They further developed a face matching engine using ada-boosted LBP features. They also presented an LDA-like scheme to further select discriminative LBP features. They both reported superior results than the state-of-the-art LBP study.

### 1.2.2. Varying Pose

In face recognition systems, the pose of the gallery images and probe images can be quite different. For example, the gallery images can be frontal faces, while probe images could have been taken from a camera placed in the corner of the room, viewing the individual from an angular position. The pose variation in gallery and probe sets affects the classification problem dramatically, and challenged researchers in the field. Briefly stating the problem, face recognition across pose develops algorithms to recognize face images with a viewpoint which have not been seen before, e.g., in traning.

In [12], Okada et al. extended linear subspaces method to parametric linear subspaces method. They stored the parametric linear subspace model representations of known individuals in which, each model can be fit to input resulting faces of known people whose head pose is aligned to input face. They have investigated two different linear models: 1) LPCMAP, which covers linear subspaces spanned by principal components of the training images and linear transfer matrices, which links the projection coefficients of training samples onto the subspaces and their corresponding 3D head angles; 2) PPLS, which is an extension of LPCMAP, by using a piecewise linear approach. It is a set of local linear models, each providing continous analysis and synthesis mappings, enabling to generalize to unknown poses by interpolation. The experimental results were shown to be robust to large 3D head pose variations covering 50 degree rotation along each axis. PPLS method was also shown to compress the data significantly, performing better than the LPCMAP approach.

In [13], Gross et al. proposed eigen light-fields approach to cope with pose variation. They used generic training data, to compute an eigenspace of the head light-fields. The projection to the eigenspace is accomplished by setting up a least-squares problem and solving for the projection coefficients. Finally, matching is completed by comparing the probe and gallery eigen light-fields. They have tested their method using a subset of CMU(PIE) database exhibiting pose variations and using a subset of the FERET database. They showed that the proposed method outperformed both the standard eigenface method and the commercial FaceIT system.

In [14], Gokberk et al. proposed a Gabor filters based method for pose estimation. They learned the best frequency and orientation of Gabor filters for feature selection and applied principal component analysis to the filter outputs. For intelligent feature selection, they introduced an intelligent sampling grid approach. They also gave a comparative study of the standard modular eigenface approach, achieving better performance on both pose estimation and face recognition.

In [15], Yue et al. extended spherical harmonics approach to encode pose information. They showed that the basis images of a rotated test image is a linear combination of the basis images at the frontal pose. They used a learning method to recover the harmonic basis images from only one image taken under arbitrary illumination conditions, with the aid of a bootstrap set consisting of 3-D face scans.

For a rotated test image under an arbitrary illumination condition, they first applied image correspondence between the test image and training images. After that, the frontal pose image was warped from the test image and a face was identified for which there existed a linear reconstruction based on basis images that was closest to the test image.

### 1.2.3. Varying Facial Expression

Besides illumination and pose variation in face images, expression variations also causes an important amount of change in facial appearance.

In [16], Donato investigated several methods for classifying twelve facial actions. He showed that Gabor filters based and Independent Component Analysis(ICA) methods performed the best among other methods such as Local Feature Analysis(LCA), LDA and Local PCA.

In [17], Tian et al. presented an Automatic Face Analysis system to analyze facial expressions based on both permanent features(brows, eyes, mouth) and transient facial features(deepening of facial furrows) in a nearly frontal-view face image sequence. They reported recognition rates of 96.4 percent for upper face action units and 96.7 percent for lower face units.

In [18], an interesting study investigating the effects of facial asymmetry in face recognition under varying expression is presented. They have showed that quantified facial asymmetry improved face recognition significantly when combined with conventional methods such as Fisherfaces and Eigenfaces.

In [19], facial expression recognition by Kernel Canonical Correlation Analysis(KCCA) is proposed. They manually locate 34 facial landmarks in each image and transform these points into a labeled graph by using Gabor filters. Moreover, for each training image, they formed a six-dimensional semantic vector describing basic expressions. Learning the correlation between the semantic vector and labeled graph vector is achieved by KCCA. They presented better results than conventional approaches like LDA and GDA.

### 1.2.4. The Occlusion

Another problem in face recognition is occlusion. A serious drawback of appearance-based systems, such as PCA, is their failure to recognize partially occluded objects.

Local approaches have proven to be better in this scheme, as they divide the face into different parts and apply a voting procedure. However, a voting scheme can not achieve an enough success since it does not consider how good a local match is.

In [20], Martinez proposed a probabilistic matching approach to overcome the occlusion effect. He divided the image into *k* local parts, and modeled each part with a Gaussian distribution. Given the mean feature vector and covariance matrix of each part, the probability of a given match could be directly associated with the sum of all *k* mahalanobis distances. He also inspected on the amount of occlusion that his approach could handle and the minimum number of local parts needed to successfully identify the partially occluded object. He showed that the recognition results were nearly same with the non-occluded facial images even the 1/3 of the face was occluded. He also reported that recognition rates decreased when the eye area was occluded instead of the mouth area.

Martinez's approach was for a partially occluded face, in [21] Kurita et al. proposed a neural network based approach that detects and also reconstructs the occluded part of the face. The network is trained with non-occluded images, during the testing original face can be reconstructed by recalling the pixel values. They reported that classification performance did not decrease even 1/3 of the face was occluded.

Moreover, Sahbi and Boujemaa have proposed a method both dealing with facial expression and occlusion effects [22]. They presented a complete framework for face recognition based on salient feature extraction in challenging conditions such as facial expression and occlusion without using an a priori or a learned model. The proposed matching process handled occlusion and facial expression effects using dynamic space warping, which aligns each feature in the query image with its corresponding feature in the gallery set. This made their approach robust to low frequency variations like occlusion, and high frequency changes like expressions, gender, etc. They used a maximum likelihood scheme to make the recognition more accurate. They reported results on ORL and ARF databases, showing that matching procedure could handle little occlusion and rotations.

### 1.2.5. Aging effects

The performance of the many state-of-the-art techniques drops when the time lapse between training and testing images is considered. This shows that the proposed methods do not take into account the aging variations. In some applications, the age of the subject can be simulated to make the system robust to aging variations.

Of the several techniques for age simulation in the literature are, coordinate transformations, exaggeration of 3-D distinctive characteristics, and facial composites; but none of these methods has been used in face recognition framework.

In [23], Lanitis et al. proposed a method based on age functions. Their work aimed at estimating the age of a person in a given image and generating age progression images. Each face image in the database is described by parameters of *b*, and for each individual the best aging function is generated depending on his/her *b*. They have also proposed a face recognition system robust to aging variation. They tested their method on a database of 12 people, with 80 images in the gallery and 85 images in the probe set. The first test they performed was on a training set of mean age 9 and a test set of mean age 24. They reported a 4% classification rate improvement with weighted apperance specific aging simulation method and a 8% improvement with weighted person specific age simulation method. In the second test, they swapped the training and test set of the previous experiment and reported a 12% improvement with weighted apperance specific aging simulation method and a 15% improvement with weighted person specific age simulation method.

Face recognition across aging variations still remains an unexplored research area. An interesting subject would be the prediction of facial apperance of wanted or missing persons.

Table 1.1 is an abbreviated version of the comparisons table in Abate et al. [24], giving a performance comparison of the methods discussed in previous sections.

**Table 1.1:** Experimental results of the discussed methods. The fourth column indicates the maximum number of samples in gallery set and probe set.

| Authors | Name | Database | Image Size | Max \|G\| - Max \|P\| | Time Lapse | Recog. Rate | Expr. | ILL. | POSE | OCCL. | AGE |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Gao et al. [8] | LEM | Bern<br>AR-Face<br>Yale | | 40-160<br>112-336<br>15-150 | no | 72.09%<br>86.03%<br>85.45% | <br><br>yes | <br>yes<br>yes | yes<br>no<br>no | no<br>no<br>no | no<br>no<br>no |
| Okada et al. [12] | Linear Subspaces | ATR-database | | 2821-804 | no | 98.7% | no | no | yes | no | no |
| Gross et al. [13] | Eigen Lights | PIE | | 5304-5304 | no | 36% | no | yes | yes | no | no |
| Martinez [20] | Martinez | AR-Face | 120x170 | 50-150 | no | 65% | no | no | no | yes | no |
| Kurita et al. [21] | Neural Network | AR-Face | 18x25 | 93-930 | no | 79% | no | no | no | yes | no |
| Lanitis et al. [23] | Age functions | PropertyDB | | 80-85 | no | 71% | yes | yes | no | no | yes |

## 1.3. Face Recognition from Intensity Images

Face recognition is such an interesting and complicated task that it has received attention of various researchers from different areas such as psychology, pattern recognition, neural networks and computer vision. Because of this fact, the literature on face recognition is diverse. It is often that a single face recognition system uses a combination of several techniques for feature representation and classification.

In [2], Zhao et al. classified intensity image based face recognition methods into three categories: Holistic matching methods, feature-based(structural) matching methods and hybrid methods.

A review of feature-based matching and hybrid methods will be given in this section, holisting matching methods will be discussed in section 2.

### 1.3.1. Feature-based(structural) Matching Methods

In these methods, local features such as eyes, nose and mouth are extracted and their locations and local statistics are given into a classifier.

In [25], Nefian et al. proposed a Hidden Markov Model(HMM) based method. In their method, the observation vectors of the HMM sytem were extracted by Karhunen-Loeve transform.

They presented a 86% recognition rate on ORL database. They also proposed a novel HMM-based face detection method in the paper.

One of the most succesful works in this area was graph matching method proposed in [26] by Wiskott et al. The graph matching approach is based on the Dynamic Link architure approach proposed in [27] by Lades et al. DLA and elastic graph matching methods are based on Gabor filters which will also be discussed in the next sections of this thesis.

### 1.3.2. Hybrid Methods

Hybrid methods are the combination of holistic and local-feature methods. In [28], Pentland et al. proposed modular Eigenfaces approach. Their work extended the eigenface approach into a multiview face recognition task. They used seperate eigenspaces for different views. They also introduced eigenfeatures which uses eyes, nose and mouth as local features, the so-called eigeneyes, eigenmouth, etc. Their results showed that local feature-based approaches could be very useful when the images contain big variations such as pose.

In [29], in order to overcome the limitations of PCA such as lack of providing local features and production of global non-topographic linear filters, Penev et al. proposed Local Feature Analysis(LFA) approach. In their proposed method, a dense set of local feedforward receptive fields defined at each point of the receptor grid, whose outputs were as decorrelated as possible, were derived. The residual correlations contained by these outputs were further used in the sparsification of the output. The final representation was a local sparse-distributed representation in which only a small number of outputs are active for any given input.

In [30] Lanitis et al. proposed a flexible appearance based method for automatic recognition. Both shape and intensity information is used to identify a face. The statistical shape model is trained on training images using PCA. In classification phase, inter-class variations of the shape model are differentiated from the within-class variations of the shape model by discriminant analysis. Local gray-level models were also built on the shape model to cope with the local appearance changes such as local occlusions. A global shape-free representation was also obtained by using mean shape and PCA. Finally, these three representations, shape parameters, shape-free parameters and local features were used together to compute a Mahalanobis distance.

In [31], Huang et al. proposed a component-based detection and recognition system. Component-based methods decomposes face into several components such as mouth, eyes, nose that are connected by a flexible geometrical model. By using components, the changes of the head pose would affect the components positions and this would be coped with the flexibility of the geometric model. A drawback of the system is the need of a high number of training images containing different pose and illumination variations. In the classification phase, SVM classifier were used. On a set of six subjects; training on 3 images and testing on 200 images, the hybrid system presented a recognition rate of 90 percent.

### 1.4. Gabor Feature Based Face Recognition Using Nearest Neighbor Discriminant Analysis

Gabor filters received much attention in the image processing field, after the pioneering work of J.G.Daugman, extending 1-D Gabor filters to 2-D [32]. The importance of the Gabor filters lie under the fact that the kernels are similar to the 2-D receptive field profiles of the mammalian cortical cells, offering spatial locality, spatial frequency and orientation selectivity [33]. Due to this fact, recognition can be made without correspondence, for instance, no manual annotations on the images is needed. The Gabor filter representation of facial images were claimed to be robust to illumination and facial expression variations [33].

