2005

Unconstrained face recognition for law enforcement applications

Richa Singh
West Virginia University

Unconstrained Face Recognition for Law Enforcement Applications

by

Richa Singh

Thesis submitted to the
College of Engineering and Mineral Resources
at West Virginia University
in partial fulfillment of the requirements
for the degree of

Master of Science
in
Computer Science

Afzel Noore, PhD, Chair
George E. Trapp, PhD
Raymond Morehead, MD

Lane Department of Computer Science and Electrical Engineering

Morgantown, West Virginia
2005

Keywords: Biometrics, Face Recognition, Gabor Wavelet, Amplitude and Phase Features,
Scanned Faces, Disguise

Abstract

Unconstrained Face Recognition for Law Enforcement Applications

by

Richa Singh
Master of Science in Computer Science

West Virginia University

Afzel Noore, PhD, Chair

This research work simulates the human visual cortex by using 2D log polar Gabor wavelet to extract the facial features. Scale and orientation independent convolution of face image with Gabor wavelet gives the features in the form of amplitude and phase. The proposed face recognition algorithm is invariant to frequency, scale, filter orientation, illumination, and contrast. We evaluated the recognition algorithm on four face databases namely FERET, CMU AMP, CMU PIE and Notre Dame Face databases. Experimental results show that using single image for training, phase feature based face recognition performs approximately 5% better than amplitude feature based face recognition.

Another facet of this research involves matching scanned and digital face images. Normalization and transformation algorithms are proposed to resample the scanned and the digital images into one common domain. Validation is performed on a face database of 500 classes containing both the scanned and digital face images.

Finally, a synthetic face database is prepared to evaluate the performance of the proposed face recognition algorithm with disguise. The database includes synthetic face images with single and multiple variations in appearance and feature. Results show that the proposed algorithm outperforms other recognition algorithms.

Acknowledgments

I would like to express my sincere gratitude and appreciation to all those who gave me the opportunity to complete this research work. First and foremost, I would like to profoundly thank my advisor, Dr. Afzel Noore, for providing a motivation, enthusiasm and critical atmosphere during many discussions we had. It was a great pleasure for me to conduct this thesis under his supervision. His enthusiasm and integral view on research has made a deep impression on me. Dr. Noore's instinct for "what constitutes good research" has helped me to shape my approach towards the research. I am really glad that I have learnt from his knowledge and creative experience.

I would like to thank my committee members Dr. G. E. Trapp and Dr. R. Morehead for all the help and support I have received from them. I am greatful to Dr. E. Eschen, Dr. A. Ross and Dr. X. Li who helped me to understand different concepts and theories used in this research. I am greatly indebted to Dr. S. K. Singh, Dr. K. K. Shukla, Dr. R. B. Lokesh, Dr. P. Gupta, and Dr. P. Mitra for helping me in my early days and for providing useful guidance. The financial support by United States National Institute of Justice is gratefully acknowledged. I also thank CVRL University of Notre Dame, Robotics Institute CMU, NIST, CMU AMP Research Lab, Equinox Corporation, and Mr. A. Bachchan for providing the face databases used in this research work.

I would like to take this opportunity to thank Jidnya, Samir, Sushil, and Swetha for helping me in my research. My special thanks to my friends Vineeta, Ratnesh, Manoj, Lalit, and Vivek for their love and support. I wish to thank my friend Mayank for sharing the innovative research ideas, helping me in this research and for always being there.

Finally, I owe my loving thanks to my parent and sisters. Without their loving support, encouragement, understanding and patience it would have been impossible for me to finish this work. I dedicate this thesis to my family.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The modern information age gives rise to various challenges that did not exist to the same extent in earlier times. Two such challenges are the organization of society and its security. The ever increasing human population and its mobility in all its facets caused an increasing demand for enhanced ways of transferring data and sharing information, which in turn, through complex communication structures, gained its own mobility - a mobility that even outgrew its human counterpart. In this context of increased mobility and world population, security and organization have become important social issues. As mobility applies to both humans and information, security includes both security of individuals and their valuables, and the integrity of data under external influence. Within this environment of increased importance of security and organization, identification and authentication methods have developed into a key technology in various areas: entrance control in buildings, access control for automatic teller machines, or in the prominent field of criminal investigation. These few examples illustrate the necessity of identification and authentication for the functioning of modern society.

One of the most widely used approaches for providing security is face recognition. For a human being, face recognition is an easy task because we learn from the surroundings and a supervised learning is used to train our brain. This training and learning process is so remarkable that we can recognize an individual with passage of time, changes in appearance and partial occlusion. Because of the near perfect positive identification performance of face recognition in human, considerable attention has been directed towards the replication of this ability at electronic level. Automatic face recognition requires the capability to identify an individual despite many variations in the appearance of face. There are billions of different faces in the world and since human face is not a unique rigid object, each of them can have variety of deformations. The variations may be due to inter-personal or intra-personal transformations. Inter-personal variations can be due to the internal deformations of an individual such as race or genetics while intra-personal variations can be due to expression, aging, hair, cosmetics, and facial accessories. Because of these variations, face recognition is still one of the most challenging problems in computer vision research community.

The process of face recognition can be divided into three steps in order to develop a systematic scheme. The first step is face detection [56], i.e., finding face in the given image. An image captured in an uncontrolled environment contains background and other objects. We first need to separate the face from other regions to perform recognition. Face detection in itself is a big challenge because the faces may vary in pose and orientation and have different color and complexion which further vary due to the lighting and background effects. For performing robust face recognition, output of the face detection process has to be accurate. Simultaneously, the system must be robust to typical image-acquisition problems such as noise, video-camera distortion, image resolution, scaling factor and other environmental constraints.

The unconstrained face images have variations due to the position of the camera, distance between the person and the camera, and scaling factors. The images may also have color inconsistencies, varying pose, and illumination effects. These effects lead to degradation in the recognition performance. Thus a normalization algorithm, which is the second step in the recognition process, should be applied to compensate these effects. Researchers have used different algorithms for normalizing face images such as histogram equalization, and diffusion methods.

The third step in the process is face recognition which includes feature extraction and matching. Detected and normalized faces are the input to this module. Features are extracted using the feature extraction algorithm and then the extracted features are matched with the stored features using the matching algorithm. In the literature, there are several algorithms for face recognition proposed by various researchers. The first documented work in face recognition came in 1966 by Bledsoe [6]. Since then, researchers have used several approaches to address the problem of face recognition such as geometrical features [12], [26], eigen faces [52], local features [35], neural networks [31], elastic bunch graph matching [55] and wavelets [32].

Face recognition algorithms can be divided into two main categories: feature based algorithms and image based algorithms. Features based algorithms extract a set of geometrical features from the face image such as the size of eyes, nose, distance between eyes and nose, and match them using correlation filters or Euclidean distance. Processing of these fiducial points is performed independently and then combined to perform recognition. Some of the feature based approaches are Geometrical Features [12], [26], Dynamic Link Architecture [28], Hidden Markov Model [43], and Convolution Neural Network [31]. These approaches are generally invariant to illumination changes, but need proper normalization for scale factor. Generally, these algorithms do not perform well for variations such as pose and expression because in these variations, we do not get the original shape and size of the features. In the case of expression variation, feature dimensions change with the change in expressions. On the other hand, image based approaches take the complete face image as input and treat the image data simultaneously without attempting to localize individual points. The face is recognized as one entity without explicitly isolating different regions in the face. Image based techniques utilize statistical analysis, neural networks and different transformations. Most of the image based algorithms are based on the appearance of an individual. Some of the image based approaches are Eigenface [2], Linear Discriminant Analysis [2], and Independent Component Analysis [1]. There are other algorithms which combine the image and the feature based approaches such as, component based algorithms [24], Local Feature Analysis [35], Eigenfaces [52] and Eigenmodules [36].

1.1 Challenges of Face Recognition

Face recognition is an unobtrusive biometric trait, whose characteristic features change with time, and the rate of change in the facial appearance and features also vary. Computer based automatic face recognition finds the three tasks of detection, normalization, and recognition incredibly difficult. Even after 40 years of research, the problem of face recognition is not completely solved. According to the FRVT 2000 [5] and FRVT 2002 [38] evaluations, automatic face recognition performance decreases approximately linearly when comparing elapsed time database and new images. The identification and the verification rates are higher for older people compared to the younger people. For identification and watch list tasks, performance decreases linearly in the logarithm of the database size or the watch list size. Finally, it has been suggested that the outdoor face recognition needs lots of improvement.

In the face recognition research community, several challenges for face recognition have been enlisted such as pose, expression, illumination, background and aging. According to our interpretation of the face recognition problem, we consider the following as the challenges of face recognition:

- Training dataset - matching the face images when a limited number of training data for each individual is available.
- Pose and Expressions - matching the face images with variations in pose and facial expression.
- Illumination - matching the face images captured under different lighting conditions.
- Background - matching the face images when noise is present due to background.
- Accessories - matching face images with and without facial accessories such as glasses, makeup and jewelries.
- Aging - matching the face images when there is a significant time difference between the database image and query image.

- Scanned images - matching the scanned face image with the digitally captured face images.
- Sketches - matching the sketch images generated by the artists with the digital face images.
- Disguise - matching the face images in the scenario when someone wants to disguise his/her identity or impersonate someone's identity using makeup tools and artifacts.

In literature, the challenges of pose, and illumination, are addressed by various researchers [58]; but some of the above listed challenges such as minimum training data [57], scanned face recognition [50], sketch face recognition [51] and disguise [40] still remain. There are several applications in which face recognition is used for identification by law enforcement agencies. In these application scenarios the above listed challenges become relevant. One such application is recognizing an individual with a missing person database. Police may use a photograph for a missing person and then try to match the given face with the faces stored in their database. The challenge in this application scenario is that we have to deal with the scanned photographs of varying resolution and the photograph of the missing person may not be recent or it may contain only partial portion of the face with variations in background and illumination. In such cases, we do not have the choice of using a large amount of data to train the recognition algorithm, and hence the number of training data required is a challenge. An example of such images is shown in Figure 1.1. In this figure, we observe that the quality of the three scanned images is not as good as the digital face images. Other than missing person database, there are some law enforcement applications which require matching to be performed between digital and scanned face photograph of a person obtained from the passport, driver license, or an identity card. Typically, matching in such applications is performed manually because there may not be enough reference photographs available. Photographs may not conform to standard sizes, there may be variations in image resolution, an individual's features change over time, and the quality of the scanned images is mediocre. Matching of faces from the scanned photographs is a research challenge which has

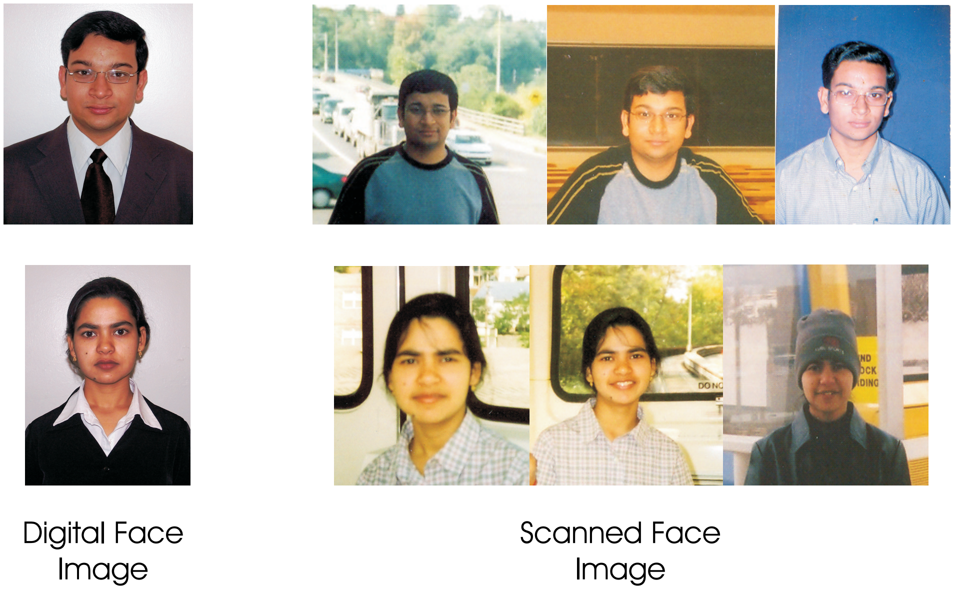

Figure 1.1: Digital and scanned images

received very little attention in the literature [22], [23], [50] and further research is needed.

Another application of face recognition is the matching of face sketches with the photograph images. Sketch images are generated by an artist using the information provided by eye-witness of a crime scene. Using sketch image, identification from the digitally stored face database of criminals is a difficult task because both the images are in entirely different domains [51]. Current systems perform the identification manually and there is a need for additional research in this area.

Face recognition or other biometric systems are developed not only to identify the genuine users but also to reliably identify impostors. Thus another challenge of face recognition is recognizing individuals when there are chances of disguised identity. If a person wants to hide his or her identity or impersonate someone else, the face recognition system should identify these cases. Figure 1.2 shows a person with different appearances. Which existing face recognition algorithm can recognize all these variations of an individual? While some of the researchers have tried to find a solution to this challenge [11], [40], [46], existing algorithms are limited to certain types of disguise.

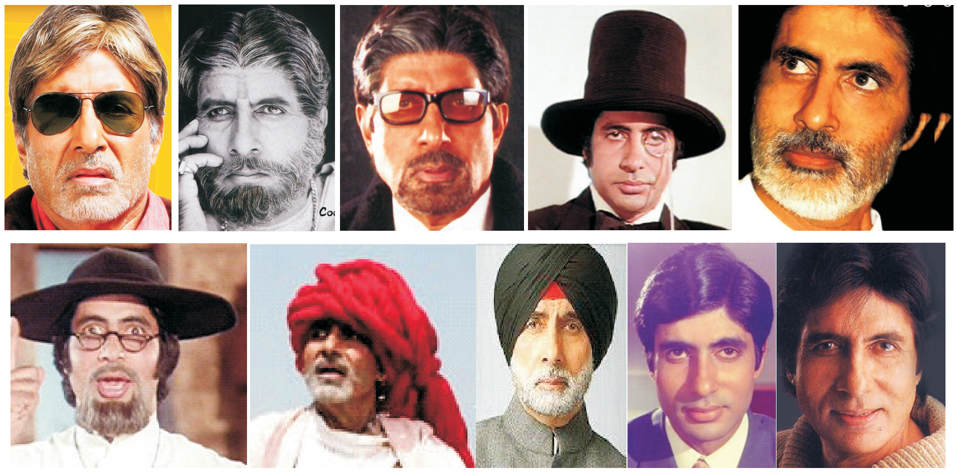

Figure 1.2: Images showing variations for disguise

1.2 Research Objective

This research focuses on the challenges of face recognition which are usually encountered in law enforcement applications. The research objectives are:

1. Design an illumination and rotation invariant face recognition algorithm for identifying faces with single training image.

2. Design a normalization and recognition algorithm such that the scanned and the digital face images can be matched efficiently.

3. Design a face recognition algorithm which can identify the variations for disguise and prepare a synthetic face database to evaluate the performance of the recognition algorithm.

1.3 Contribution of the Thesis

In this research work, we have made an attempt to simulate the visual cortex of the human mind by relating some of the results published by psychology researchers [13], [16], [54], with computer vision knowledge. We have used 2D log-polar Gabor wavelet [18] for face recognition because the representation of the 2D log polar Gabor wavelet is similar to the representation of visual cortex of the human mind [15].

We propose face recognition algorithm which uses 2D log-polar Gabor to extract the textural features in the form of amplitude and phase features. The algorithm has been applied to address the challenges of minimum training dataset, illumination, scanned images, disguise and accessories for face recognition.

1. Minimum Training Dataset: The proposed face recognition algorithm is evaluated using only one image for training. The algorithm is validated on different face databases including FERET [37], CMU-AMP Face Expression Database [60], CMU-PIE Database [47], and Notre Dame Face Database [10], [19]. The proposed algorithm is also compared with the standard face recognition algorithms such as Local Feature Analysis [35], Correlation Filters [44], $E(PC)^2A$ [57], and SVD perturbation [8].

2. Scanned Face Recognition: We applied the proposed face recognition algorithm to match a scanned face image with a digital face image. We also proposed a preprocessing and normalization algorithm to transform both the scanned and the digital face images into a common domain using histogram equalization [21], Multiscale Retinex [39], and Eigenspace decomposition [52]. Face recognition algorithm is then applied on the local features of both the images and face templates are generated. These face templates are matched using weighted matching scheme. A comparison is performed with face recognition algorithm based on the Dissimilarity Measure [50], Texture Features [48], Correlation Filter [44] and Fisher Linear Discriminant Analysis [2].

3. Disguise: Since there is no face database available in public domain which contains variations in disguise, we have created a synthetic face database for disguise using the software Faces v2.0 [59]. This database contains the variations in hair style, facial accessories, glasses, aging, cap and hat, beards and mustaches, and combination of these variations. The proposed face recognition algorithm is applied on these images using single image for training. A comparison has also been performed with other

existing algorithms such as Principal Component Analysis [52], Geometrical Features [12], and Local Feature Analysis [35].

1.4 Organization of the Thesis

Subsequent chapters provide the details of challenges we have undertaken in this research work with literature survey, proposed new algorithms and experimentally validated the results. Chapter 2 explains the significance of 2D log-polar Gabor wavelet and the visual cortex of human mind. The proposed face recognition algorithm using amplitude and phase feature extraction and matching is also presented. Chapter 3 explains the challenges with scanned face recognition. The proposed face normalization and recognition algorithm for scanned images and the results are described in this chapter. Chapter 4 shows the challenge of disguise in face recognition and presents the synthetic disguise face database along with the experimental results using the proposed face recognition algorithm.

Chapter 2

Face Recognition using Log-polar Gabor Wavelet

An image (from Latin word imago) is an artifact that reproduces the likeness of some object. Several studies have been undertaken in the computer vision community related to feature detection, classification based on properties such as geometrical features, and textural features. The problem of face recognition comes under the realm of computer vision. As described in Chapter 1, a number of challenges exist even after 40 years of research in the field. Researchers have proposed several approaches to address the challenges of face recognition. One way to approach the problem of face recognition is to identify how visual cortex of the human mind works and simulate similar processing using a computer.

2.1 Overview of Gabor Wavelet

Researchers from psychology and physics have simulated and analyzed the human visual cortex by conducting certain experiments. It has been reported that the properties of a human visual cortex can be represented by the Gabor wavelet [13], [16], [54]. In 1946, D. Gabor [20] established the uncertainty principle where the product of uncertainties in frequency Δf and time Δt must exceed a fixed constant. If the signal is transmitted through a bank of bandpass filters, the closest frequencies which can be distinguished are given by

$\Delta f = 1/\Delta t$. In Fourier space, the product of time and bandwidth gives the maximum amount of information that can be transmitted. Gabor showed that the minimum area in Fourier space is achieved by the Gaussian modulated complex exponential functions of the form

$$\phi_{jk}(t) = exp\left[-\pi(t-j\Delta t)^2/\alpha^2\right] exp\left[2\pi ik\Delta f(t-j\Delta t)\right] \tag{2.1}$$

where $\Delta f \Delta t = 1$. In this equation, the first exponential term represents a Gaussian envelop centered on $j\Delta t$ and the second term represents the conjugate exponential form of the trigonometric functions of frequency $k\Delta f$ which is a periodic function, and denotes the locality of the Gaussian envelop. Equation 2.1 can also be represented as,

$$\phi_{jk}(t) = exp\left[-\pi(t-j\Delta t)^2/\alpha^2\right]\left[cos[2\pi ik\Delta f(t-j\Delta t)] + isin\left[2\pi ik\Delta f(t-j\Delta t)\right]\right. \tag{2.2}$$

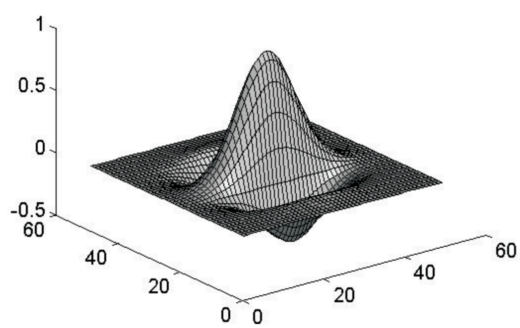

Figure 2.1: Even and odd components of the 2D Gabor filter

As $\alpha \rightarrow \infty$, Gabor wavelet reduces to Fourier transform and no locality is shown at this point. Experiments performed by the researchers show that the receptive fields of simple cells in a visual cortex are similar to the Gaussian modulated sinusoids and 97% of them are statistically indistinguishable from the odd-symmetric or even symmetric parts of a 2D

Gabor elementary function [15]. This quadrature phase relationship can be computed from the odd and the even symmetric parts of the 2D Gabor wavelet [15], as can be seen in Equation 2.2. Figure 2.1 shows the odd and the even symmetric parts of the 2D Gabor wavelet. Gabor wavelet is a non-orthogonal wavelet and according to Daugman [16], non-orthogonal representations are ubiquitous in biological sensory and motor systems.

Another property of neurophysiological and psychological system is that they show a log-polar distribution of response selectivity in cells in the visual cortex, which shows the orientation half-bandwidth of $\pm 15^0$ and the frequency bandwidth of 1.5 octaves [17]. To account for this property, we have used log-polar Gabor wavelet for representing the face image in which radial distance represents the spatial frequency and the polar angle represents the orientation. Log polar Gabor wavelet is a form of Gabor wavelet [4], [18] which is based on polar coordinates and the dependency of directional independent variance (σ) on the polar coordinate is realized by a logarithmic scale. Thus the functional form of 2D log polar Gabor filter can be represented as

$$G_{r_0,\theta_0}(r,\theta) = exp\left(-2\pi^2\sigma^2\right)\left[(ln(r) - ln(r_0))^2 + (ln(r)sin(\theta - \theta_0))\right] \tag{2.3}$$

and the position of filter in the Fourier domain is defined by

$$r_{00} = \sqrt{2}, r_{0i} = 2^i * r_{00}, \theta_{0i} = i * \frac{2\pi}{N_\theta} \tag{2.4}$$

where r_{00} is the smallest possible frequency, N_θ is the number of filters on the unit circle, and at index L, σ_L and s_L are further defined by

$$\sigma_L = \frac{1}{ln(r_0)\pi sin(\pi/N_\theta)}\sqrt{\frac{ln2}{2}} \tag{2.5}$$

$$s_L = \frac{ln(r_0)\pi sin(\pi/N_\theta)}{ln2}\sqrt{\frac{ln2}{2}} \tag{2.6}$$

Log-Gabor wavelet has no DC component and has an extended tail. According to Field [18], log-Gabor functions with extended tail are able to encode the images more efficiently compared to Gabor wavelet, because Gabor wavelet would over represent the low frequency components and under represent the high frequency components. Log-polar Gabor wavelet also provides invariance to rotation and scaling.

2.2 Face Recognition Using Amplitude and Phase Features

In this section, we propose a face recognition algorithm which is based on amplitude and phase information extracted from a face image. The algorithm is described as follows:

Let the size of face image $F(x, y)$ be $N \times N$. Face image is transformed into its log-polar form $F(r, \theta)$

$$F(r,\theta) = F\left(\left[\frac{N}{2}\right] + \left[\theta cos\left(\frac{2\pi r}{s}\right)\right], \left[\frac{N}{2}\right] - \left[\theta sin\left(\frac{2\pi r}{s}\right)\right]\right) \tag{2.7}$$

where $r = 0, ..., s-1$, $\theta = 0, ..., [N/2]-1$, and s is the factor by which the image is sampled from 0^0 to 360^0 to produce its equivalent polar form. Polar face image is expressed into the 2D Fourier domain,

$$F(\rho,\phi) = \Sigma_r \Sigma_j F(r,\theta_j) exp\left[-i\frac{2\pi r}{\mu}\rho + \frac{2\pi j}{\upsilon}\phi\right] \tag{2.8}$$

where $\theta_j = 2\pi j/\upsilon$, $0 \leq r < \nu$ and $0 \leq j < \upsilon$ are the radial and angular frequency resolutions respectively. 2D Fourier transform of the face image is convolved with the 2D

Fourier transform of the log polar Gabor wavelet $G(\rho, \phi)$. The Inverse Fourier Transform (IFT) of the convolved face image is computed and the output $F_g(r, \theta)$ is a complex valued matrix containing the amplitude and the phase information. Amplitude and phase features are computed from the matrix $F_g(r, \theta)$ using Equation 2.10 and 2.11 respectively. Phase features are quantized using Equation 2.12 to generate a binary representation of phase features, referred as phase template. The amplitude and quantized phase features are shown in Figure 2.2,

$$F_g(r, \theta) = IFFT(\Sigma_r \Sigma_j F(\rho, \phi) * G(\rho, \phi)) \tag{2.9}$$

$$A(r, \theta) = \sqrt{ReF_g(r, \theta) + ImF_g(r, \theta)} \tag{2.10}$$

$$P(r, \theta) = tan^{-1}\left(\frac{ImF_g(r, \theta)}{ReF_g(r, \theta)}\right) \tag{2.11}$$

$$B_p[r, \theta] = \begin{cases} [1, 1] & if \quad 0^0 < P(r, \theta) \leq 90^0 \\ [0, 1] & if \quad 90^0 < P(r, \theta) \leq 180^0 \\ [0, 0] & if \quad 180^0 < P(r, \theta) \leq 270^0 \\ [1, 0] & if \quad 270^0 < P(r, \theta) \leq 360^0 \end{cases} \tag{2.12}$$

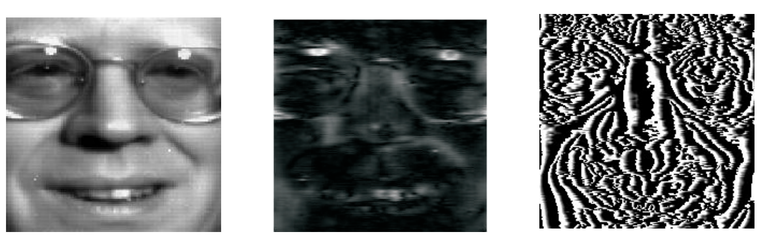

Figure 2.2: (a) Face image (b) Amplitude features, (c) Phase features [61]

Amplitude features are matched using the steps described below:

- Two amplitude features, A^1 and A^2 are divided into z number of frames, each of size $k \times l$.