In this thesis, a new combination of a Gabor feature based method is proposed, Gabor+NNDA. It applies the NNDA method [34] to the augmented Gabor feature vectors obtained by the Gabor filter representation of facial images. To make use of all the features provided by different Gabor kernels, each kernel output is concatenated to form an augmented Gabor feature vector. The feasibility of the method has been succesfully tested on the Yale database [4].

The effectiveness of the proposed method is shown in terms of both absolute performance indices and by a comparative performance study against popular face recognition methods such as the combination of Gabor and Eigenfaces method and the combination of the Gabor and Fisherfaces method [33] on a subset of FERET database [35] containing 600 facial images of 200 subjects exhibiting both illumination and facial expression variations.

.

### 1.5. Organization of the Thesis

In Section 2, popular dimensionality reduction methods such as Principal Component Analysis, the so-called Eigenfaces approach [3], Linear Discriminant Analysis, the so-called Fisherfaces approach [4], and a recent method Nearest Neighbor Discriminant Analysis(NNDA) [34] is discussed. In Section 3, 2-D Gabor filters based face recognition is discussed and previous works on Gabor filters based feature extraction and classification is investigated. In Section 4, Gabor+NNDA approach is proposed. In section 5, a comparative study based on the performance values of the Gabor+Eigenfaces, Gabor+Fisherfaces and Gabor+NNDA methods on Yale database and FERET database is presented. In Section 6, conclusions and future work is given.

## 2. DIMENSIONALITY REDUCTION WITH HOLISTIC METHODS

These methods use the whole face region as the input to a recognition system. One of the most popular representation is the Eigenpictures approach, proposed by Kirby and Sirovich [36]. Later, Turk and Pentland proposed Eigenfaces [3] as the first application of eigenpictures in face identification and detection.

### 2.1. Eigenfaces

In Eigenfaces approach, the aim is to find the optimal linear transformation matrix that will maximize the total scatter of the data. The columns of the transform matrix are called principal components, the so-called Eigenfaces. These principal components are the basis vectors of the new subspace, that correspond to the maximum-variance directions in the original image space.

Suppose that we have $N$ training images $\{x_1, x_2, ..., x_N\}$, each of which is represented as column vectors in the $n$-dimensional image space. We search for the optimal $d$-dimensional subspace, that will maximize the total variance of the data.

Hence, we have to find the optimal linear transformation matrix $W \in \mathbb{R}^{d \times n}$, that will transform each $n$-dimensional vector into $d$-dimensional vectors in the new subspace, where $d<n$.

This projection can be mathematically stated as,

$$y_k = W^T x_k, \qquad \textbf{(2.1)}$$

where $k$ = 1,2,…,$N$ and each input vector $x$ is $n$-dimensional, whereas the output vectors $y$, are $d$-dimensional. W is a linear, orthonormal projection matrix.

Total scatter matrix, the so-called covariance matrix, is calculated as,

$$S_T = \sum_{k=1}^{N} (x_k - m)(x_k - m)^T, \quad \textbf{(2.2)}$$

where *m* is the mean of the training images and calculated as follows,

$$m = \frac{1}{N} \sum_{k=1}^{N} x_k. \quad \textbf{(2.3)}$$

The optimal projection matrix that yields the solution is the result of the following optimization problem,

$$W_{opt} = \arg\max_{W} \left| W^T S_T W \right| = [w_1 \mid w_2 \mid ... \mid w_d]. \quad \textbf{(2.4)}$$

where, *w*'s are the column vectors(Eigenfaces) corresponding to the *d* largest eigenvalues of the total scatter(covariance) matrix. The training algorithm of the Eigenfaces approach is given in Figure 1.3.

1. Align the training images $x_1, x_2, \ldots, x_N$ ; each *x* is a *n*-dimensional column vector.
2. Calculate the mean face. $m = \frac{1}{N} \sum_{k=1}^{N} x_k$
3. Calculate the difference images. $\beta_i = x_i - m$
4. Calculate the covariance matrix.

   $S_T = \frac{1}{N} \sum \beta_i \beta_i^T = BB^T, B = [\beta_1 \quad \beta_2 \quad ... \quad \beta_N]$
5. Compute the eigenvalues and eigenvectors of $S_T$. $S_T V = \lambda V$
6. Keep the largest *m* nonzero eigenvalues. Order the eigenvectors from high to low, corresponding to the selected *m* eigenvalues.

   $W = [w_1 \mid ... \mid w_d]$

**Figure 1.3.** Training algorithm of the Eigenfaces approach.

The recognition algorithm is given in Figure 1.4.

1. Project each training vector onto the eigenspace. $x'_k = W^T(x_k - m); k = 1,..,N$
2. Project the test image onto the eigenspace. $y = W^T(Y - m)$
3. Compute the distance from *y* to each training vector in the eigenspace, using a distance measure. $\delta_k = \|y - x'_k\|; k = 1,..,N$
4. The training vector giving the mininum distance is chosen to identify the test image *Y*.

**Figure 1.4.** Recognition algorithm of the Eigenfaces approach.

The *B* matrix, which is composed of the difference images, is of size $n \times n$. When a 100x100 image is considered, *n* becomes as high as 10,000. Thus, this high dimensionality causes a great deal of computational inefficiency. To overcome this problem, Turk and Pentland proposed a solution in which, eigenvectors of the matrix $B^T B$ are used instead [3]. It can easily be noticed that $B^T B$, which is of *MxM* size, is much more computationally efficient, where $M << n$.

Eigenfaces method offers an optimal representation of faces in the sense of mean-square error. Since it is a global method, it provides reduced sensitivity to noise, blurring and small occlusions in the images [2].

However, PCA-based methods are prone to localization errors, so the alignment step is important. In [24], Martinez proposed a method for modeling localization error. In [37], Yambor has investigated different eigenvector selection mechanisms in PCA, and reported results on FERET images and on a cat&dog database [37]. It has been shown that discarding the leading eigenvector corresponding to the greatest eigenvalue, gave slightly beter recognition results. It was showed that the leading eigenvector contained illumination-variant information, thus degrading the performance when there exist illumination variations among training and testing images [37].

### 2.2. Fisherfaces

Offering a better discrimination criterion, Fisherfaces method was proposed by Belhumeur et al. [4]. The method is aimed at maximizing the between-class variations while miniziming the within-class variations among the images.

Considering a *c*-class problem, between-class scatter matrix is defined as,

$$S_b = \sum_{i=1}^{c} N_i (m_i - m)(m_i - m)^T, \tag{2.5}$$

where $N_i$ is the number of samples in class *i*, $m_i$ is the mean vector calculated from the samples of class *i*, and $m$ is the global mean calculated from all training images. Similarly, the within-class scatter matrix is defined as,

$$S_w = \sum_{i=1}^{c} \sum_{k=1}^{N_i} (x_k - m_i)(x_k - m_i)^T. \tag{2.6}$$

The optimal linear projection matrix *W* is the matrix that will maximize the ratio between the between-class scatter matrix and the within-class scatter matrix.

$$W_{opt} = \arg\max_{W} \frac{\left|W^T S_b W\right|}{\left|W^T S_w W\right|} = [w_1 \mid w_2 \mid ... \mid w_d]. \tag{2.7}$$

This projection matrix is constructed by the *d* greatest eigenvectors calculated from the $S_b S_w^{-1}$ matrix. As the dimensionality of $S_w$ is $n \times n$ and its rank is usually less than *N*-*c*, where c is the number of classes and *N* is the number of total samples in training set, $S_w$ becomes singular and its inverse can not be computed. In Fisherfaces approach, PCA is first applied to reduce the dimensionality to *N*-1. After applying PCA, the projection matrix is constructed over the PCA-transformed data, resulting with a total number of *d* eigenvectors, where *d* = *c*-1. A disadvantage of this approach is the maximum number of features extracted are limited by the number of classes, because the rank of $S_b$ is at most *c*-1. Another drawback is the small sample size problem.

As the distributions of classes are estimated by samples, one training image per sample is not sufficient, thus the method requires more than one sample for each class in training. A good study comparing PCA and LDA can be found at [38].

### 2.3. Nearest Neigbor Discriminant Analysis

Nearest neighbor discriminant analysis(NNDA) is a multi-exemplar, nonparametric feature extraction method proposed by Qiu and Wu [34]. It is an efficient eigen decomposition method similar to LDA. It forms the between-class and within-class scatter matrices in a nonparametric way and it does not depend on the nonsingularity of the within-class scatter matrix.

#### 2.3.1. NNDA Criterion

Considering a $c$-class problem with classes $C_i\{i=1,2,\ldots,c\}$, the extra-class and intra-class neighbor of a sample $x_n \in C_i$ is defined as follows respectively,

$$x_n^E = \arg\min_z \left\| z - x_n \right\|, \forall z \notin C_i \quad \textbf{(2.7)}$$

$$x_n^I = \arg\min_z \left\| z - x_n \right\|, \forall z \in C_i, z \neq x_n \quad \textbf{(2.8)}$$

The nonparametric extra-class and intra-class distances are defined as follows respectively,

$$\Delta_n^E = x_n - x_n^E, \quad \textbf{(2.9)}$$

$$\Delta_n^I = x_n - x_n^I. \quad \textbf{(2.10)}$$

The nonparametric between-class and within-class scatter matrices are defined as follows respectively,

$$S_B = \sum_{n=1}^{N} w_n (\Delta_n^E)(\Delta_n^E)^T, \quad \textbf{(2.11)}$$

$$S_W = \sum_{n=1}^{N} w_n (\Delta_n^I)(\Delta_n^I)^T, \quad \textbf{(2.12)}$$

where, $w_n$ is defined as

$$w_n = \frac{\left\|\Delta_n^I\right\|^\alpha}{\left\|\Delta_n^I\right\|^\alpha + \left\|\Delta_n^E\right\|^\alpha}. \quad \textbf{(2.13)}$$

$w_n$ is introduced to emphasize the samples in class boundaries and deemphasize the samples in class centers.

Utilizing the fact that the accuracy of the nearest neighbor classification can be directly computed by the equation,

$$\Theta_n = \left\|\Delta_n^E\right\|^2 - \left\|\Delta_n^I\right\|^2, \quad \textbf{(2.14)}$$

Qiu and Wu come to a solution for the computation of the projection matrix *W*,

$$W = \arg\max_W tr(W^T (S_B - S_W)W). \quad \textbf{(2.15)}$$

Thus, the columns of the projection matrix are the *m* leading eigenvectors of $S_B - S_W$, corresponding to the *m* greatest eigenvalues.

As $S_B - S_W$ is in high dimensionality, it is not computationally efficient to compute the eigenvectors of that matrix, instead, they offered to apply PCA first, to reduce the dimension to *N*-1, and then apply NNDA in the *N*-1 dimensional PCA space.

### 2.3.2. Stepwise Dimensionality Reduction

To keep the nonparametric extra-class and intra-class differences of the high dimensional space consisted with the projected extra-class and intra-class differences, Qiu and Wu proposed stepwise dimensionality reduction process. In this scheme, nonparametric extra-class and intra-class differences are recomputed in the current dimensionality. The algorithm of stepwise nearest neighbor discriminant analysis is given in Figure 2.1.

- Given $D$-dimensional samples $\{x_1,...,x_N\}$, $d$-dimensional discriminant subspace is expected to be found.
- Supposing that the projection matrix $\hat{W}$ is found in $T$ steps, the dimensionality of samples is reduced to $d_t$ in step $t$, and $d_t$ meets the consitions $d_{t-1}>d_t>d_{t+1}$ , $d_0=D$ and $d_T=d$.
- For $t=1,\ldots,T$

  (1)calculate the nonparametric between-class scatter $S_B^t$ and within-class scatter matrix $S_W^t$ in the current $d_{t-1}$ dimensional space;

  (2)calculate the projection matrix $\hat{W}_t$ ; $\hat{W}_t$ is $d_t \times d_{t-1}$ matrix.

  (3)project the samples by the projection matrix $\hat{W}_t$ ,

  $x' = \hat{W}_t^T \times x$.