- Correlation V_1 between the two corresponding frames is computed using Equation 2.13. A circular shift of 8 pixels in all directions is applied to make the process shift invariant. Using the frame matching threshold λ_1, which is the intermediate matching score for the frames, M_s^A is calculated. Final amplitude matching score M^A is obtained by dividing the intermediate matching score by the number of frames.

$$V_i = \frac{A_i^1 \oplus A_i^2}{k * l} \tag{2.13}$$

$$M_s^A = \begin{cases} M_s^A + 1, & if \quad V_i \geq \lambda_1 \\ M_s^A, & if \quad V_i < \lambda_1 \end{cases} \tag{2.14}$$

To match the two phase templates,

- Phase templates are divided into m frames each of size $p \times q$.
- Corresponding frames from the two phase templates are matched using hamming distance as shown in Equation 2.15. A circular shift of 8 pixels in all directions is applied and the minimum hamming distance is computed to make the process shift invariant.

$$D_i = \frac{\Sigma B_i^1 \otimes B_i^2}{p * q} \tag{2.15}$$

where B_i^1 and B_i^2 are the i^{th} frames for the two templates, and D_i is the corresponding distance measure.

- Phase matching score M_s^P is calculated using Equation 2.16,

$$M_s^P = \begin{cases} M_s^P + 1, & if \quad D_i \geq \eta_1 \\ M_s^P, & if \quad D_i < \eta_1 \end{cases} \tag{2.16}$$

$$M^P = \frac{M_s^P}{m} \tag{2.17}$$

where η_1 is the frame matching threshold, M_s^P is the intermediate matching score for frames and m is the number of frames.

- A match occurs if the matching score $M^P < \eta_2$, where η_2 is the phase matching threshold for the face template.

Here we transform the problem of recognizing a face into an efficient test of statistical dependence operating on amplitude variation and statistical independence operating on phase variation. Phase and amplitude are the two textural features that provide invariance to several image transformations. Figure 2.3 shows the relationship between phase P and amplitude A of an image $F(x, y)$. Venkatesh and Owens [53] show that the degree of phase is independent of overall magnitude of the face image which provides invariance to the changes in illumination and contrast. Another advantage of the proposed face recognition algorithm is that amplitude and phase features are invariant to frequency, scale, and orientation of the filter. Thus the proposed face recognition algorithm is invariant to frequency, scale, filter orientation, illumination and contrast.

2.3 Performance for Single Training Face Images

Face recognition algorithms such as Local Feature Analysis [35] (LFA) and Correlation Filters [44] (CF) need a large set of training images. These algorithms learn from the training data and use the feature set for matching. However if the training set is small then performance of these algorithms decrease. Typically, for law enforcement applications only a single face image is available for training. In [8] and [57], face recognition algorithms are proposed to solve the problem of single training images using $E(PC)^2A$ and SVD Perturbation respectively. These algorithms achieve an accuracy of around 85% on the FERET database

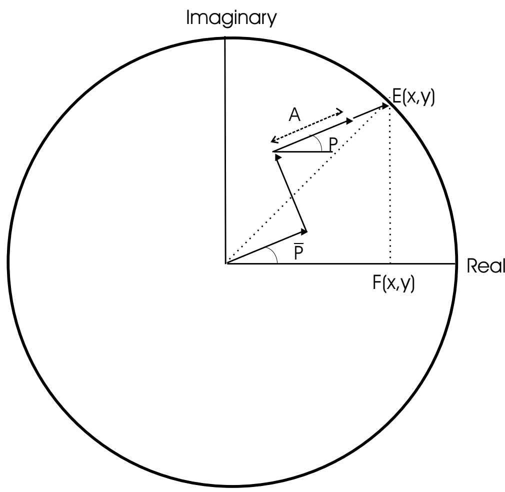

Figure 2.3: Relationship between amplitude (A) and phase (P) of an image

[37]. In this section we present the results of face recognition obtained using amplitude and phase features extracted from Gabor wavelets. Results are presented for single training image in the database, and also with varying the number of training images in the database. A comparison of the proposed approach with other existing algorithms is also presented.

The proposed algorithm is tested on frontal face images from the colored FERET database [37]. From this database we have chosen 3000 frontal face images from 600 individuals with variations in pose from 0 to 10 degrees. Using this database, the number of training images for each individual is varied from one to four. Figure 2.4 shows the receiver operating characteristics (ROC) plot of the amplitude and the phase features. We found that phase features perform better than amplitude features with an approximate difference of around 5% in the equal error rates (EER) of the two.

The proposed algorithm is compared with LFA [35] and CF [44] whose performance is known to be good when large training set is used. Figure 2.5 shows that when the number of training image is four, identification accuracy of all the algorithms is comparable. However, as the number of training images is gradually reduced, the accuracy of LFA and CF based algorithms drops significantly compared to the proposed algorithm. For example, experimental results show that with a single training image, accuracy of the LFA algorithm is reduced to 64.21% resulting in a drop of 27.95%, while accuracy of the CF algorithm is reduced to 71.66%, resulting in a drop of 22.21%. Comparatively, the 2D log Gabor algorithm maintains a higher level of accuracy at 89.14%, resulting in a drop of only 5.05%.

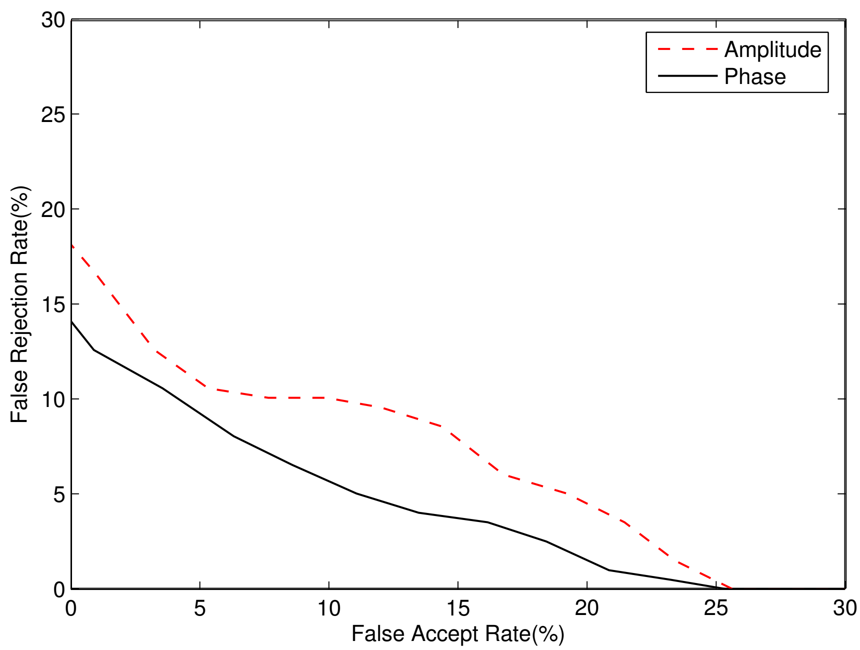

Figure 2.4: Accuracy of the proposed algorithm using the amplitude features and phase features on FERET face database [37]

We next compare our results with two recent face recognition algorithms, $E(PC)^2A$ [57] and SVD Perturbation [8], that use single training images. The same set of images from FERET database [37] are used for comparison. Experimental results show that the proposed algorithm has a FRR of 9.17%, FAR of 1.69%, and an overall accuracy of 89.14%. An improvement of 5% in accuracy is achieved compared to the $E(PC)^2A$ algorithm [57] which has an accuracy of 84.55%. The 2D log Gabor algorithm also shows an improvement of 3% compared to the SVD Perturbation algorithm [8] which has an accuracy of 86.24%.

We also validated the performance of the proposed face recognition algorithm on the CMU AMP Face Expression Database [60] which contains variation in expressions. Figure

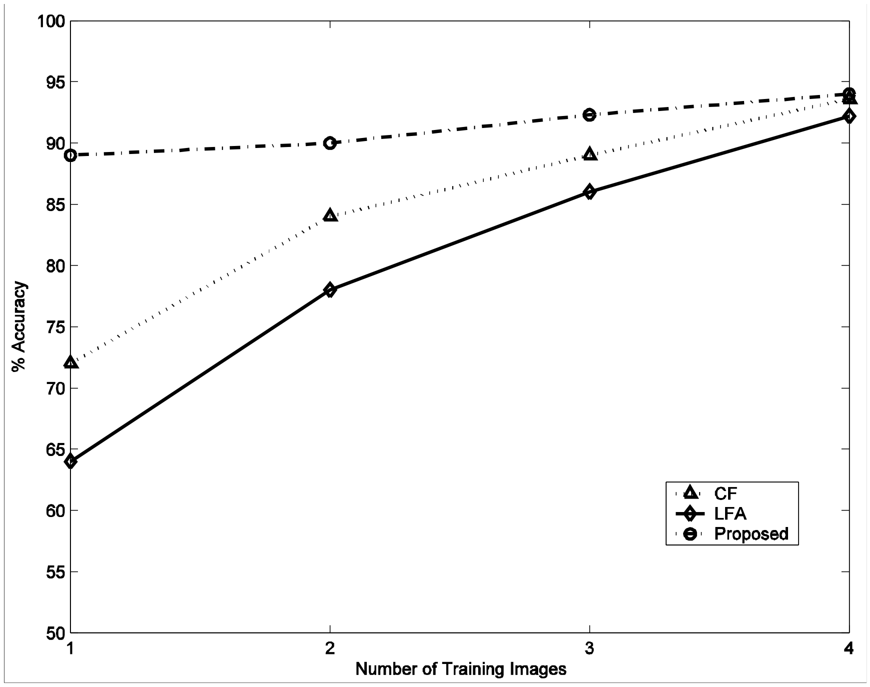

Figure 2.5: Accuracy of algorithms on varying the number of training images

2.6 shows performance of the amplitude and the phase features with single training image in the database. This experiment further reinforced that the phase features give better performance compared to the amplitude features. The experiment shows the proposed face recognition algorithm is robust to different facial expressions.

For further validation, additional experiments have been performed using the CMU PIE Database [47] which contains variations in pose, expression, and illumination. We selected frontal images from this database having variations in illumination, pose and expression. ROC plot is shown in Figure 2.7. Performance of the face recognition algorithm was also tested on visible images from the Notre Dame Database [10], [19]. This database contains variations in expressions with significant time difference between images. The ROC plot is shown in Figure 2.8.

Experiments on these databases demonstrate that the phase features perform better than the amplitude features. These results are obtained using only one image per person to train the face recognition algorithm. The availability of only one image for training is typical in

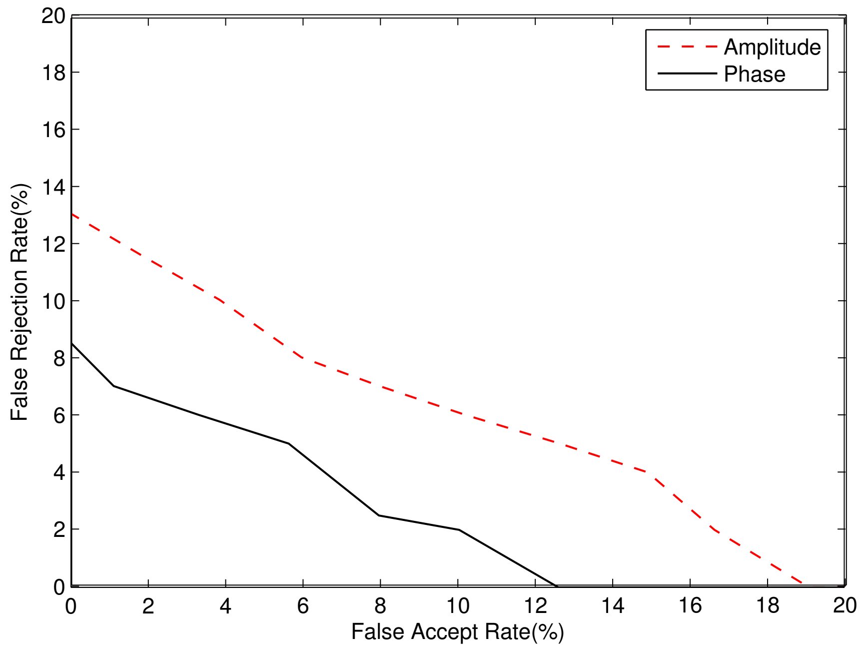

Figure 2.6: ROC plot for CMU AMP facial expression database

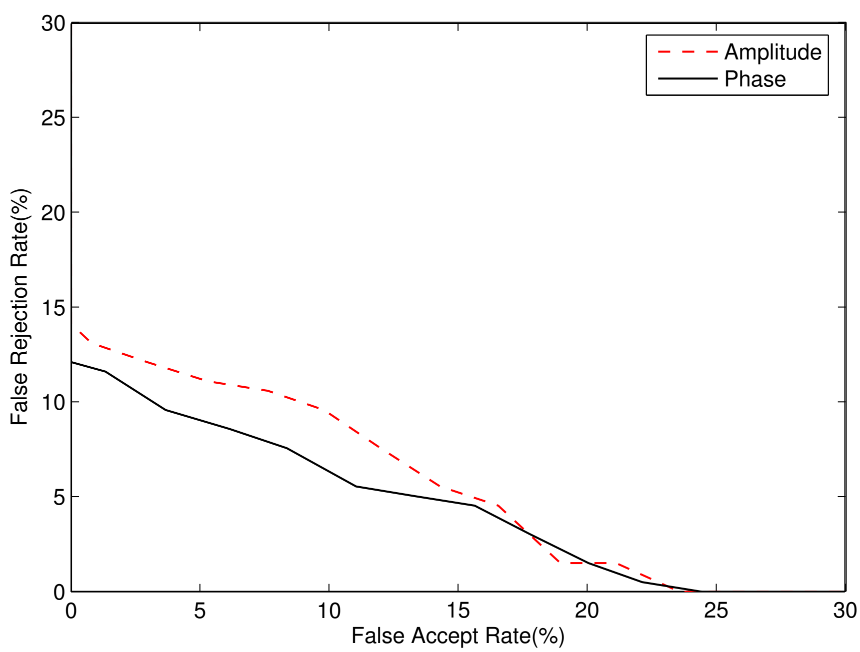

Figure 2.7: ROC plot for CMU PIE database

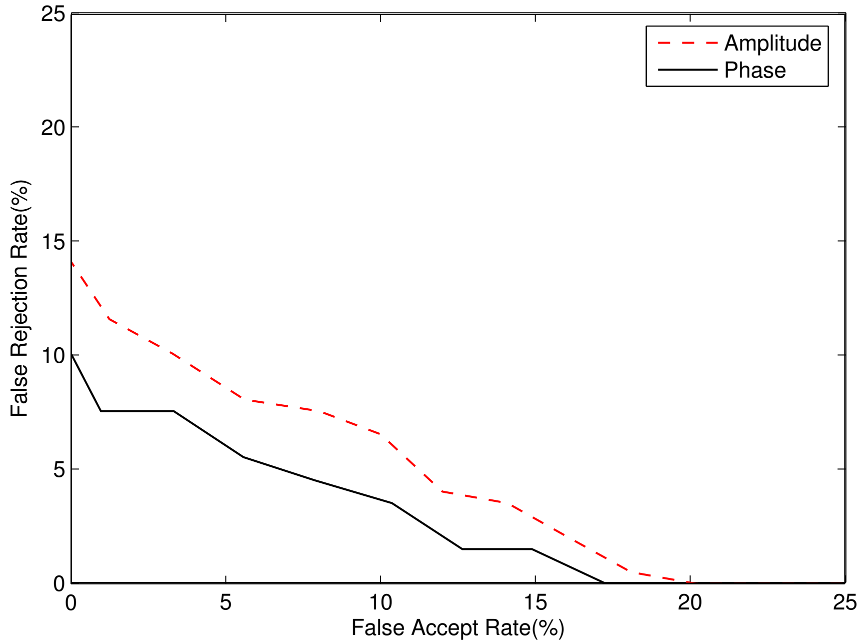

Figure 2.8: ROC plot for Notre Dame database

a law enforcement application. The level of performance shows that the proposed algorithm is robust to some of the challenges such as illumination, minimum training data and to a certain extent, variations in pose and expression.

Chapter 3

Face Recognition using Scanned Images

3.1 Introduction

Law enforcement applications usually require matching to be performed between digital face photograph in a given database and scanned photo of a missing person, from passport, driver license, or an identity card. Typically, the matching is performed manually since automatic matching is very challenging. This is because there may be only one reference photograph available, photographs may not conform to standard sizes, there are variations in image resolution, an individual's features change over time, and the quality of scanned images is mediocre. In [50], an algorithm using preprocessing techniques and dissimilarity measure is used to match document face image with a digital face image.

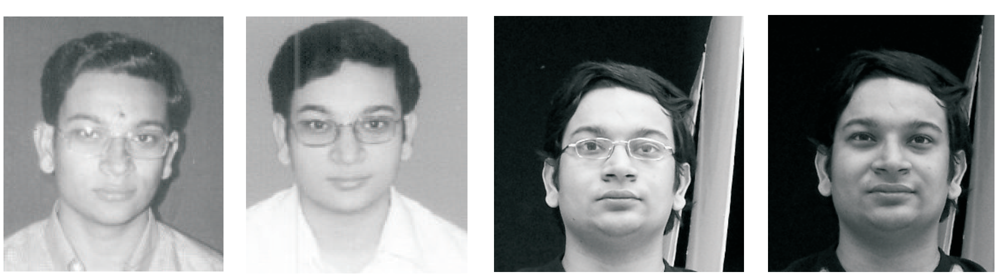

Figure 3.1: Scanned and digital face images of an individual

Figure 3.1 shows an example of the scanned and the digital images of an individual, over a period of time. The first two are scanned images and the last two are digital images and all of them are captured in different environment and lighting conditions. The first scanned image contains shadow effect, while the second scanned image has uniform lighting conditions. The two digital images are captured under normal lighting conditions. Difference among the images can be seen visually. For an automatic face recognition system, it is very challenging to match these two sets of images. Also scanning of the images cause irregularities, which sometimes distort the facial features. This is a very simple example of variations found in images captured in unrestricted environment, similar to those encountered in the law enforcement applications.

An efficient face recognition algorithm should be able to handle all these types of variations. There are several regions in a face such as cheeks and forehead that do not contribute much to the interclass variability and intra-class similarity of two face images. For example, shape of the mouth changes significantly when a person smiles while the eye and the nose hardly undergo any change. The proposed face normalization and recognition algorithm takes advantage of these static and dynamic properties to match a scanned face with a digital face.

Based on all these observations, we first transform the digital and the scanned images into a common domain using Multiscale Retinex and Histogram equalization. For matching, we use the phase information of the local textural facial features and combine them with a weighting scheme for different features to make the algorithm robust.

3.2 Matching Scanned Face Image to Digital Face Image

3.2.1 Preprocessing and Image Enhancement

In preprocessing step, we first normalize the digital and the scanned face images captured under different conditions. Face is detected from both the images using a triangle based face detection algorithm [49] in which first the eyes and nose coordinates are marked and then using the three coordinates the face is detected. Further, the detected face is aligned with respect to the three coordinates to handle the scaling and rotation in the image. We used affine transformation [21] to align the face images which is a combination of image rotation and translation.

Histogram equalization [21] is performed on detected and aligned face images to enhance the intensity. Further, Multi-scale Retinex (MSR) algorithm [39] is used for quality enhancement and removal of deterioration due to scanning. Multi-scale Retinex is a modified form of the Single-scale Retinex (SSR) [30]. Retinex is an image enhancement algorithm that provides dynamic range compression and color constancy as described below.

Let $F_i(x, y)$ be the histogram equalized face image distribution in the i^{th} spectral band. The SSR output $F_R(x, y)$ is written as,

$$F_R(x, y) = logF_i(x, y) - log\left[S(x, y) \otimes F_i(x, y)\right] \tag{3.1}$$

where $\otimes$ denotes the convolution operation, $S(x, y)$ is the surround function and $S(x, y) = Ke^{-r^2/c^2}$. Here $r = \sqrt{x^2 + y^2}$, c is the Gaussian surround space constant. Depending on the value of c, retinex either performs dynamic range compression at the cost of poorer color rendition or good color rendition at the cost of dynamic range compression. Small value of c provides good dynamic range compression and large value of c provides good color rendition. We empirically determined that for smaller values of c, the recognition performance is better. The recognition performance with different values of c is shown in Figure 3.6. K is selected such that

$$\int\int S(x, y)dxdy = 1 \tag{3.2}$$

MSR is a weighted sum of the outputs of several different SSR outputs. It can be expressed as

$$F_{MSR_i} = \sum_{n=1}^{N} \omega_n F_{Rn_i} \tag{3.3}$$

where N is the number of scales, F_{Rn_i} is the i^{th} component of the n^{th} scale and ω_n is the weight of the n^{th} scale. Here the surround function is also modified to incorporate the scale factor. Modified surround function can be written as

$$S_n(x, y) = Ke^{-r^2/c_n^2} \tag{3.4}$$

Output of histogram equalization and MSR of both the scanned and the digital face images for different values of c are shown in Figure 3.2. After normalization, we transform the normalized digital face image into scanned image domain using the transformation algorithm described in Section 3.2.2.

Figure 3.2: Scanned and digital face images obtained after applying multiscale retinex

3.2.2 Digital Image to Scanned Image Transformation

Even after preprocessing, sometimes the difference between digital and scanned image of the same person is greater than the difference in digital image of two different persons. To overcome this difference, a transformation algorithm is proposed to transform the digital face image into a scanned face like image. An Eigen space is created using scanned face image, and digital face image is projected into this space for transformation. The algorithm is described as follows:

- Let $T_D = [\overrightarrow{T}^1_D, \overrightarrow{T}^2_D, ..., \overrightarrow{T}^n_D]$ denote the enhanced digital training images and $T_P = [\overrightarrow{T}^1_P, \overrightarrow{T}^2_P, ..., \overrightarrow{T}^n_P]$ denote the enhanced scanned photo images.
- Average digital face image $\overrightarrow{a}_D$ and average photo face image $\overrightarrow{a}_P$ are computed for the digital and the photo training set respectively.
- Eigenvector matrix E_D is then computed from the digital training set.
- Let the input digital face image be $\overrightarrow{F}_D$. Average digital face $\overrightarrow{a}_D$ is subtracted from $\overrightarrow{F}_D$ ($\overrightarrow{F}_D = \overrightarrow{F}_D - \overrightarrow{a}_D$).
- $\overrightarrow{F}_D$ is then projected in the eigen-space of $\overrightarrow{E}_D$ to compute the weight vector $\overrightarrow{\omega}_D$.
- Digital face image is reconstructed using the following equation,

$$\overrightarrow{F}^R_D = T_P V_D D_D^{-1/2} * \overrightarrow{\omega}_D \tag{3.5}$$

  where V_D is the eigenvector matrix of digital face image training set, D_D is the diagonal eigen value matrix.

- Average photo face $\overrightarrow{a}_P$ is added with the reconstructed digital face image to obtain the transformed face image $\overrightarrow{F}^T_D$,

$$\overrightarrow{F}^T_D = \overrightarrow{F}^R_D + \overrightarrow{a}_p \tag{3.6}$$

As shown in Figure 3.3, quality of the reconstructed transformed digital face image is close to the scanned face image.