**Figure 2.1.** Stepwise NNDA algorithm

They also extended the method from 1-NN to $k$-NN utilizing $k$-NN classification criterion in the training phase. If the majority(no less than $\lceil k/2 \rceil + 1$) of a sample's $k$-nearest neighbors belong to the same class with it, then the sample will be classified correctly. The intra-class $(\lceil k/2 \rceil + 1)$-th nearest neighbor of the sample $x$ is defined as $x^I_{(\lceil k/2 \rceil + 1)}$ and similarly, the $\lceil k/2 \rceil$-th extra-class nearest neighbor of the same sample is defined as $x^E_{\lceil k/2 \rceil}$. If the distance from $x$ to $x^I_{(\lceil k/2 \rceil + 1)}$ is shorter than than the distance from $x$ to $x^E_{\lceil k/2 \rceil}$, $x$ will be classified correctly by $k$-nearest neighbor classifier. Thus, the nonparametric extra- and intra-class differences are rewritten as follows,

$$\Delta^E = x - x^E_{\lceil k/2 \rceil}, \tag{2.16}$$

$$\Delta^I = x - x^I_{(\lceil k/2 \rceil + 1)}. \tag{2.17}$$

They performed 2 experiments on the FERET database. In both of the experiments, they presented better results than traditional LDA and PCA-based methods such as Eigenfaces, Fisherfaces, NLDA and traditional methods like NDA and Bayes.

### 2.3.3. Discussions on NNDA

As NNDA does not have to estimate distributions from samples, it does not suffer from the small sample size problem. Moreover, the number of extracted features is not limited with number of classes, because $S_B$ is of full rank. It also does not require the nonsingularity of $S_W$ as no inversion operation for $S_W$ need to be applied. However, Qiu and Wu did not give a suggestion on how to select the step size and the alpha parameter of the weighting scheme. The approach is time consuming in the training phase, due to the stepwise procedure, but once the projection matrix is calculated, it is as efficient as LDA or PCA in the recognition phase.

## 2.4. Similarity and Distance Measures

In subspace face recognition methods, it should be decided which projected training face image is closest to or most similar to the projected query face image.

The training vector that exhibits the most similarity or the closest distance, identifies the query image. Yambor has discussed several questions about several of these measures [37]. She showed that $L_2$ norm and Cosine angle measures provided same results both in the subspace and in the original dimensionality, while $L_1$ norm, Mahalanobis distance and correlation measures produced different results. Several distance and similarity measures is briefly discussed in this section.

***$L_1$* norm:** $L_1$ norm is a distance measure and is also called as city block distance. Considering $x$ and $y$ as $N$-dimensional column vectors, $L_1$ norm is defined as follows,

$$L_1(x, y) = \sum_{i=1}^{N} |x_i - y_i|. \qquad \textbf{(2.18)}$$

***$L_2$* norm:** $L_2$ norm is a distance measure and is also called as Euclidean distance. It is the sum of squared distances of two vectors. Considering $x$ and $y$ as $N$-dimensional column vectors, $L_2$ norm is defined as follows,

$$L_2(x, y) = \sum_{i=1}^{N} (x_i - y_i)^2 = (x - y)^T (x - y). \qquad \textbf{(2.19)}$$

**Cosine Angle:** 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,

$$\cos(x, y) = \frac{x}{\|x\|}\frac{y}{\|y\|}. \quad \textbf{(2.20)}$$

**Mahalanobis Distance:** Mahalanobis distance is an eigenspace distance measure, in which, for each vector dimension, the vector values and the eigenvalue of that dimension is producted and the results are summed up. The mathematical definition of mahalanobis distance is as follows,

$$Mah(X, Y) = -\sum_{i=1}^{m} X_i Y_i C_i, \quad \textbf{(2.21)}$$

where $X$ and $Y$ are $m$-dimensional vectors in eigenspace and $C_i = \frac{1}{\sqrt{\lambda_i}}$.

# 3. TWO-DIMENSIONAL GABOR FILTERS BASED FACE RECOGNITION

## 3.1. Introduction

After the pioneering work of Daugman [32], extending 1-D Gabor filters to 2-D, Gabor filters have been extensively 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. Gabor filters, whose kernels are similar to 2-D responses of visual neurons of the mammalians, have shown to offer desirable characteristics of spatial localization, spatial frequency and orientation selectivity. It has been shown that local feature processing approaches with spatial-frequency analysis are better to cope with local distortions such as illumination, expression and pose variations than both holistic and analytic approaches. Among these methods, Gabor filters give the optimized resolution in space-frequency localization and result with illumination, expression and pose invariant image features. The motivation of gabor filters usage can said to be three-fold [40]:

- **Biological motivation.** The responses of Gabor filters are similar to the 2-D receptive field profiles of mammalian cortical cells.
- **Mathematical motivation.** Gabor filters are shown to be optimal for measuring local spatial frequencies.
- **Emprical motivation.** Gabor filters have shown to be robust to distortions in other pattern recognition tasks such as texture segmentation, handwritten numeral recognition and fingerprint recognition.

In this section, 2-D Gabor filters and 2-D Gabor filters based feature representation is discussed.

### 3.2. Two-dimensional Gabor Filters

Gabor filters(kernels) are a set of filters $\psi_{\vec{k}}$, where $\vec{k} = k_{\mu,v}$; $\mu$ indicates the orientation and $v$ indicates the scale of the kernel. Each kernel is a product of a Gaussian envelope function and a complex plane wave. Gabor kernels, in image cooordinates $z = (x, y)$, are defined as follows,

$$\psi_{\mu,v}(z) = \frac{\left\| k_{\mu,v} \right\|^2}{\sigma^2} e^{\left(-\|k_{\mu,v}\|^2 \|z\|^2 / 2\sigma^2\right)} \left[ e^{izk_{\mu,v}} - e^{-\sigma^2/2} \right]. \quad \textbf{(3.1)}$$

The wave vector $k_{\mu,v}$ is defined as follows,

$$k_{\mu,v} = k_v e^{i\phi_\mu}, \quad \textbf{(3.2)}$$

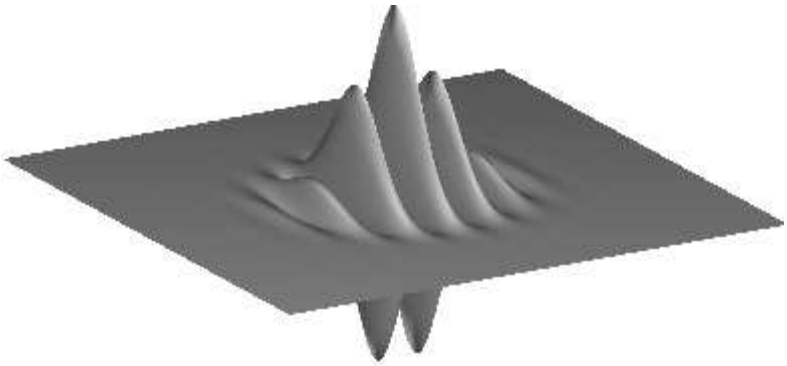
where $k_v = k_{\max} / f^v$ and $\phi_\mu = \pi\mu/8$. $k_{\max}$ is defined as maximum frequency and $f$ is the spacing factor between kernels in the frequency domain [27]. Lades et al. investigated $\sigma = 2.\pi$, $f = \sqrt{2}$ and $k_{\max} = \pi/2$ yielding with optimal results along with 5 scales, $v \in \{0,..,4\}$, and 8 orientations, $\mu \in \{0,..,7\}$, and. In [39], Shen et al. also discussed tuning the Gabor kernel parameters and after two experiments, they showed that 5 scales and 8 orientations yielded optimal recognition performance.

The first exponential term in the square brackets in Equ.3.1 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). Figure 3.1 is an example from [27], visualizing the 3-D shape of real and imaginary part of the kernel.

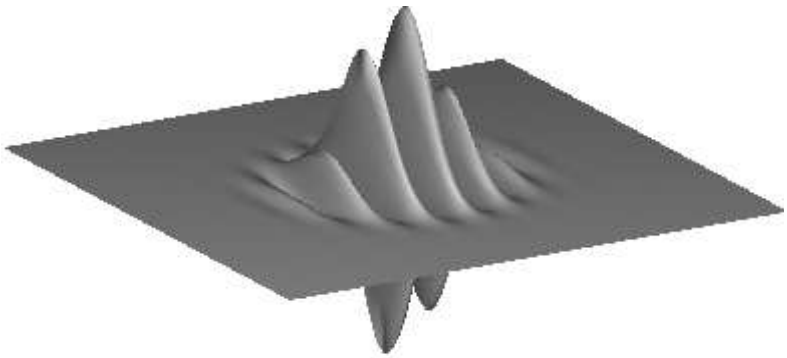
The response of the kernel in Fourier domain is defined as follows,

$$(F\psi_{k_{\mu,v}})(k_{0_{\mu,v}}) = e^{-\frac{\sigma^2.\left(k_{0_{\mu,v}}-k_{\mu,v}\right)^2}{2.\left(k_{\mu,v}\right)^2}} - e^{-\frac{\sigma^2.\left(\left(k_{0_{\mu,v}}\right)^2+\left(k_{\mu,v}\right)^2\right)}{2.\left(k_{\mu,v}\right)^2}}. \quad \textbf{(3.3)}$$

The first Gaussian determines a band-pass filter. The second exponential removes the DC component of the kernel [27]. The kernels $\psi_{k_{\mu,v}}$ are all self-similar, i.e., they can be generated from the mother wavelet, by scaling and rotating with the wave factor $k_{\mu,v}$ [33]. 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 function [27]. Figure 3.2. shows the 64x64 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 $k_{\max} = \pi/2$.

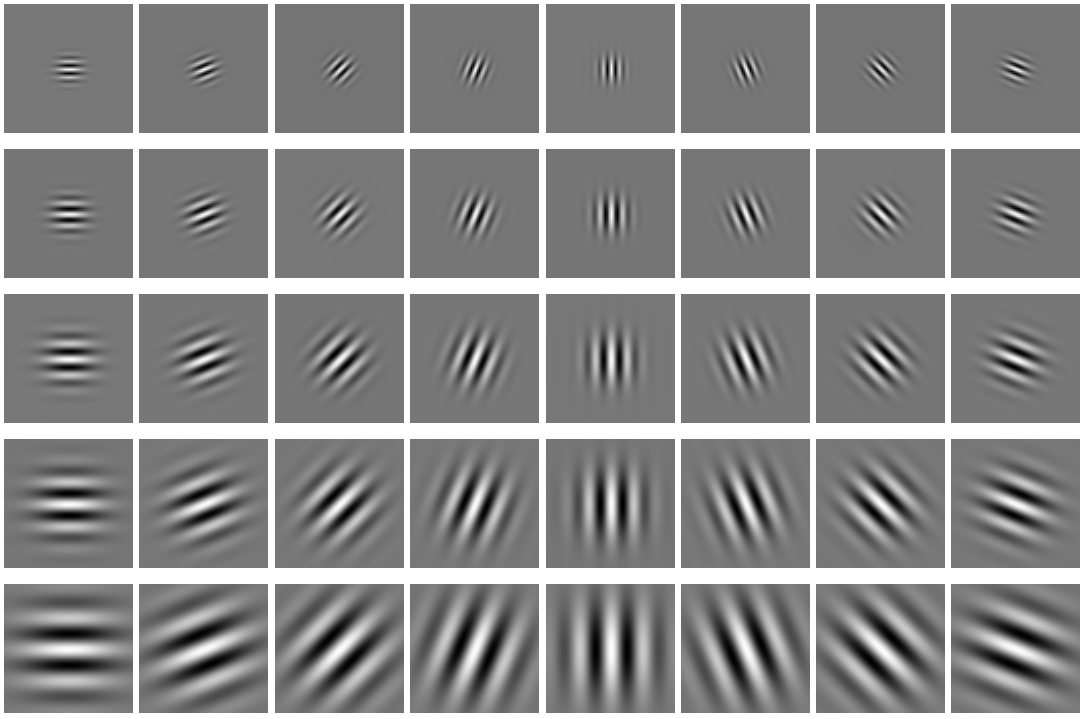


**(a)**

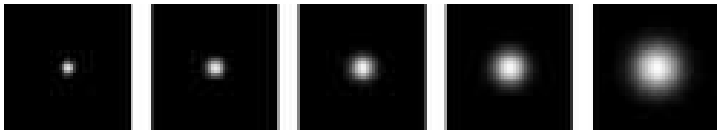


**(b)**

**Figure 3.1.** 3-D visualization of a Gabor kernel. (a) The cosine part(real part) of the kernel with $\mu = 0.72, v = 45^{\circ}$. (b) The sine part(imaginary part). The kernel has a size of 128 units in each of the first two dimensions.



**(a)**



**(b)**

**Figure 3.2.** 2-D Gabor kernels of 5 scales and 8 orientations. (a) The real part of the Gabor kernels at five different scales and eight orientations with the parameters : $\sigma = 2.\pi$ , $f = \sqrt{2}$ , $k_{\max} = \pi/2$ . (b) The magnitude of the Gabor kernels at five different scales.

### 3.3. Two-dimensional Gabor Filters Based Feature Representation

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

$$G_{\mu,v} = I(x,y) * \psi_{\mu,v}, \quad \textbf{(3.4)}$$

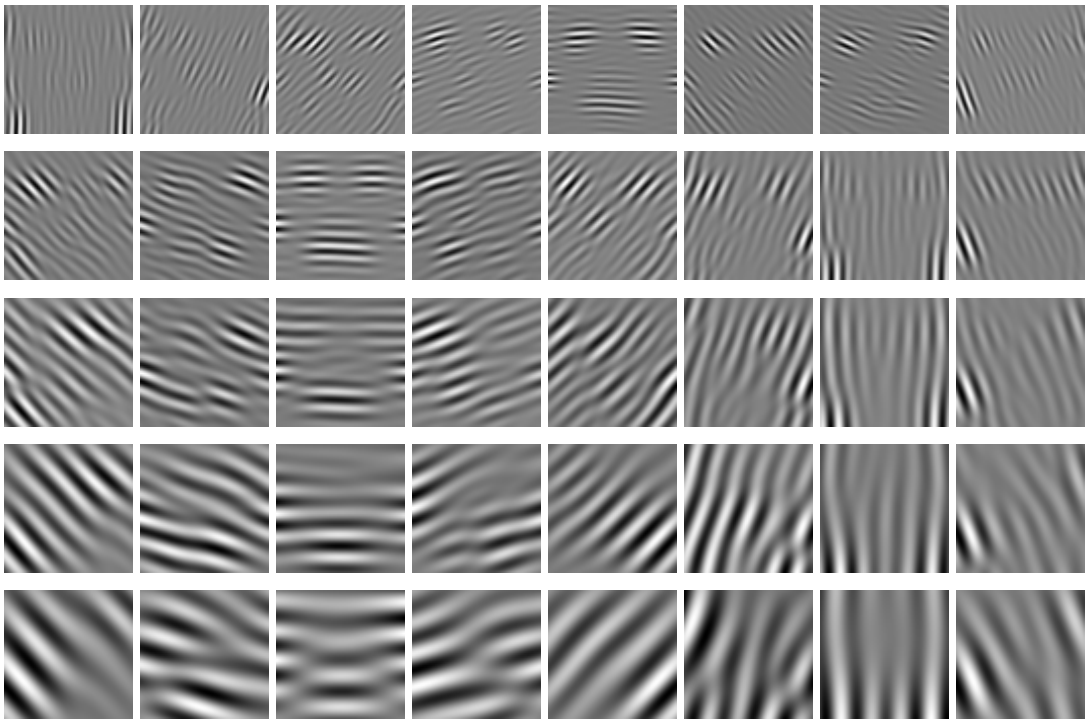
where *z = (x,y)* and * denotes the convolution operator. $G_{\mu,v}(z)$ is the convolution result corresponding to the Gabor kernel at orientation $\mu$ and scale *v*. The convolution results of Gabor kernels with five scales and eight orientations forms the set of Gabor filters representation of an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

$$S = \left\{ G_{\mu,v}(z) : \mu \in \{0,..,7\}, v \in \{0,..,4\} \right\}. \quad \textbf{(3.5)}$$

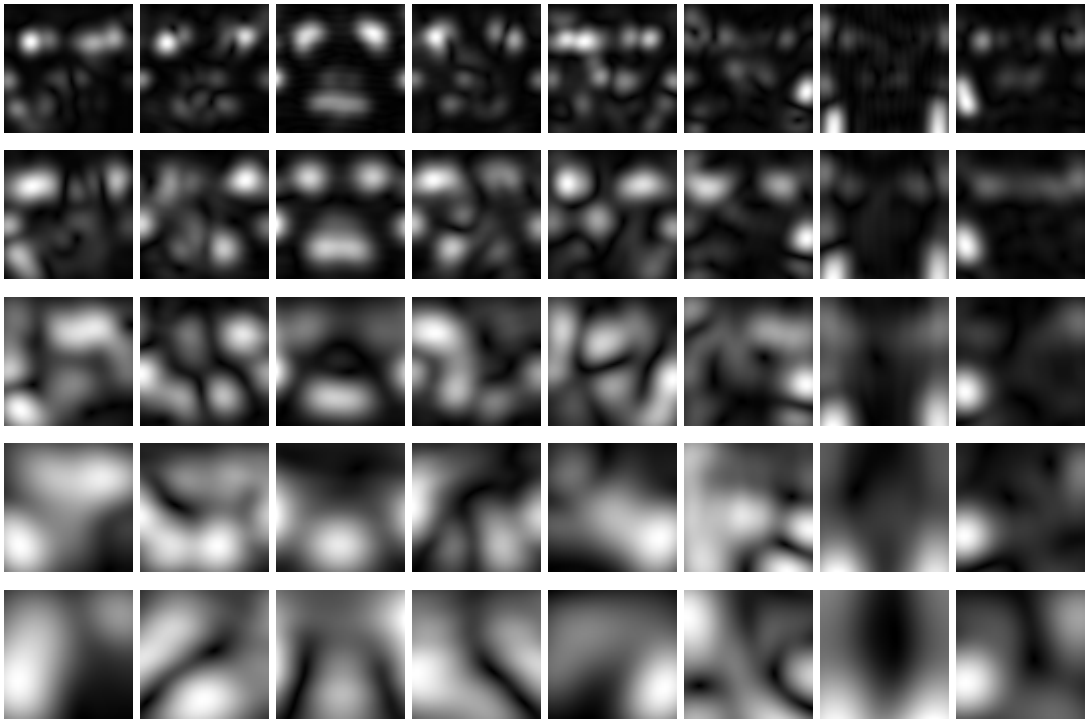
This feature set is called as Gabor jets in [26, 27]. To perform the time-consuming convolution operation efficiently in Fourier domain, this set can be derived using Fast Fourier Transform(FFT) and then taking the inverse Fourier transform back. This operation is defined as,

$$G_{\mu,v}(z) = F^{-1}\left\{ F\{I(z)\} F\left\{ \psi_{\mu,v}(z) \right\} \right\}, \quad \textbf{(3.6)}$$

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 [27], or the augmented form of the magnitude of the convolution results from full images are used as feature vectors [33]. 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 3.3 shows the Gabor filters representation(the real part and the magnitude) of a 64x64 sample image from the ORL database.



**(a)**



**(b)**

**Figure 3.3.** Gabor filters 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.

### 3.4. Previous Work on Gabor Feature Based Face Recognition

A considerable amount of research has been conducted on Gabor filters based face recognition due to the effectiveness of Gabor kernels in image processing and representation. These studies can be organized under two categories: analytic and holistic approaches [40].

#### 3.4.1. Analytic Approaches

Analytic approaches make use of the Gabor jets obtained from the pre-located feature points in the facial image. Several approaches differing in selection of these important points for Gabor jets extraction exist in the literature. Of these methods, elastic graph matching-based and non-graph matching-based methods is discussed in this section.

##### 3.4.1.1. Elastic Graph Matching Based Methods

The Dynamic Link Architecture(DLA) [27] and Elastic Bunch Graph Matching(EBGM) [26] are two well-known elastic graph matching-based methods. These methods use a two-step approach to build the representing graph $g_I$ for face image *I*. In the first step, a model graph $g_M$ is shifted within the input image keeping its rigid form undistorted. The rigid graph is initialized at an arbitrary position on the image and the position of the input graph is updated until a predefined cost function $S(g_I, g_M)$ is minimized. The next move is to deforme the individual vertices of the graph. To catch the local distortions or in depth pose variations, the vertices are visited in a random order and shifted by a random vector $\vec{v}$ within a topological constraint $\overline{C}$. Because of this deformation of vertices, the method is called *elastic* graph matching.

In DLA [27], a rectangular subgrid is placed on each face image of the gallery. Each node of this subgrid is labeled by the Gabor jet representation in the image coordinates. During the recognition, a resulting sparse graph is formed adaptively by the best matching model graph.

The edges between vertices $\vec{v}_i$ and $\vec{v}_j$ of an image or a model graph is labeled with the Euclidean distance as follows,

$$\vec{\delta}_{ij} := \vec{v}_i - \vec{v}_j, \qquad (i, j) \in E \tag{3.6}$$

where $E$ is the set of edges. The edge labels of the image graph is compared with the edge labels of the model graph with the following quadratic function,

$$S_e\left(\vec{\delta}_{ij}^{I}, \vec{\delta}_{ij}^{M}\right) := \left(\vec{\delta}_{ij}^{I} - \vec{\delta}_{ij}^{M}\right)^2. \tag{3.7}$$

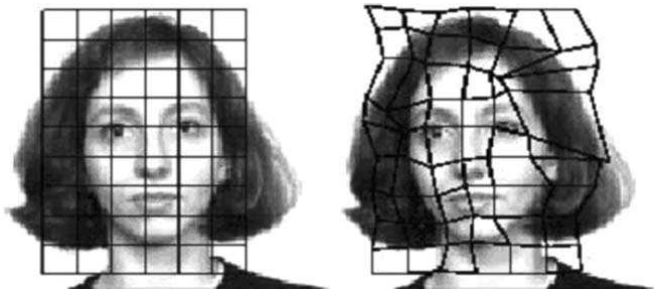
The optimal matching of an image graph with a model graph in DLA is the search of a set of vertex labels which optimizes the vertex labels and edge labels minimizing the following cost function,

$$\begin{aligned} C_{total}\left(\vec{v}_i^{I}\right) &:= \lambda C_e + C_v \\ &= \lambda \sum_{(i,j)\in E} S_e\left(\vec{\delta}_{ij}^{I}, \vec{\delta}_{ij}^{M}\right) - \sum_{i \in V} S_v\left(J^I\left(\vec{v}_i^{I}\right), J_i^M\right), \end{aligned} \tag{3.8}$$

where $J^I$ is the Gabor jet representation of the image and $v_i^I$ is the optimal set of vertex labels in the model graph, minimizing the cost function $C_{total}$. $\lambda$ is a control parameter which tries to protect the rigidity of the image graph, penalizing the distortions. Similarity function $S_v$ compares Gabor jets of image graph and with the jets of model graph using magnitude information only. The similarity function they used for jet comparison is the Cosine angle measure and it is defined as,

$$S_v\left(J^I, J^M\right) := \frac{J^I J^M}{\left\|J^I\right\|\left\|J^M\right\|}. \tag{3.9}$$

The deformation of the elastic graph in DLA, is demonstrated in Figure 3.4 [27].



(a) (b)

**Figure 3.4.** Face images represented by graphs (a) a model graph (b) the deformed graph representing a new face image.