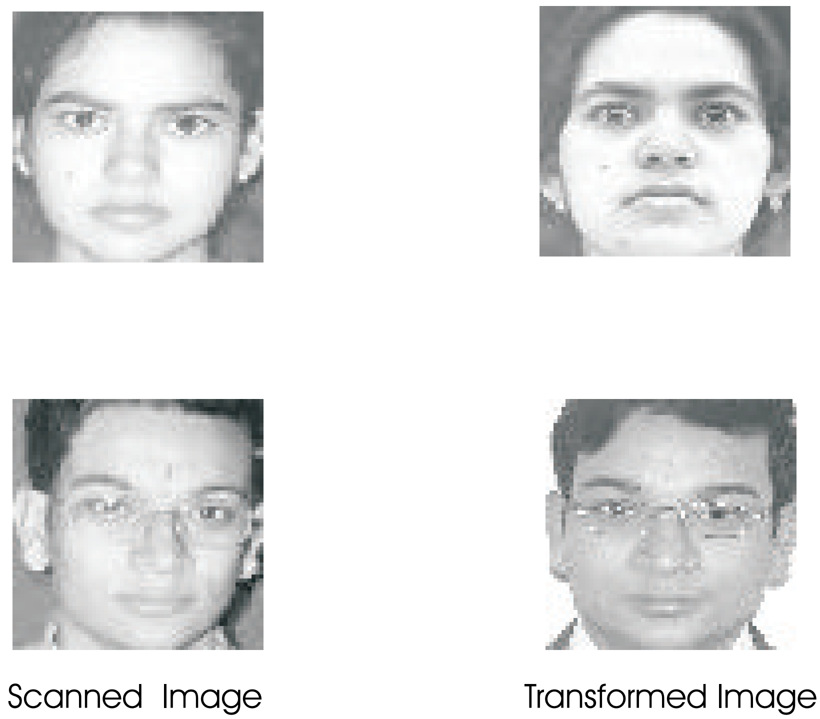

Figure 3.3: Digital image transformed to scanned image domain

3.2.3 Face Recognition using Local Textural Features

Since there are several regions in the scanned and the digital images that do not have similar texture, we apply the proposed phase feature extraction based face recognition algorithm only at the local features. Texture feature extraction algorithm extracts local features from the face image which add to uniqueness of the face. Local edge and corner detection algorithm [45] is first applied on the preprocessed face image to detect corner points in the face. Neighboring corner points are then grouped in a circular form to represent the local regions. Figure 3.4 shows local features extracted in the preprocessed scanned and the transformed digital images. Local edges and corners are detected using the corner detection algorithm [45]. Corners lying in close vicinity are grouped to represent a local region, for example, the corners near the left eye are grouped to represent the left eye.

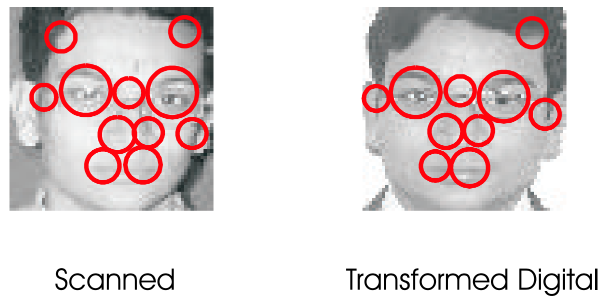

Figure 3.4: Local features extracted from scanned and transformed digital face images

This gives a set of prominent regions for the face. The number of regions can vary at different instances for individuals with change in pose, facial expression, occlusion and illumination. Following steps are performed to extract textural features from these local regions:

- Let $f_1, f_2, ..., f_n$ be the local facial regions. Each region is converted into the dimensionless polar coordinate system so that they become invariant to imaging distance and rotation.

- Let $f'_1, f'_2, ..., f'_n$ be the features in polar coordinate system. These features are convolved with the 2D log polar Gabor wavelet in frequency domain. If 2D log polar Gabor filter is G and facial feature is f'_i, then convolution can be represented as,

$$f''_i = IFFT(FFT(f'_i) * FFT(G)) \tag{3.7}$$

  where FFT stands for Fast Fourier Transform and IFFT stands for Inverse Fast Fourier Transform. The output f''_i contains both amplitude and phase information.

- Using Equation 2.12, phase quantization is applied to encode phase information in the form a binary template of size $p \times q$. Combining all the n features for one face gives a feature B^1 set which can be represented as $B^1_1, B^1_2, ..., B^1_n$.

One way to match the two sets of encoded patterns is to label the features and match them accordingly. But this method requires large training data and increases the computation time. So we propose an efficient matching strategy described below.

- Let B^1 and B^2 be the textural feature sets to be matched. Create a distance matrix D_{mn} of size (m, n) using the following equation,

$$D_{m\times n} = \begin{bmatrix} \frac{\sum B_1^1 \otimes B_1^2}{p\times q} & \dots & \frac{\sum B_1^1 \otimes B_n^2}{p\times q} \\ \vdots & \vdots & \vdots \\ \frac{\sum B_m^1 \otimes B_1^2}{p\times q} & \dots & \frac{\sum B_m^1 \otimes B_n^2}{p\times q} \end{bmatrix} \tag{3.8}$$

  where m and n are the number of facial features in the image 1 and 2 respectively, and $p \times q$ is the dimension of each encoded binary template.

- Minimum score for each column is preserved and remaining values are discarded. These minimum scores $d_1, d_2, ..., d_n$ are preserved based on the assumption that the corresponding features give the minimum match score, i.e., the score obtained by matching two nose will be less compared to the score of matching a nose with a mouth.

- These matching scores are combined using a weighted sum rule

$$S = \frac{\sum_{i=1}^{N} \omega_i d_i}{n} \tag{3.9}$$

  where ω_i is the weighting factor. The values of weighting factors depend on their corresponding matching scores and are computed from a lookup table (Table 3.1) where $\omega_1 < \omega_2 < \omega_3 < \omega_4 < \omega_5 < \omega_6 < \omega_7$. In this manner, more emphasis is given to the features which are matched with higher confidence or are mismatched with

Table 3.1: Lookup table used for assigning the weights

Range of Scores	0.0 - 0.1	0.1 - 0.2	0.2 - 0.3	0.3 - 0.4	0.4 - 0.5	0.5 - 0.6	0.6 - 1.0
Weighting Factor	ω_1	$\omega 2$	ω_3	ω_4	ω_5	ω_6	ω_7

higher confidence compared to the ones with matching score near the threshold. This matching strategy thus reduces the intra-class variation and increases the inter-class distances. The values of ω_i are empirical and the values of various weights we have used in our experiments are $\omega_1 = 0.012$, $\omega_2 = 0.074$, $\omega_3 = 0.259$, $\omega_4 = 0.281$, $\omega_5 = 0.576$, $\omega_6 = 0.613$, and $\omega_7 = 0.999$.

- A person is matched if the matching score is less than the matching threshold.

3.2.4 Experimental Results

Since there is no face database available in public domain which contains both the scanned and the digital images, we created a frontal face database of 500 individuals containing the scanned and the digital images. Digital images are captured in four different sessions with a time interval of three months in each session. Images have variations in facial expression, lighting, occlusion and camera properties. In addition, we have extended the database to include the scanned and the digital face images of various resolutions to validate the performance of the proposed algorithm. In the experiments, one scanned face image per person is used for training and the remaining scanned and digital face images are used for testing. Results in Chapter 2 show that the phase features perform better than the amplitude features. In this experiment we therefore use only the phase features for matching.

ROC plot in Figure 3.5 shows performance of the proposed algorithm at different resolution of face images with and without preprocessing, and digital image to scanned image transformation. We found that the performance increases with increase in resolution. The identification accuracy of 98.11% was achieved with a resolution of 160×120. Any further increase in resolution only marginally improved the identification accuracy. This graph also shows that the proposed preprocessing and transformation algorithm enhances the performance of face recognition algorithm by at least 8% compared to without preprocessing and transformation.

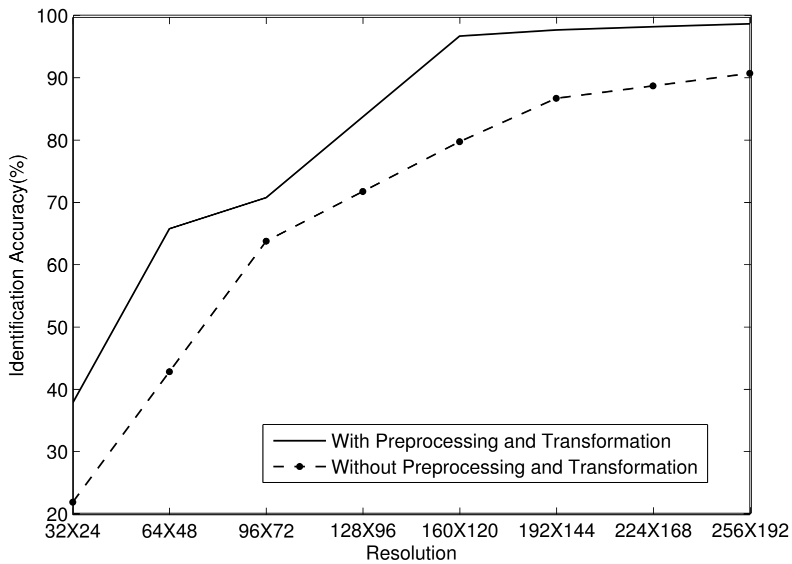

Figure 3.5: Accuracy of the face recognition algorithm at different resolutions

One major factor which affects the performance of Multiscale Retinex algorithm is the value of Gaussian surround constant c. We analyzed the performance of face recognition by varying this constant with the objective to find the optimal value, or range of values for c that would achieve the best illumination invariance. Figure 3.6 shows that when the value of $c = 4$, identification accuracy of around 98% is achieved. It is evident that retinex based normalization significantly improves the recognition rates. Lower values of c are better for illumination correction and as c increases, recognition rate decreases. Only exception occurs with the value of $c = 1$, where the recognition rates are much lower. This is explained by the fact that the images are overly grayed out with very small c, thereby hindering the face recognition algorithm from classifying the images correctly.

Performance of the proposed algorithm is compared with four existing face recognition

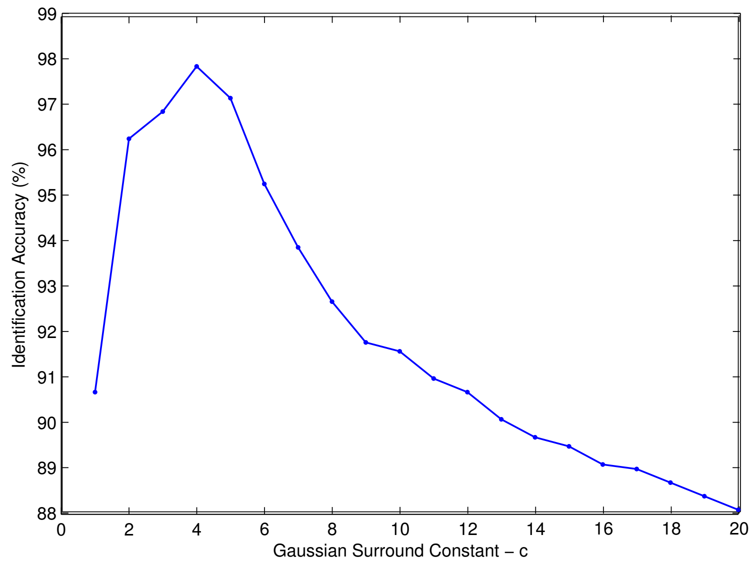

Figure 3.6: Performance of face recognition on varying the Gaussian surround constant c

algorithms namely Dissimilarity Measure (DM) [50], Texture Features (TF) i.e. the phase based face recognition algorithm described in Chapter 2, Correlation Filter (CF) [44] and Fisher Linear Discriminant Analysis (FLDA) [9]. The comparison is made using a single scanned face image for training. The ROC plot in Figure 3.7 shows that performance of the proposed algorithm is superior compared to the other four algorithms.