The system presented in [26] was inspired by [27] but differed in three ways. Firstly, they included phase information to discriminate between the jets exhibiting same magnitude levels and to make accurate jet localization with displacements. Second, they used object-adapted graphs instead of the rectangular graphs of DLA to better cope with in-depth rotations. Each vertex of the object-adapted graph is a specific facial landmark and is called as fiducial point. Third, instead of one-to-one graph matching of the DLA system, they presented a bunch graph matching model which combines jets of a small set of individual faces. This allowed to find the fiducial points in one matching instead of matching each image graph with each model graph, reducing the time complexity significantly. A Gabor jet $J_j$ , computed at an image point $\vec{x}$ can be written as $J_j = a_j e^{i\phi_j}$ , with magnitude $a_j\left(\vec{x}\right)$ which varies slowly with position, and phases $\phi_j\left(\vec{x}\right)$ which rotates by a rate specified by the wave factor $\vec{k}_j$. Considering the jets $J$ and $J'$ computed on object locations with small relative displacement $\vec{d}$, the phase shifts can be approximately compensated for by the term $\vec{d}.\vec{k}_j$ [26]. Utilizing the fact, they offered the following phase sensitive jet similarity function,

$$S_\phi\left(J, J'\right) = \frac{\sum_j a_j a'_j \cos(\phi_j - \phi'_j - \vec{d}.\vec{k}_j)}{\sqrt{\sum_j a_j^2 \sum_j a'^2_j}}. \qquad \textbf{(3.10)}$$

In this scheme, $\vec{d}$ has to be estimated. They achieved this by maximizing S in its Taylor expansion around d = 0.

In the graph model they introduced, a labeled graph consists of *N* vertices connected by *E* edges. Each vertex is located at facial landmarks $\vec{v}_n$, *n* = 1,…, *N*, called fiducial points, e.g., the pupils, the corners of the mouth, the tip of the nose, the top and bottom of the ears, etc. This graph is object-adapted since its geometrical structure is adapted to the structure of the object. An example is demonstrated in Figure 3.5 [26]. Similar to the DLA approach, vertices are labeled with Gabor jets $J_n$ and edges are labeled with euclidean distance vectors $\vec{\delta}_e := \vec{v}_n - \vec{v}_{n'}$, *e* = 1,…, *E* , where edge *e* connects vertex *n* with $n'$.

Instead of representing each feature combination by a seperate graph, they proposed a combination of a set of *M* individual model graphs $g^{B_m}$ (m = 1,..,*M*) into a stack-like structure, called a face bunch graph(FBG). It has the same grid structure with the previous graph structure, but each vertex now consists of jets of the same fiducial point from *M* individual model graphs. A set of jets coresponding to one fiducial point is called a bunch. The vertices of the corresponding FBG *B* are labeled with bunches of jets $J_n^{B_m}$ and its edges are labeled with average distances $\vec{\delta}_e^{\,B} = \sum_m \vec{\delta}_e^{\,B_m} / M$ .

With a sufficient set of graphs in a FBG(approximately 70 graphs), the elastic graph matching procedure automatically generates large galleries of model graphs. Matching a FBG on a new image is performed by maximizing a graph similarity between an image graph and the FBG of identical pose. It is similar to the DLA graph similarity measure but facilitates to use the phase information for jet comparison. For an image graph $g_I$ with vertices *v* = 1 ,…, *V* and edges e = 1 ,…, *E* and a FBG *B* with model graphs m = 1, …, *M* the FBG graph similarity measure is defined as follows,

$$S_B\left(g^I, B\right) = \frac{1}{V}\sum_v \max_m \left(S_\phi\left(J_v^I, J_v^{B_m}\right)\right) - \frac{\lambda}{E}\sum_e \frac{\left(\vec{\delta}_e^{\,I} - \vec{\delta}_e^{\,B}\right)^2}{\left(\vec{\delta}_e^{\,B}\right)^2}, \qquad \textbf{(3.11)}$$

where, $\lambda$ determines relative importance of jet similarities and the rigidity term.

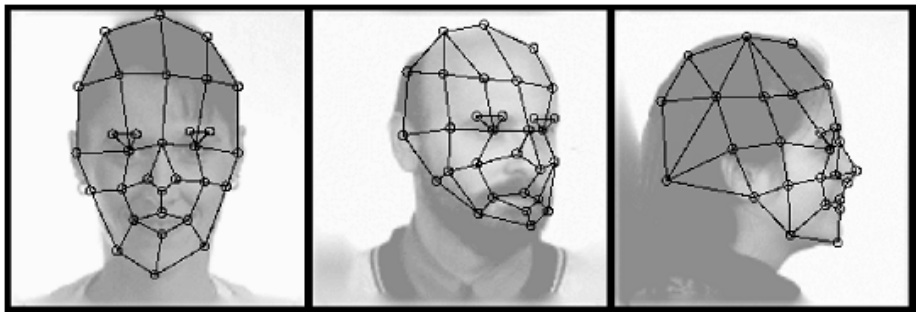


**Figure 3.5.** Object-adapted grids for different poses. The vertices are positioned automatically by elastic graph matching. In general, the matching finds the fiducial points accurately, however in the center face image, chin was not found accurately and leftmost vertex was misplaced.

In a further work [41], Wiskott investigated the role of topographical constraints in face recognition. The simple graph models using rectangular subgrids with different topographical constraints are compared with a face bunch graph system. He reported that constraints are useful when the variations in illumination, background and scale do not change. He also compared jet similarity measures with magnitude information only and with phase information included. He showed that phase information included jet similarity measures yielded better performance when the illumination did not vary significantly.

To further improve the elastic graph matching algorithm, Mu et al proposed the group shift/local deformation approach [42]. In this scheme, vertices of the graph are clustered into groups and each cluster moves together in the rigid shifting stage. In the second stage, individual vertices of groups deform to fine-tune their positions. The fine-tuning stage allowed relative shift among the vertices in the same group and hence gave flexibility to the deformation. They also showed that the local voting based nearest neighbor classifier along with the group shifting deformable graph alogorithm achieved better recognition rates than the standard elastic graph matching algorithm.

In [43], Liao and Li used a 17 node bunch graph to locate 17 facial feature points such as pupils, mouth, nose tip, ears, etc. To build the bunch graph, they used a total of 70 images with manually marked correct facial feature points. Once the FBG is built, the facial feature points are detected correctly in the new image by EGM procedure. Finally, a graph-adjusting strategy was proposed to correct the few amount of incorrect matches.

In recognition, they utilized Two-Layer Nearest Neighbor(TLNN) and Modular Nearest Feature Line(MNFL) approach and achieved better results than the standard Nearest Neighbor and Nearest Feature Line(NFL) classification methods.

#### 3.4.1.2. Non-Graph Matching Based Methods

Due to the time consuming process of EGM approach, researchers proposed different techniques for feature point location.

In [44], Escobar et al. proposed log-polar image usage for Gabor feature extraction. Firstly, 16 different fiducial points are manually located, then their coordinates are transformed to log-polar coordinates. Gabor jets are calculated on the log-polar transformed feature coordinates. Finally jets are compared using the similarity measure described in DLA [27]. The main advantage of the proposed system was the fact that face images before Gabor filtering had smaller sizes and hence speeding up the process. They presented better results than EBGM approach on the Yale database, when the training size is greater than 2.

In [45], facial feature points were detected using color and edge information, followed by corner detection using SUSAN operator. They have found 12 useful Gabor filters based on a filter arrangement theory and they used these filters both for feature localization and feature extraction.

In [46], Hjelmas applied a set of 24 Gabor filters to the facial image and magnitudes of the Gabor responses on each image point is summed up to result in the filtered image. The filtered image is multiplied by a Gaussian to focus on the center area of the face and avoid extracting features at the face contour. The Gaussian weighted image is then searched for peaks to locate the important feature points for face recognition. Then he used 40 Gabor filters to extract features from the pre-located feature points. Finally in recognition he used euclidean distance for classification.

### 3.4.2. Holistic Methods

While analytic methods compute Gabor jets on pre-located feature points and use this information as features, holistic methods extract features from the whole face. Since the feature vector consists of all useful information extracted from different frequencies, orientations and locations, this representation produces discriminant features for face recognition.

Similar to the classical holistic based approaches, images need to be normalized in size and orientation. In [47], Shan showed that this scheme is much more robust to mis-alignment in normalization procedure than gray scale intensity-based approaches.

In the holistic Gabor filters based scheme, similar to the jet computation, a set of Gabor filters are applied at each pixel location. An augmented feature vector is formed by concatenating each magnitude response from each pixel location and each image is represented by this long feature vector. The resultant high dimensional feature vectors are usually downsampled before being fed into a classifier.

In [16], Donato et al. realized this procedure and compared several methods for facial action classification. They applied 40 Gabor filters to each image and formed augmented feature vectors as described above. They applied downsampling with an 16 scale factor to reduce the feature dimension to $40.n.n/16$ , where $n \times n$ is the size of the input images. They reported cosine similarity measure as giving the best results for Gabor features classification. The results they reported was superior to PCA and local-PCA approaches, while Independent Component Analysis(ICA) was competitive with their Gabor filter results.

In [33], Liu et al. presented a similar work to Donato et al. They applied 40 Gabor kernels(5 scale and 8 orientations) at each pixel location of the image, after normalizing each image to zero-mean and unit variance. Each filter response $O_{\mu,v}(z)$ is downsampled by a factor of $\rho$ and normalized to zero-mean and unit variance. As $O_{\mu,v}^{(\rho)}$ denoting the downsampled and normalized filter output with orientation $\mu$ and scale *v*; then, the augmented Gabor feature vector $X^{(\rho)}$ is defined as follows,

$$X^{(\rho)} = \left( O_{0,0}^{(\rho)^t} \mid O_{0,1}^{(\rho)^t} \mid ... \mid O_{7,4}^{(\rho)^t} \right)^t, \qquad \textbf{(3.10)}$$

where, *t* is the transpose operator.

The resulting downsampled and normalized feature vector is said to be encompassing all the elements of the Gabor filters representation set, $S = \left\{ O_{\mu,v}(z) : \mu \in \{0,..,7\}, v \in \{0,..,4\} \right\}$, as important discriminating information. In the proposed scheme, feature dimesion was as high as 655,360 before downsampling, as 128x128 input images were used.

After downsampling with a factor of 64, dimensionality was reduced to 10,240. In the proposed work, they apply PCA to the augmented Gabor feature vector first. Then a further feature extraction is performed by Enhanced Fisher Model(EFM). They compared proposed approach with a combination of Gabor and Fisherfaces and a combination of Gabor and Eigenfaces method on a 200 subject subset of FERET database, using city block distance(L1-norm) measure. Gabor+EFM approach both outperformed Gabor+Eigenface and Gabor+Fisherface methods in recognition rate. They also evidenced that L1 distance measure and cosine similarity measure both gave better performances than mahalanobis distance and L2 distance measures.

In [48], Liu et al. applied a similar strategy to their previous work [33]. They extract Gabor features from the whole image and apply PCA to the augmented feature vector. In the next step, they apply Independent Component Analysis(ICA) to extract the Independent Gabor Features(IGF). In the classification stage, they apply Bayes MAP classification tool, achieving %98.5 recognition rate with 180 features on a 200 subject subset of FERET database and %100 recognition rate on ORL database with 88 features.

Shen and Bai proposed a kernel-space based approach with a similar Gabor filters utilization [39]. Augmented Gabor feature vectors are formed in the same way with Liu's approach, but these features are mapped to the kernel space using kernel functions and then Generalized Disriminant Analysis(GDA) is utilized in the kernel space. They showed %97.5 recognition rate on the same 200 class subset of FERET database exhibiting illumination and expression variations, selecting random 2 images for training and 1 image for testing They experimented cosine similarity measure giving the best performance in classification.

A combination of a Gabor feature based method is proposed in the next section, Gabor+NNDA.

# 4. GABOR FEATURE BASED FACE RECOGNITION USING NEAREST NEIGHBOR DISCRIMINANT ANALYSIS

## 4.1. Introduction

Inspired by the work of Liu et al. [33], Gabor features are extracted in a holistic manner. PCA is applied in the next step to further reduce the high dimensional Gabor features and NNDA [34] is applied as a final step of feature extraction, on the Gabor+PCA features. In the classification stage L2 distance measure is applied.