In this chapter, we have addressed with the challenge of scanned face recognition. The scanned and the digital images are preprocessed to transform them into a common domain. First, histogram equalization and Multiscale Retinex are applied on both the digital and the scanned images for illumination normalization. For further compensating the quality of the scanned image, eigen space based transformation is applied on the digital image. This transformation projects the preprocessed digital image into the eigenspace formed by the preprocessed scanned face images. The resulting digital image is thus of similar quality as the scanned image. For feature extraction and matching, the face recognition algorithm described in Chapter 2 is modified to apply at the local level. Local textural features are extracted from the local regions of both the scanned and the digital face images. These

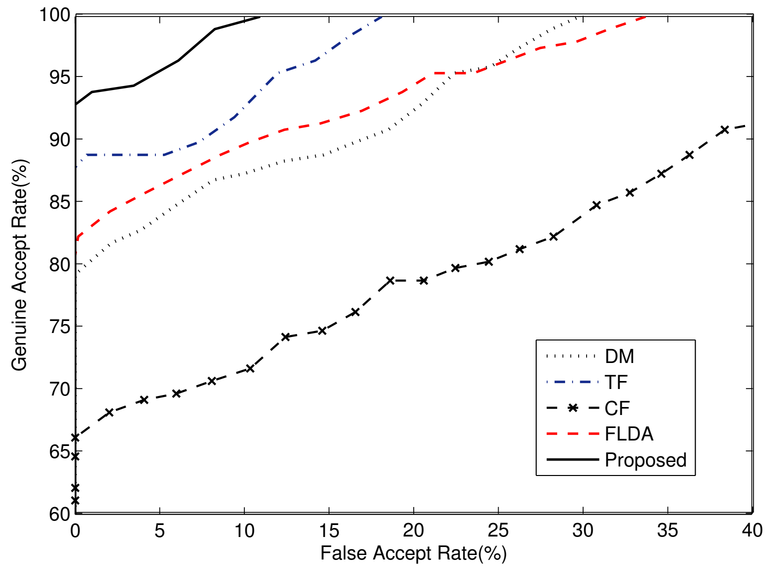

Figure 3.7: Performance comparison using ROC curves

features are encoded using 2D log polar Gabor wavelet and matched using the weighted matching strategy described in Section 3.2.3. Results show an accuracy of around 98% on a database containing both the digital and the scanned images from 500 classes.

Chapter 4

Recognition of Faces with Variations in Disguise

4.1 Introduction

Despite the success of automatic face recognition techniques in many practical applications, face recognition has difficulties in detecting identical twins and disguised faces, which can be critical in high-end security applications. Most of the research papers on face recognition claim excellent performance for user cooperative face recognition. But very few papers [11], [40], [46] discuss the possibility where someone wants to hide their identity or impersonate someone else using disguises.

Chellappa et al. [40] have studied the facial similarity for several variations including disguise by forming two eigenspaces from two halves of the face, one using the left half and other using the right half. From the test image, optimally illuminated half face is chosen and is projected into the eigenspace. This algorithm is tested on National Geographic database [40] which consists of variations in smile, glasses, glasses with varying left right light and an accuracy of around 39% for best two matches is reported. Another approach using eigeneyes is presented in [46] to handle several challenges of face recognition including disguise. Using the AR database [34] the algorithm shows an accuracy of around 87.5%; but the AR database [34] does not contain many examples of disguise except for variations in glasses and scarves.

The advantage of the algorithm is that alterations in the facial features excluding eye region does not affect the accuracy.

This chapter deals with the recognition of faces with variations in disguise. A synthetic database of various disguises is prepared to test the performance of the phase based face recognition algorithm described in Chapter 2. Experimental results show that on a disguise face database the proposed algorithm gives a higher accuracy compared to some well known face recognition algorithms when tested on the same database [12], [35], and [52].

4.2 Database for Disguise

There are various ways to alter the facial features or appearance to disguise oneself without considering the other behavioral characteristics of the human body. The changes in the original feature which causes one person to impersonate another individual can be classified into following three categories:

- Appearance
- Feature
- Multiple Variations

The appearance and features of a human face can be changed by changing the hair style (Figure 4.1), glasses (Figure 4.2), wearing a cap/hat (Figure 4.3), beard and mustache (Figure 4.4), altering lips, eyebrows, eyelashes, nose using make-up (Figure 4.5), and texture of the skin by make-up to show aging effects or wrinkles (Figure 4.6). A combination of any of these can also be applied for disguise (Figure 4.7) which is represented as Multiple Variations. The degree of variation of these changes also affects the performance of face recognition. To the best of our knowledge, there is no face database in public domain which consists of these variations and can be used for face recognition. So, we prepared a synthetic face database to capture the variations for disguise using the software Faces v2.0. This database consists of 40 frontal face images of 100 individuals.

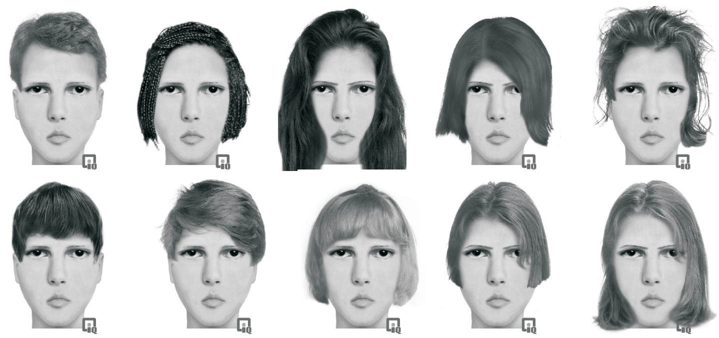

Figure 4.1: Variations in hair style

Figure 4.2: Variations in glasses

Figure 4.3: Variations in cap/hat

Figure 4.4: Variations in beard and mustache

Figure 4.5: Variations in lips/eyebrow/nose

Figure 4.6: Variations in aging and wrinkles

Figure 4.7: Multiple variations

Figure 4.8 shows examples of disguise generated using the software Faces to alter the appearance of a person. The changes in hair can bring a lot of change in the appearance. The hair color can be changed to appear old or young, the hair style can be changed, hairs can be shaved and even wigs can be used to change the appearance frequently. The degree of variation up to which the hairs are changed matters significantly. For example, in the first two images of Figure 4.8, when the degree of variation is low, most of the recognition algorithms can easily and accurately identify the person. However, if the hair style is changed dramatically such that a significant part of the face is occluded, then it might lead to performance degradation. Beards and mustaches can also be used for disguise. These are the features which significantly change the facial appearance, and the performance of most of the algorithms drops even with medium degree of variation.

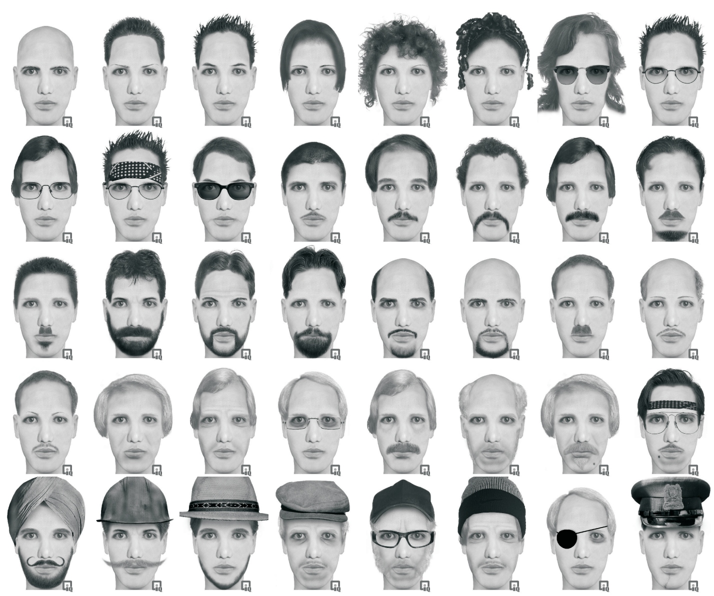

Figure 4.8: Face image of same person with 40 variations

However, most of the recognition algorithms can easily handle low variation in beards and mustaches. Another feature which causes substantial change in appearance is the use of glasses. Use of transparent glass is considered as low variation; but colored or sunglasses is considered to be a high variation because the eyes and its neighboring facial features are hidden. This affects the performance of recognition algorithms. Similarly, the use of a hat or a cap is considered to be a low variation disguise if it does not cover the facial features, but cowboy style hat can be considered to be a medium level of variation in disguise. The changes in lips, eyebrows and eyelashes with the help of make-up do not affect the accuracy of most of the recognition algorithms. However, changes in nose shape by using some artifacts or surgical alterations can cause the performance of feature based recognition algorithms to decrease. Finally, changes in the skin texture to show aging effects or wrinkles also contribute in defying the appearance based face recognition algorithms. A combination of different disguise features as shown in the last few images of Figure 4.8 further exacerbate the problem of recognition.

4.3 Experimental Results

Experiments are carried out on a face database which is created using commercial software Faces v2.0 [59] to generate the effect of face disguise. For the above mentioned disguise scenario, no database is available in public domain. We used Faces v2.0, which is widely used in face image generation in forensic applications, to generate face images of 100 individuals having 40 frontal face images for each class. Among these 40 images, 4 images have minor variations, 22 images have variations in only one feature and rest of the 14 images contain variations in two or more features. Figure 4.8 shows an example of these 40 variations within one class. The phase-based face recognition algorithm described in Chapter 2 is used for feature extraction and matching. The results in Chapter 2 show that the phase features outperform the amplitude features. We have therefore used only the phase features for matching disguised images.

We selected one normal face image from each class as the database image and rest of the images were used as query image. Of the query images, two to three images contain minor variations and are treated as 'No Variation' sample. Rest of the query images are disguise samples. We performed recognition assuming that a person is already enrolled in the system and wants to deceive the system. We calculate the recognition accuracy which measures the number of times the algorithm is able to correctly identify the disguise. More specifically,

Table 4.1: Identification accuracy of the algorithms for different variations

Variation	Identification Accuracy			
	Proposed	PCA [52]	GF [12]	LFA [35]
No Variation	98.2 %	89.3 %	95.8 %	97.1 %
Hair Style	94.9 %	85.7 %	94.1 %	94.8 %
Bear + Mustache	84.6 %	59.1 %	86.2 %	81.3 %
Glasses	85.2 %	70.9 %	55.1 %	84.4 %
Cap/Hat	94.7 %	82.6 %	90.9 %	91.7 %
Lips/Eyebrows/Nose	97.1 %	87.4 %	78.6 %	96.3 %
Aging/Wrinkles	95.4 %	77.5 %	81.8 %	92.9 %
Multiple Variations	71.2 %	19.7 %	49.1 %	61.3 %

system is able to find the correct match if the person is in the database. If a person tries to impersonate another individual, then the system indicates a reject. Table 4.1 shows the identification accuracy of the algorithm for different variations.

We also evaluated the performance of three popular algorithms using the disguise database. The algorithms used for comparison are Eigen face [52], Geometrical Feature (GF), based [12] and Local Feature Analysis (LFA) [35]. Recognition accuracies are calculated in the same scenario as described earlier to compare the performance of the four algorithms. Results of comparison are shown in Table 4.1.

Eigen face algorithm [52] is an appearance based face recognition algorithm. In our experiments we found that this approach is not capable of handling variations that affect the appearance of a person such as medium variation in beard and mustache. With multiple combinations of disguise, the accuracy decreases drastically due to increase in both the false acceptance and the false rejection. Geometrical feature based recognition algorithm [12] uses mixture distances of the facial features for matching. This algorithm works on the distance between geometrical features and correspondences between them and hence works only in cases when this information is preserved. Thus it works well for changes that do not hide or alter the facial features. It does not give correct results for high variation in beards or mustaches, wearing dark eye glasses, and multiple variations in disguise. Local feature analysis [35] uses local features to represent the face and match them. In case of disguise, local feature analysis withstands low variation in features but major changes in face lead to false results. This algorithm can handle minor variations in hair style, light colored eye glasses, medium degree of variation in eyebrow, eyelashes and lips and certain occlusions which do not hide or change prominent features of the face. It does not give good results for textural variations such as aging, high variations in beards or mustaches, hair style which covers facial features, dark eye glasses, and occlusions which cover the local features of the face.

Experimental results for different algorithms vary depending on the features altered. Phase feature based face recognition algorithm is able to handle low to medium variation in phase information such as change in hair style, beard, mustache, transparent glasses, cap/hat, lips, nose, eyebrows and aging effects. We also found that accuracy decreases due to medium to high variations in the face texture. In case of high variations such as disguise with turban, beard, and mustache, the algorithm is unable to find corresponding match in the database. This drops the accuracy due to false rejection, but at the same time none of the faces are found to be falsely matched with any other face. Thus for multiple variations in disguise, accuracy of the proposed algorithm outperforms other algorithms and is found to be 71.2%. Figure 4.9 shows the performance of four face recognition algorithms on the synthetic face database. The images used for this experiment contain only minor variations. We performed this experiment to analyze the performance of the individual algorithms on the synthetic database with minimum variations. For every algorithm, this also provides the relative difference in performance on the images with no variations and images with minor variations.