## 4.2. Gabor Feature Representation

A set of 40 Gabor kernels are used with the following parameters: $\sigma = 2.\pi$, $k_{\max} = \pi/2$ , and $f = \sqrt{2}$. Each input 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 input image. Thus, a total of 40 $n \times n$ magnitude image is obtained. Let the kernel with orientation $\mu$ and scale *v* be denoted by $\psi_{\mu,v}(z)$, with *z = (x,y)* denoting an image coordinate. The convolution of the image *I(z)* with $\psi_{\mu,v}(z)$ is defined as,

$$G_{\mu,v}(z) = I(z) * \psi_{\mu,v}(z). \quad \textbf{(4.1)}$$

Therefore the set $S = \{G_{\mu,v}(z) : \mu \in \{0,..,7\}, v \in \{0,..,4\}\}$ forms the Gabor filters representation of the image *I(z)*.

To benefit from different spatial frequencies(scales), spatial localities and orientation selectivities, these representation results are concatenated and an augmented Gabor feature vector *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 vector is defined as follows,

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

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

### 4.3. Dimensionality Reduction and Discriminant Analysis of Gabor Features with PCA and LDA

The augmented Gabor feature vector defined in Equ.4.2 resides in a high dimensionality of $\mathbb{R}^{\mathbb{N}}$, where $\mathbb{N}$ is the dimensionality of vector space. In a typical application, $\mathbb{N}$ 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” [49]. PCA is the optimal dimensionality reduction technique in the sense of mean-square error.

Let

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

where $\varepsilon(.)$ is the expectation operator.

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

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

where $\Phi \in \mathbb{R}^{D\times D}$ is an orthogonal eigenvector matrix and $\Lambda \in \mathbb{R}^{D\times D}$ a diagonal eigenvalue matrix with diagonal elements in decreasing order $(\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_D)$. 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 is defined as,

$$y^{(\rho)} = T^t x^{(\rho)}, \quad \textbf{(4.5)}$$

where, $T = [\phi_1\phi_2...\phi_d], d < D$ and $T \in \mathbb{R}^{d\times D}$.

The lower dimensional feature vector $y^{(\rho)} \in \mathbb{R}^d$ 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(FLD) approach.

FLD 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, FLD approach discussed in section 2.2 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 $X^{(\rho)}$.

### 4.4. Discriminant Analysis of Gabor Features with NNDA

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

As NNDA approach is detailed in section 2.3, only the training algorithm of Gabor+NNDA approach is presented in this section. Training algorithm is given in Figure 4.1.

1. Given *D* dimensional samples $\{x_1 \mid x_2 \mid ... \mid x_N\}$, *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 $G_{i,\mu,\nu}(z); \mu \in \{0,..,7\}, \nu \in \{0,..,4\}, i \in \{1,..,N\}$.
4. Downsample each filter output $G_{i,\mu,\nu}(z)$ with a factor of $\rho$ to achieve $G^{(\rho)}_{i,\mu,\nu}(z)$, and normalize the final $G^{(\rho)}_{i,\mu,\nu}(z)$ to zero-mean and unit variance.
5. Concatenate rows(or columns) of each resultant $G^{(\rho)}_{i,\mu,\nu}(z)$ to form an augmented feature vector $x^{(\rho)}_i$. $x^{(\rho)}_i = \left\{G^{(\rho)^t}_{i,0,0} \mid G^{(\rho)^t}_{i,0,1} \mid ... \mid G^{(\rho)^t}_{i,7,4}\right\}^t$
6. Form the final Gabor feature matrix $X^{(\rho)}$ by assembling each $x^{(\rho)}_i$ in columns, side by side. $X^{(\rho)} = \left\{x^{(\rho)}_1 \mid x^{(\rho)}_2 \mid ... \mid x^{(\rho)}_N\right\}$
7. Apply PCA on Gabor feature matrix $X^{(\rho)}$ to learn the PCA projection matrix $T_{pca}$. $T_{pca} = [\phi_1 \phi_2 ... \phi_{D-1}], T \in \mathbb{R}^{D-1 \times D}$
8. Project feature matrix $X^{(\rho)}$ with the learned PCA model. $Y_{pca} = T^t_{pca} X^{(\rho)}$
9. Apply NNDA on $Y_{pca}$ to learn the NNDA projection matrix. $T_{nnda} = [\varphi_1 \varphi_2 ... \varphi_d], T \in \mathbb{R}^{d \times D-1}$
10. Project Gabor+PCA features $Y_{pca}$ with the learned NNDA model. $Y = T^t_{nnda} Y_{pca}$

**Figure 4.1.** Gabor+NNDA training algorithm

In classification, a new test image $x'$ is projected to the Gabor+NNDA feature space with the linear projection matrix $T_{nnda}$ as follows,

$$y' = T^t_{nnda} x'. \quad \textbf{(4.6)}$$

Then, a simple distance measure is applied to identify $y'$ with the label of the closest feature vector in Gabor+NNDA space, using L1 distance measure or L2 distance measure defined as,

$$\delta_{L1}(y', y) = \sum_{i=1}^{m} | y_i' - y_i | \quad , \quad \delta_{L2}\left(y', y\right) = \left(y' - y\right)^t \left(y' - y\right). \qquad \textbf{(4.7)}$$

Both of the performances of L1 and L2 distance measures will be investigated in the next section. The feasibility of Gabor+NNDA will be shown by comparing it by NNDA on Yale database [4]. The efficiency of the proposed method will also be presented by comparing it with Gabor+Eigenfaces and Gabor+Fisherfaces methods on a 200 subject subset of FERET database.

## 5. EXPERIMENTS AND RESULTS

Experiments in this section is performed on 2 different databases. First set of experiments is performed on Yale database [4], and the next set of experiments is performed on a subset of FERET Database [35] which consists of 600 images of 200 subjects.

For Gabor filters implementation, a public Gabor filters API utilizing Intel OpenCV library is used [50]. NNDA, PCA and LDA are implemented on Matlab 7.0.1 R14 platform. All the implementation and tests are performed on an Intel Pentium 4 CPU 3.20 GHz , 1.0 GB RAM PC.

### 5.1. Experiments and Results on Yale Database

#### 5.1.1. Description of the Yale Database

Yale database [4] is a popular face database that was prepared in Yale University CS department. The database contains 165 GIF images of 15 subjects. There are 11 images per subject, one for each of the following facial expressions or configurations: center-light, w/glasses, happy, left-light, w/no glasses, normal, right-light, sad, sleepy, surprised, and wink. 8 Images of 2 subjects is shown in Figure 5.1. Original Yale images are of size 320x243. The 64x64 scaled images prepared by Deng Cai [51] were used in the experiments.

**Figure 5.1.** Eight of a total of 11 images of 2 subjects from Yale Database. Images exhibit both expression, lighting variations and occlusion(glasses).

### 5.1.2. Experiments

***Experiment#1*** In the first experiment, the aim is to show that the proposed Gabor+NNDA method outperforms its predecessor, NNDA. 5 random partitions are formed by randomly selecting 5 images for training and the remaining 6 images for testing in each trial. Average results of the 5 experiments is given.

In this experiment, the *step size* of NNDA is set to 5. The reduced subspace is of 14 dimensions. Note that the original dimensionality is 64x64 = 4,096. While step size is constant, (*k, alpha*) tuple is changed as follows,

$(k, alpha) : k \in \{1,3,5\}, alpha \in \{0,1,2,3\}$ . The result is shown in Figure 5.2.

It can be clearly observed that Gabor+NNDA outperformed NNDA on all (*k, alpha*) tuples . Gabor+NNDA reaches 90 percent accuracy in 14 feature dimension, along with parameters *step size* = 5, *alpha* = 1, *k* = 5 .

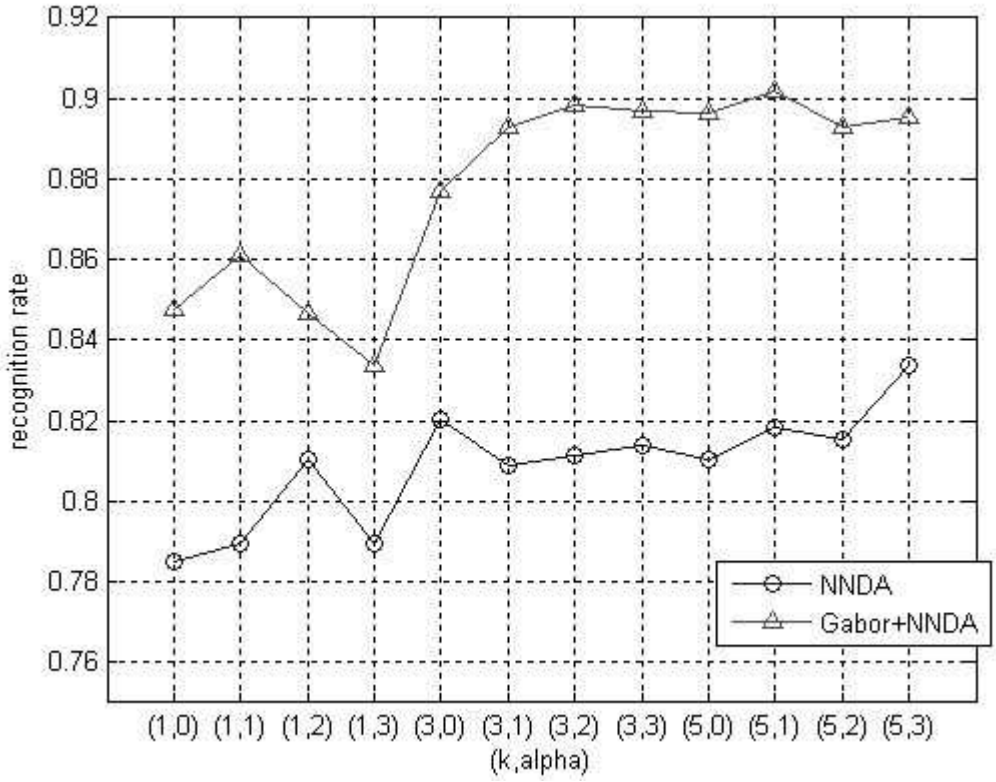



**Figure 5.2.** Relative performance of Gabor+NNDA and NNDA on Yale database. Reduced feature dimension is 14. Step size is set to 5.

## 5.2. Experiments and Results on FERET Database

### 5.2.1. Description of the FERET Database

The FERET database was collected in 15 sessions between August 1993 and July 1996. The database contains 1564 sets of images for a total of 14,126 images that includes 1199 individuals and 365 duplicate sets of images. 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 imposed on facial expression and with at least two frontal images shot at different times during the same photo session.

Three sets of experiments is performed on a subset of FERET database. The experiments involve 600 face images of 200 subjects such that each subject has three images of size 256x384 with 256 gray scale levels. First, the centers of the eyes are manually detected, then rotation and scaling transformations align the centers of the eyes to predefined locations. Finally, face image is cropped to the size of 128x128 to extract the facial region, which is further normalized to zero-mean and unit variance. Figure 5.3. shows some example images that are used in the experiments.

**Figure 5.3.** Example images used in the FERET experiments. Images both show different lighting conditions and facial expressions. In the figure, the top two rows show the examples of training images used in the experiments, and the bottom row shows the example of test images.

### 5.2.2. Experiments

***Experiment#1*** In this experiment, performances of L1 distance measure and L2 distance measure after Gabor+NNDA feature extraction is investigated. To compare the distance measures, *step size* is set to 13, *alpha* and *k* are set to 1. In Figure 5.4., recognition rates with L1 and L2 distances on the same data is shown. The results show that L2 distance measure gives better performance than L1 distance measure, on the classification of Gabor features. The average recognition rates on the Figure 5.4., with L1 and L2 distance measure is tabulated in Table 5.1.

**Table 5.1.** The average recognition rates and standar deviations of L1 and L2 distance measures on Gabor+NNDA features, using a 200 subject subset of FERET Database.

| | Mean | Standard deviation |
|---|---|---|
| L1 Distance Measure | 0.8843 | 0.1827 |
| L2 Distance Measure | 0.9021 | 0.1813 |

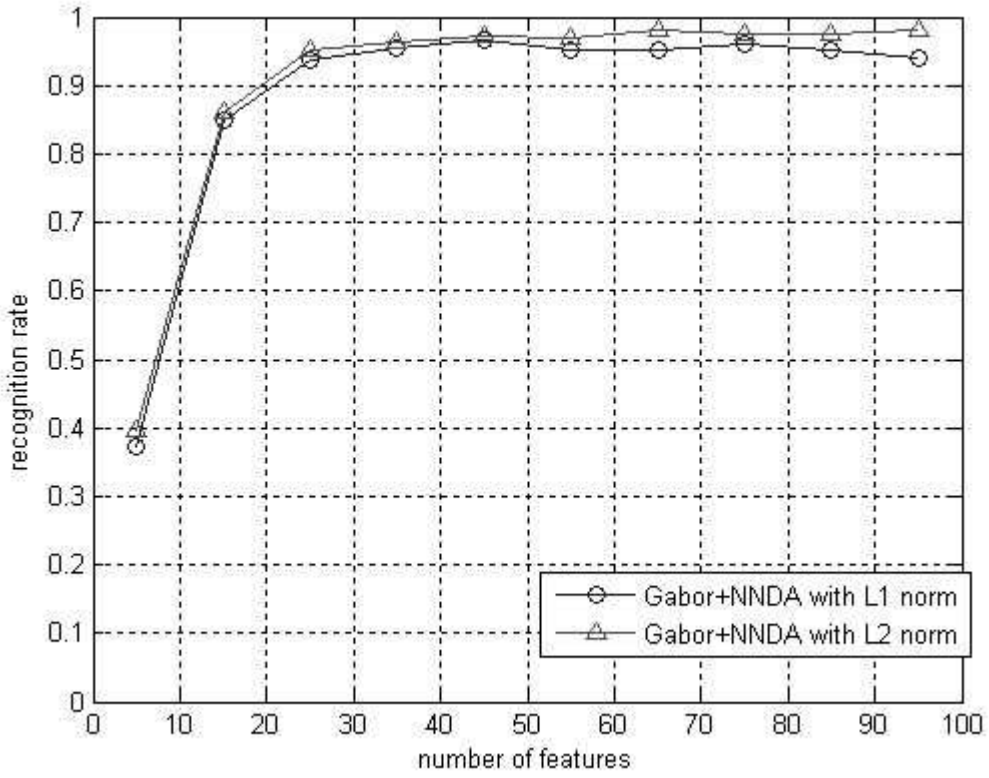



**Figure 5.4.** Performance comparison of L1 and L2 distance measures on Gabor+NNDA features. L2 distance measure performs better than L1, with a slight increase of 2 percent in average.

***Experiment#2*** The next experiment is the performance comparison of the two standard methods Gabor+Eigenfaces and Gabor+Fisherfaces, with the proposed method Gabor+NNDA. L2 distance measure is used for recognition. Same configurations with the previous experiment, *step size* = 13, *alpha* = 1, *k* = 1, is applied. Figure 5.5. shows the results. Gabor+NNDA achieves the highest recognition rate with 98 percent in 65 feature dimensions. Gabor+Fisherfaces achieves a 92.6 percent accuracy and Gabor+Eigenfaces achieves 40.6 in the same feature dimension. The average recognition rates of the figure is tabulated in Table 5.2.

**Table 5.2.** Average recognition rates of 2 standard methods Gabor+Eigenfaces and Gabor+Fisherfaces, and the proposed Gabor+NNDA, using a 200 class subset of FERET Database. The maximum number of feature dimension is 95.

| | Mean | Standard Deviation |
|---|---|---|
| Gabor+Eigenfaces | 0.3051 | 0.1610 |
| Gabor+Fisherfaces | 0.8270 | 0.2194 |
| Gabor+NNDA | 0.9021 | 0.1813 |

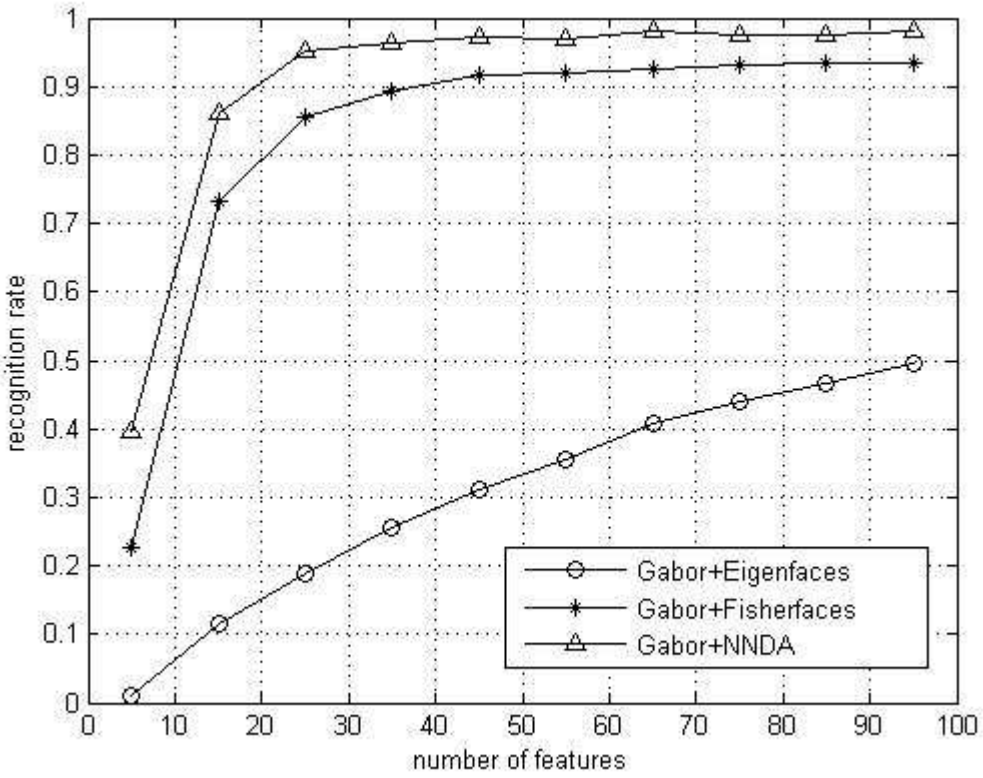



**Figure 5.5.** Comparative face recognition performance of Gabor+Eigenface, Gabor+Fisherface and Gabor+NNDA, on the augmented Gabor feature vector $X^{(\rho)}$ downsampled by a factor of 64, i.e., $\rho = 64$.

***Experiment#3*** The final experiment is aimed to judge the effect of the change of the *alpha* and *step size* parameters on the recognition performance. Due to the fact that no suggestions are given on the selection of *step size* and *alpha* in the original NNDA work, a performance comparison experiment is performed to investigate their effects on recognition. Firstly, the effect of *alpha* is investigated while *step size* is set to 5 and *k* is set to 1. Recognition rates while *alpha* is changing from 0 to 6 is given in Figure 5.6. It can generally be inferred that recognition rate is higher with small *alpha* values. In feature dimensions greater than 75, it is observed that the rate is increased while *alpha* is changing from 4 to 6.

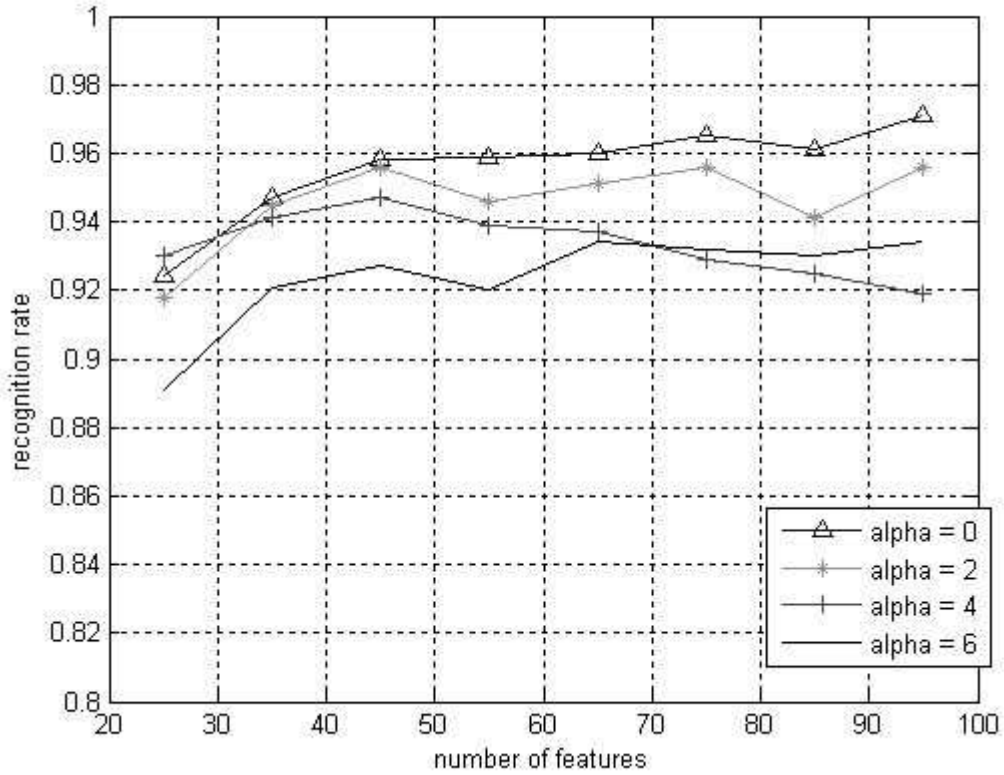



**Figure 5.6.** The effect of *alpha* parameter on the recognition performance of Gabor+NNDA features. It is observed that recognition rate is higher when *alpha* is small.

The next step is to set alpha to a constant value and increase the step size to investigate the effect of step size on recognition performance. Figure 5.7. shows the relation of step size with recognition performance. *Alpha* and *k* is set to 1 in the experiment. It can be clearly seen, when the feature dimension is less than 75, that *step size* is proportional with the recognition rate. However, no evidence about about the upper limit of this relation is obvious.

Only step sizes of 7, 9, 11 and 13 is experimented as the training time increases significantly with the increasing step size, i.e., training with the Gabor features in step sizes greater than 15 required more than 2 minutes on our Intel Pentium 4 CPU 3.20 GHz platform.

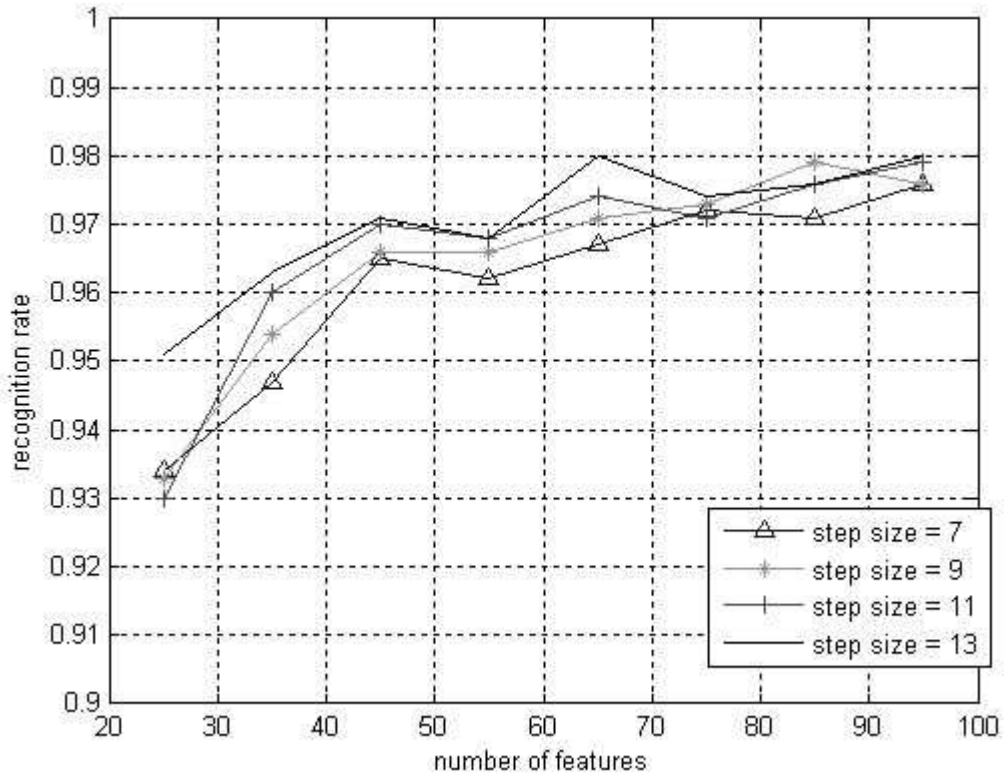



**Figure 5.7.** The effect of *step size* parameter on the recognition performance of Gabor+NNDA features. It is observed that recognition rate increases proportional to the *step size.*

## 6. CONCLUSIONS AND FUTURE WORK

In this thesis, a review on popular Gabor feature based face recognition algorithms is given and a new feature extraction method, Gabor+NNDA, is proposed. 2-D Gabor filters have been used in both image processing and computer vision, after the pioneering work of Daugman [32], extending 1-D Gabor filters to 2-D. 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 [34]. Moreover, it does not suffer from the singularity of the within-class scatter matrix of the data.