Figure 4.10 to Figure 4.23 show the feature templates obtained by applying the face recognition algorithm and the ROC plots for each variation. These ROC plots also show the comparison among PCA, Geometrical Features, Local Feature Analysis and the pro-

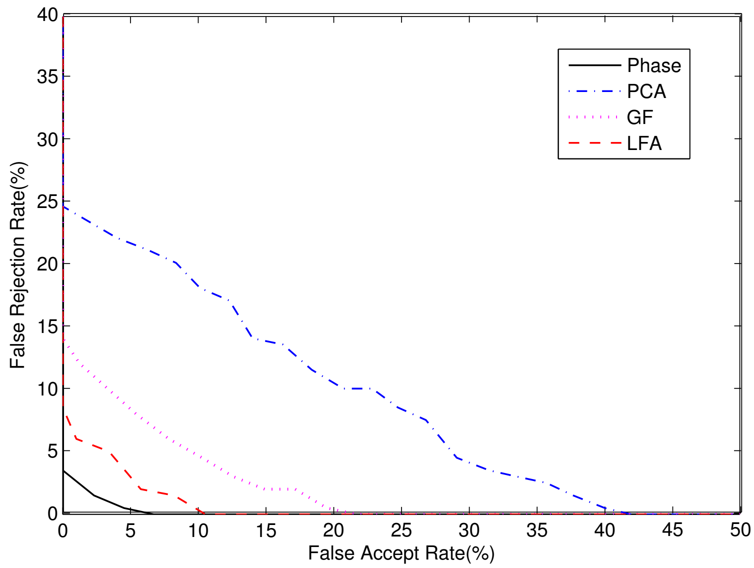

Figure 4.9: ROC for minor variations in the images

posed phase based face recognition algorithm. For variation in hair style, glasses, cap/hat, lips/eyebrow/nose, aging/wrinkles and multiple variations, the proposed texture based face recognition algorithm gives the best performance followed by LFA, and then the geometrical features and PCA. These variations change the appearance of a person and hence affect the performance of the appearance based face recognition algorithms such as PCA. With beard and mustache, geometrical feature based algorithm performs the best followed by the proposed phase based face recognition algorithm.

Figure 4.10: Template for variation in hair style

In this chapter, we have addressed the challenge of face disguise. The proposed phase based face recognition algorithm described in Chapter 2 is applied on a synthetic face data-

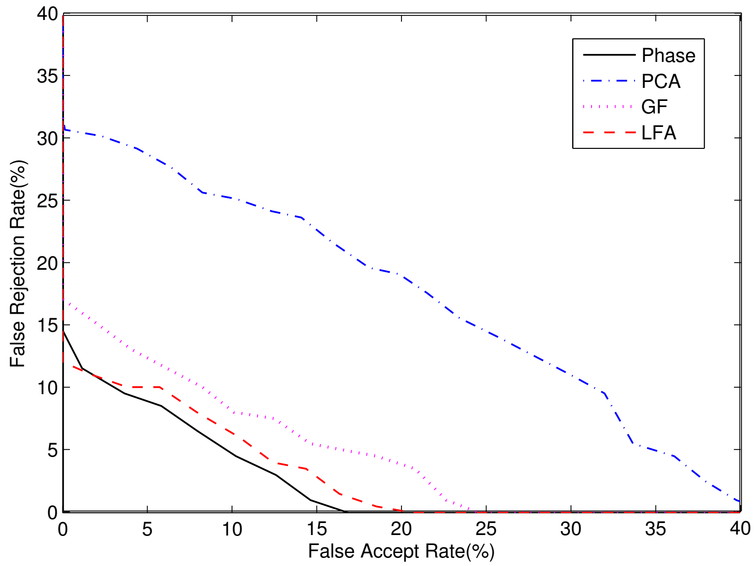

Figure 4.11: ROC for variations in hair Style

Figure 4.12: Template for variation in beard and mustache

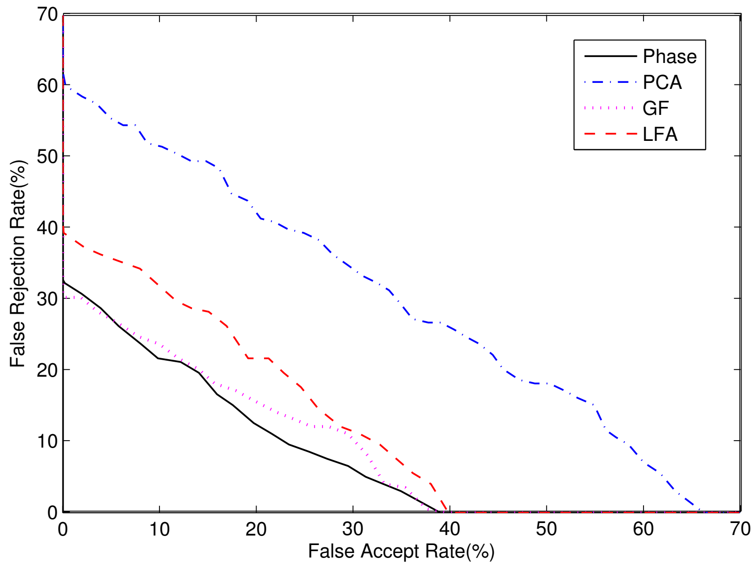

Figure 4.13: ROC for variations in beard and mustache

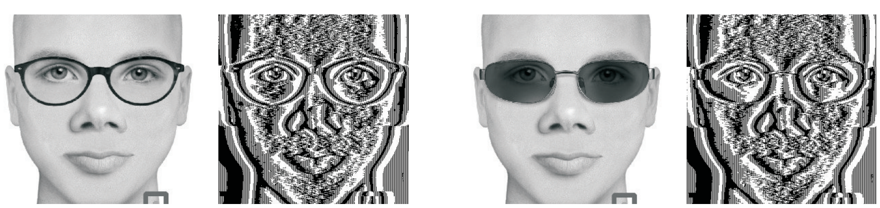

Figure 4.14: Template for variations in glasses

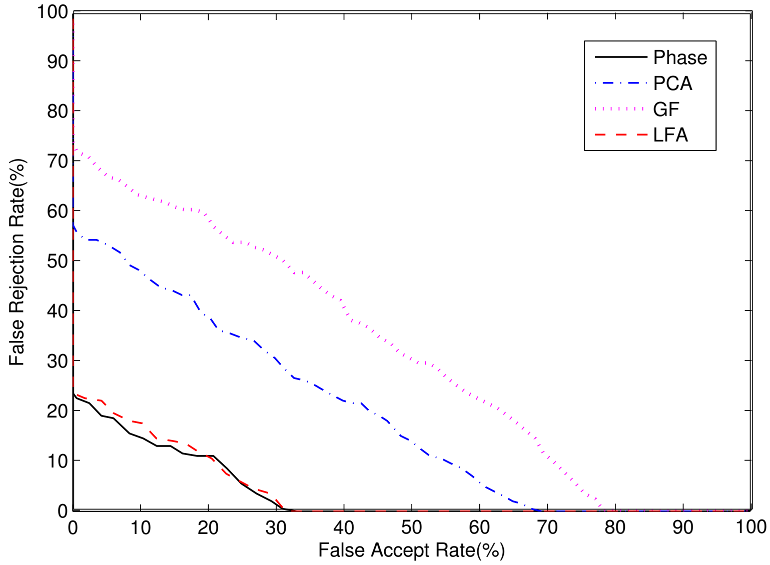

Figure 4.15: ROC for variations in glasses

Figure 4.16: Template for variations in cap/hat

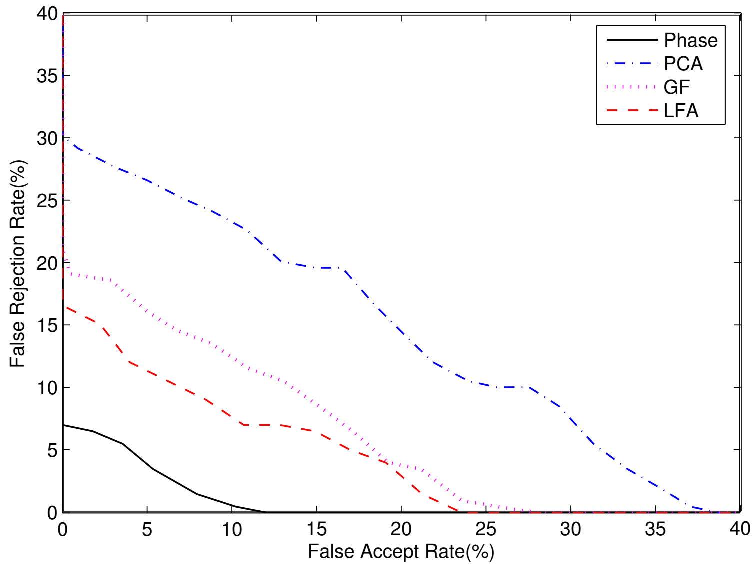

Figure 4.17: ROC for variations in cap/hat

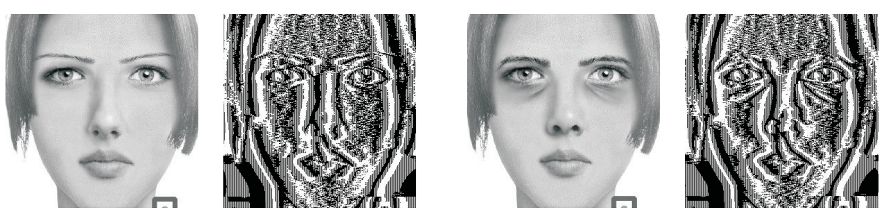

Figure 4.18: Template for variations in lips/eyebrow/nose

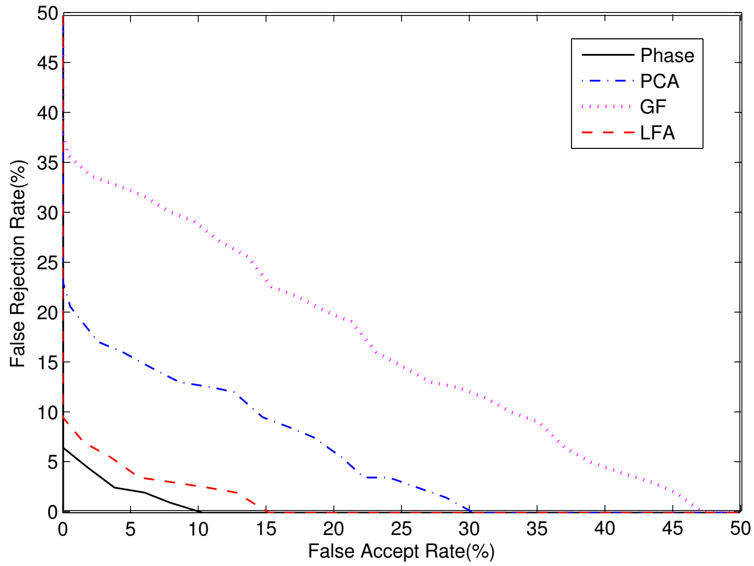

Figure 4.19: ROC for variations in lips/eyebrow/nose

Figure 4.20: Template for variations in aging/wrinkles

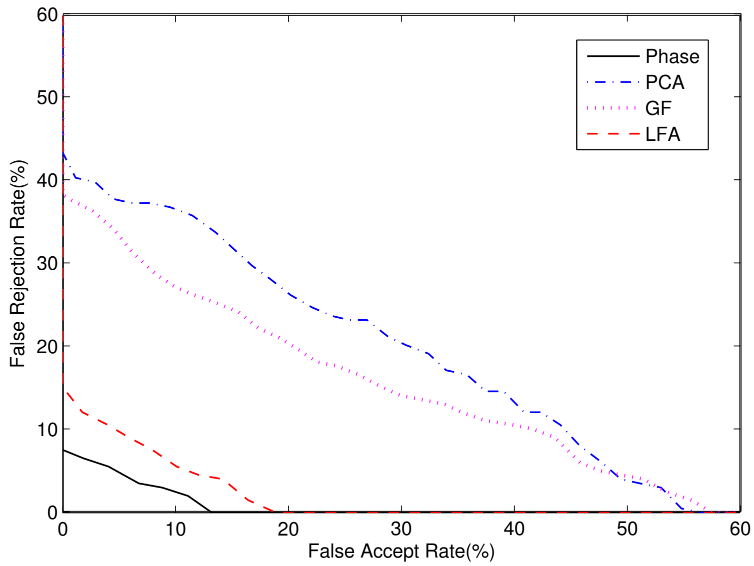

Figure 4.21: ROC for variations in aging/wrinkles

Figure 4.22: Templates for multiple variations

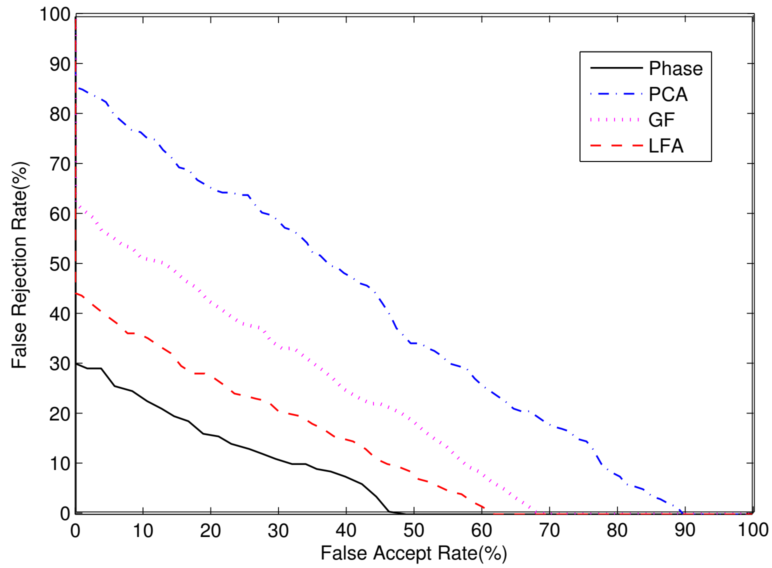

Figure 4.23: ROC for multiple variations

base created for testing the algorithms for disguise. The commercial software Faces v2.0 is used to generate the synthetic faces with several possible variations. The database contains images with variations in appearance, features and multiple variations. Three existing face recognition algorithms namely Eigenface [52], Geometrical Features [12], and Local Feature Analysis [35] are compared with the proposed phase based face recognition algorithm. Experimental results show that the proposed face recognition algorithm performs better than the existing face recognition algorithm in most of the cases.