Gabor+NNDA extracts important discriminant features both utilizing the power of Gabor filters and NNDA. The efficiency of the approach is shown both as relative and absolute performance indices. It achieved a 98 percent classification accuracy with 65 features, outperforming both standard methods such as Gabor+Fisherfaces and Gabor+Eigenfaces [33], using a 200 class subset of FERET database exhibiting illumination and expression variations. Also, the effects of training parameters of NNDA method such as *alpha* and *step size*, on Gabor+NNDA features classification is investigated. It is concluded that the recognition rate is high for small *alpha* values. This result is consistent with the theoretical findings. For greater alpha values, the denominator of the Equ.2.13 grows faster than the nominator, and hence, the weighting parameter which emphasizes the samples in class neighborhoods more than the samples in class centers, gets smaller. It is also concluded that *step size*, the parameter that controls the stepwise dimensionality reduction process, needs to be high for better performances.

But as the time complexity increases directly proportional with *step size*, it may not be efficient to train the system with *step sizes* greater than 15 as it required about 2 minutes to train the system with 15 step size in our configuration. It should be noted that no suggestion is given in the original NNDA work [34], on how to select parameters *alpha*, *k* and *step size*. An optimization scheme such as Evolutionary Computing can be suggested to provide the optimal parameter selection for training. But this can also increase the time complexity of the system. Due to the effectiveness of kernel approaches in Gabor feature based face recognition [39], NNDA can be extended to a kernel approach and Gabor features then can be utilized in kernel NNDA space. Moreover, instead of applying simple distance measures like L1 and L2 norm in classification of Gabor+NNDA features, more sophisticated classification schemes such as Support Vector Machines or Neural Networks can be applied.

## REFERENCES


[1] **Jain, A. K. and Li**, **S. Z.,** 2005. *Handbook of Face Recognition*, Springer-Verlag New York, Inc., Secaucus, NJ.

[2] **Zhao, W., Chellappa, R., Phillips, P. J., Rosenfeld, A.,** 2003. Face Recognition: A Literature Survey, *ACM Computing Surveys*, v.35 n.4, pp.399-458.

[3] **Turk, M. And Pentland, A.,** 1991. Eigenfaces for Recognition, *Journal of Cognitive Neurosicence*, Vol. 3, No. 1, pp.71-86.

[4] **Belhumeur, P. N., Hespanha, J., Kriegman, D.,** 1997. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection, *IEEE Trans. PAMI, Special Issue on Face Recognition*, 19(7), pp.711—20.

[5] **Moses, Y., Adini, Y., Ullman, S.,** 1994. Face Recognition: The Problem of Compensating for Changes in Illumination Direction, *ECCV(1)*, pp.286-296.

[6] **Gonzalez, R. C., Woods, R. E.,** 2001. Digital Image Processing, Addison-Wesley Longman Publishing Co., Inc., Boston, MA.

[7] **Adini, Y., Moses, Y., Ullman, S.,** 1997. Face recognition: The problem of compensating for changes in illumination direction, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp.721-732.

[8] **Gao, Y., Leung, M. K. H.,** 2002. Face Recognition Using Line Edge Map, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, v.24, n.6, pp.764-779.

[9] **Georghiades, A.S., Kriegman, D.J., Belhumeur, P.N.,** 1998. Illumination cones for recognition under variable lighting: Faces, *In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.52.

[10] **Basri, R., Jacobs, D. W.,** 2003. Lambertian Reflectance and Linear Subspaces, *IEEE Trans. PAMI*, vol 25, no. 2, pp.218-233.

[11] **Li**, **S. Z., Chu, R., Liao, S., Zhang, L.,** 2007. Illumination Invariant Face Recognition Using Near-Infrared Images, *IEEE Trans. PAMI*, vol. 29, no. 4, pp.627-639.

[12] **Okada, K., von der Malsburg, C.,** 2002. Pose-Invariant Face Recognition with Parametric Linear Subspaces, *In Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition(FGR'02)*, 20-21 May 2002, Washington D.C., USA, pp.71-76

[13] **Gross, R., Matthews, I., Baker, S.,** 2002. Eigen Light-Fields and Face Recognition Accross Pose, *In Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition(FGR'02)*, 20-21 May 2002, Washington D.C., USA, pp.3-9.

[14] **Gokberk, B., Akarun, L., Alpaydın, E.,** 2002. Feature Selection for Pose Invariant Face Recognition, *In Proceedings of the 16th International Conference on Pattern Recognition*, Quebec City, Canada, pp.306-309.

[15] **Yue, Z., Zhao, W., Chellappa, R.,** 2005. Pose-Encoded Spherical Harmonics for Robust Face Recognition Using a Single Image, *In Proceedings of the second International Workshop of Analysis and Modeling of Faces and Gestures*, October 16, Beijing, China, pp.229-243.

[16] **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.

[17] **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.

[18] **Liu, Y., Schmidt K. L., Cohn, J. F., Weaver, R. L.,** 2002. Facial Asymmetry Quantification for Expression Invariant Human Identification, *In Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition(FGR'02)*, 20-21 May 2002, Washington D.C., USA, pp.208.

[19] **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.

[20] **Martinez, A.,** 2002. Recognizing Imprecisely Localized, Partially Occluded and Expression Variant Faces from a Single Sample per Class. *IEEE Trans. PAMI*, vol. 24, no. 6, pp.748-763.

[21] **Kurita, T., Pic, M., Takahashi, T.,** 2003. Recognition and Detection of Occluded Faces by a Neural Network Classifier with Recursive Data Reconstruction, *In IEEE Conference on Advanced Video and Signal Based Surveillance*, 21-22 July 2003, Miami, FL, USA, pp.53.

[22] **Sahbi, H., Boujemaa, N.,** 2002. Robust Face Recognition Using Dynamic Space Warping, *Biometric Authentication, LNCS 2359*, Springer-Verlag, Berlin, Heidelberg, pp.121-132.

[23] **Lanitis, A., Taylor, C. J., Cootes, T. F.,** 2002. Towards Automatic Simulation of Aging Effects on Face Images, *IEEE Trans. on PAMI*, vol. 24, no.4, pp.442-455.

[24] **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.

[25] **Nefian, A. V., Hayes III, M. H.,** 1998. Face Detection and Recognition Using Hidden Markov Models, *In Proceedings of the 1998 IEEE International Conference on Image Processing*, Chicago, Illinois, USA, October 4-7, v.1, pp.141-145.

[26] **Wiskott, L., Fellous, J., Krüger, N., von der Malsburg, C.,** 1997. Face Recognition by Elastic Bunch Graph Matching, *IEEE Trans. on PAMI*, vol. 19, no. 7, pp.775-779.

[27] **Lades, M., Vorbruggen, J. C., Buhmann, J., Lange, J., von der Malsburg, C., Wurtz, R. P., Konen, W.,** 1993. Distortion Invariant Object Recognition in the Dynamic Link Architecture, *IEEE Transactions on Computers*, vol. 42, no. 3, pp.300-311.

[28] **Pentland, A., Moghaddam, B., Starner, T.** 1994. View-Based and Modular Eigenspaces for Face Recognition, *In Proc. Of IEEE Conf. On Computer Vision and Pattern Recognition*, June 1994, Seattle, WA, USA.

[29] **Penev, P. S., Atick, J. J.,** 1996. Local Feature Analysis: A General Statistical Theory for Object Representation, *Network: Computation in Neural Systems*, 7:3, pp.477-500.

[30] **Lanitis, A., Taylor, C. J., Cootes, T. F.,** 1995. Automatic Face Identification System Using Flexible Appearance Models, *Image Vis. Comput. 13*, pp.393-401.

[31] **Huang, J., Heisele, B., Blanz, V.,** 2003. Component-Based Face Recognition with 3D Morphable Models, *In Proc. of International Conf. On Audio- and Video-Based Person Authentication*.

[32] **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.

[33] **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.

[34] **Qiu, X., Wu, L.,** 2006. Nearest Neighbor Discriminant Analysis, *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 20, no. 8, 1245-1260.

[35] **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.

[36] **Kirby, M., Sirovich, L.,** 1990. Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces, *IEEE Trans. Pattern Anal. Mach. Intel.*, vol. 12, no. 1, 103-108.

[37] **Yambor, W. S.,** 2000. Analysis of PCA-Based and Fisher Discriminant-Based Image Recognition Algorithms, *MSc. Thesis*, Colorado State University, Fort Collins, CO, USA.

[38] **Martinez, A. M., Kak, A. C.,** 2001. PCA versus LDA., *IEEE Trans. Pattern Anal. Mach. Intel.*, vol. 23, no. 2, 228-233.

[39] **Shen, L., Bai, L., Fairhurst, M. C.,** 2007. Gabor wavelets and General Discriminant Analysis for Face Identification and Verification, *Image Vision Comput. 25(5)*, 553-563.

[40] **Shen, L., Bai, L.,** 2006. A review on Gabor wavelets for face recognition, *Pattern Anal. Appl.*, vol. 9, no. 2-3, 273-292.

[41] **Wiskott, L.,** 1999. The role of topographical constraints in face recognition, *Pattern Recognition Letters*, vol. 20, no. 1, 89-96.

[42] **Mu, X., Hassoun, M. H., Watta, P.,** 2003. Combining Gabor features: summing vs. voting in human face recognition, *IEEE International Conf. On Systems, Man and Sybernetics*, 5-8 Oct. 2003, vol. 1, 737-743.

[43] **Liao, R., Li, S. Z.,** 2000. Face Recognition Based on Multiple Facial Features, *In Proc. of the Fourth IEEE International Conf. on Automatic Face and Gesture Recognition*, p.209.

[44] **Escobar, M. J., Ruiz-del-Solar, J.,** 2002. Biologically-based Face Recognition using Gabor Filters and Log-Polar Images, *In International joint conference on neural networks-IJCNN*, Honolulu, USA, 1143-1147.

[45] **Wu, H., Yoshida, Y., Shioyama, T.,** 2002. Optimal Gabor Filters for High Speed Face Indetification, *In Proc. Of Internation Conf. on Pattern Recognition*, 107-110.

[46] **Hjelmas, E.,** 2000. Feature-Based Face Recognition, *In Proceedings of Norwegian Image Processing and Pattern Recog. Conference*.

[47] **Shan, S., Gao, W., Chang, Y., Cao, B., Yang, P.,** Review the Strength of Gabor Features for Face Recognition from the Angle of Its Robustness to Mis-Alignment, *ICPR(1)*, 338-341.

[48] **Liu, C., Weschler, H.,** 2003. Independent Component Analysis of Gabor features for Face Recognition, *IEEE Trans. Neural Networks*, vol. 14, no. 4, 919-928.

[49] **Edelman, S.,** 1999. Representation and Recognition in Vision, *MIT Pres*, Cambridge, MA, USA.

[50] **Zhou, M.** http://www.personal.rdg.ac.uk/~sir02mz/

[51] **Cai, Deng.** http://www.cs.uiuc.edu/homes/dengcai2/Data/FaceData.html

## BIOGRAPHY

Kadir KIRTAÇ was born in Istanbul, Turkey, in 1983. He completed his high school education in İstanbul Kabataş Erkek Lisesi(Anatolian High School), in 2001. He received his B.Sc. degree in computer engineering from Gebze Institute of Technology, in 2006.