Chapter 5

Conclusion and Future Work

The problem of recognizing faces under variations still remain largely unsolved. Variation due to illumination, pose, background, environmental constraints, aging, and other factors cause degradation in the performance of face recognition algorithms. Besides these challenges, several other application oriented challenges also lead to degradation in the performance of face recognition systems. One such example is matching scanned image of a missing person with digital image. Another major problem with face recognition algorithms is with intruders who deliberately want to fool the system either by impersonating someone else or disguising their own identity.

In this research work, we have undertaken the problem of face recognition with the approach of simulating human visual cortex. Psychological and neurophysiological researchers have shown that the response of 2D Gabor wavelet is similar to the response of visual cortex of the human mind. Log polar form of the 2D log Gabor makes it invariant to rotation and scaling. Based on these research findings, we use 2D log polar Gabor wavelet to extract facial features in the form of amplitude and phase. Using the proposed face recognition algorithm, we transform the problem of face recognition into an efficient test of statistical dependence operating on amplitude variation and statistical independence operating on phase variation. These two features provide invariance to illumination, contrast, brightness and up to certain extent facial accessories as well. Through experimental validation we establish that phase features give better performance compared to amplitude features.

To match a scanned face image with a digital face image, first the multiscale retinex based normalization and the eigen space projection based transformation algorithms are proposed. We then presented a modification on the proposed face recognition algorithm to focus only on the prominent local regions of a face image. The modified face recognition algorithm is evaluated on a face database of 500 individuals which contains both the scanned and digital face images. An identification accuracy of around 98% is achieved. The proposed algorithm also outperforms the existing algorithms by atleast 5%.

The proposed phase based face recognition algorithm is also validated using a synthetic face database prepared for simulating the disguise scenario. In this simulation, we considered the variations due to change in appearance and features and showed the performance of recognition algorithms. In this experiment we found that the proposed phase-based face recognition algorithm gives the best results in comparison to other existing algorithms.

5.1 Future Work

There are several ways by which the performance of face recognition can be further improved. Face is a 3D object and the image is its 2D representation. If 3D information of the face can be obtained from 2D face image, then the performance can be improved. One way could be to use mosaicing techniques in which multiple variations of a face image are captured and seamlessly stitched together to get a mosaiced face image. The mosaiced image is a derived representation of the 3D image in the form of a 2D image, without reconstituting the complete 3D structure of the face. Face mosaicing is also useful for multiple pose based face recognition.

Further, a 2D face deformation model can be used to model the variations in expression and aging. By studying the deformation patterns and the elasticity of face, a deformation model can be generated to model the different variations of a face due to expression, and aging.

Finally a large real time disguise face database is necessary to test and validate the performance of a face recognition algorithm for non-cooperative users.

References

[1] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," IEEE Transaction on Neural Networks, Vol.13, pp. 1450 - 1464, 2002.

[2] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 711 - 720, 1997.

[3] E. Bigun, J. Bigun, B. Duc, and S. Fischer, "Expert conciliation for multi modal person authentication systems by bayesian statistics," Proceedings of International Conference on Audio- and Video-based Person Authentication, pp. 311 - 318, 1997.

[4] J. Bigun, and J. M. Du Buf, "N-folded symmetries by complex moments in Gabor space and their applications to unsupervised texture segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 16, No. 1, pp. 80 - 87, 1994.

[5] D. Blackburn, M. Bone, and P. J. Phillips, "Face recognition vendor test 2000," Technical Report, 2001.

[6] W. W. Bledsoe, "Man machine facial recognition," Panoramic Research Inc., Palo Alto Cal., Rep. PRI, August 1966.

[7] K. Chang, K. Bowyer, V. Barnabas, and S. Sarkar, "Comparison and combination of ear and face images in appearance based biometrics," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 9, pp. 1160 - 1165, 2003.

[8] S. Chen, D. Zhang, and Z. H. Zhou, "A new face recognition method based on SVD perturbation for single example image per person," Applied Mathematics and Computation, 2005 (http://lamda.nju.edu.cn/datacode/face/singleface.htm).

[9] S. Chen, J. Liu, and Z. H. Zhou, "Making FLDA applicable to face recognition with one sample per person," Pattern Recognition, Vol. 37, No. 7, pp. 1553 - 1555, 2004.

[10] X. Chen, P. J. Flynn, and K. W. Bowyer, "Visible - light and infrared face recognition," ACM Workshop on Multimodal User Authentication, pp. 48 - 55, 2003.

[11] Y. D Cheng, and M. J Tarr, "How to rob a bank and get away with it: recognizing disguised faces," Journal of Vision, Vol. 3, No. 9, pp. 834a, 2003.

[12] I. J. Cox, J. Ghosn, and P. N. Yianilos, "Feature-based face recognition using mixture-distance," Proceedings of the International Conference on Computer Vision and Pattern Recognition, pp. 209 - 216, 1996.

[13] J. G. Daugman, "Two-dimensional spectral analysis of cortical receptive field profiles," Vision Research, Vol. 20, Issue 10, pp. 847 - 856, 1980.

[14] J. G. Daugman "Spatial visual channels in the fourier plane," Vision Research, Vol. 24, Issue 9, pp. 891 - 910, 1984.

[15] J. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," Journal of the Optical Society of America A, Vol. 2, No. 7, 1160 - 1169, 1985.

[16] J. G. Daugman, "Representational issues and local filter models of two-dimensional spatial visual encoding," In: D. Rose and V. G. Dobson (Editors), Models of the Visual Cortex, John Wiley and Sons, pp. 96 - 107, 1985.

[17] J. G. Daugman, "An information-theoretic view of analog representation in striate cortex," In: E. L. Schwartz (editor), Computational Neuroscience, MIT Press, pp. 403 - 423, 1988.

[18] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," Journal of the Optical Society of America, Vol. 4, pp. 2379 - 2394, 1987.

[19] P. J. Flynn, K. W. Bowyer, and P. J. Phillips, "Assessment of time dependency in face recognition: an initial study," Proceedings of the Audio- and Video-Based Biometric Person Authentication, pp.44 - 51, 2003.

[20] D. Gabor, "Theory of communication," Journal of IEE, Vol. 93, pp. 429 - 457, 1946.

[21] R. C. Gonzalez, and R. E. Woods, "Digital image processing," 2nd Edition, Prentice Hall, 2002

[22] V. Govindaraju, D. B. Sher, R. K. Srihari, and S. N. Srihari, "Locating human faces in newspaper photographs," Proceedings of the IEEE, pp. 549 - 554, 1989.

[23] V. Govindaraju, "Locating human faces in photographs," International Journal of Computer Vision, Vol. 19, No. 2, pp. 129 - 146, 1996.

[24] J. Huang; P. C. Yuen, C. W. Sheng, and J. H. Lai, "Component-based LDA method for face recognition with one training sample," Proceedings of the IEEE International Workshop on the Analysis and Modeling of Faces and Gestures, pp. 120 - 126, 2003.

[25] J. Jones, and L. Palmer, "An evaluation of the two dimensional Gabor filter model of simple receptive fields in cat striate cortex," Journal of Neurophysiology, Vol. 58, No. 6, pp. 1233 - 1258, 1987.

[26] T. Kanade, "Picture processing system by computer complex and recognition of human faces," Doctoral Dissertation, Kyoto University, November, 1973.

[27] J. Kittler, G. Matas, K. Jonsson, and M. Sanchez, "Combining evidence in personal identity verification systems," Pattern Recognition Letters, Vol. 18, No. 9, pp. 845 - 852, 1997.

[28] C. Kotropoulos, I. Pitas, S. Fischer, and B. Duc, "Face authentication using morphological dynamic link architecture," Proceedings of the First International Conference on Audio- and Video-Based Biometric Person Authentication, Lecture Notes In Computer Science, Vol. 1206, pp. 169 - 176, 1997.

[29] A. Kumar, D. C. M. Wong, H. Shen, and A. K. Jain, "Personal verification using palmprint and hand geometry biometric," Proceedings of the International Conference on Audio- and Video-based Person Authentication, pp. 668 - 675, 2003.

[30] E. H. Land, and J. J. McCann, "Lightness and retinex theory," Journal of the Optical Society of America, Vol. 61, No. 1, pp. 1 - 11, 1971.

[31] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: a convolutional neural-network approach," IEEE Transaction on Neural Network, Vol. 8, pp. 98 - 113, 1997.

[32] C. Liu, and H. Wechsler, "Face recognition using independent Gabor wavelet features source," Proceedings of the Third International Conference on Audio- and Video-Based Biometric Person Authentication, Lecture Notes In Computer Science, Vol. 2091, pp. 20 - 25, 2001.

[33] B. MacLennan, "Gabor representations of spatiotemporal visual images," University of Tennessee, Knoxville, Technical Report CS-91 - 144, 1991.

[34] A. M. Martinez, and R. Benavente, "The AR face database," CVC Tech. Report #24, 1998.

[35] P. Penev, and J. Atick, "Local feature analysis: A general statistical theory for object representation" Network: Computation in Neural Systems, Vol. 7, pp. 477 - 500.

[36] A. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eigenspaces for face recognition," Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 84 - 91, 1994.

[37] P. J. Philips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face recognition algorithms," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 10, pp. 1090 - 1104, 2000.

[38] P. J. Philips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and J. M. Bone, "FRVT 2002: evaluation report," 2003.

[39] Z. Rahman, D. J. Jobson, and G. A. Woodell, "Multiscale retinex for color image enhancement," Proceedings of IEEE International Conference on Image Processing, Vol. 3, pp. 1003 - 1006, 1996.

[40] N. Ramanathan, A. R. Chowdhury, and R. Chellappa, "Facial similarity across age, disguise, illumination and pose," Proceedings of International Conference on Image Processing, Vol. 3, 1999 - 2002, 2004.

[41] A. Ross, and A. K. Jain, "Information Fusion in Biometrics," Pattern Recognition Letters, Vol. 24, No. 13, pp. 2115 - 2125, 2003.

[42] A. Ross, and R. Govindarajan, "Feature level fusion using hand and face biometrics," Proceedings of the SPIE Conference on Biometric Technology for Human Identification, Vol. II, pp. 196 - 204, 2005.

[43] F. Samaria, and S. Young, "HMM based architecture for face identification," Image and Vision Computing, Vol. 12, pp. 537 - 583, 1994.

[44] M. Savvides, B. V. K. Vijaya Kumar, and P. Khosla, "Face verification using correlation filters," Proceedings of the Third IEEE Automatic Identification Advanced Technologies, pp. 56 - 61, 2002.

[45] F. Shen, and H. Wang, "A local edge detector used for finding corners," Proceedings of the Third International Conference on Information, Communications and Signal Processing, 2001.

[46] P. Q. Silva, and A. N. C. Santa Rosa, "Face recognition based on eigeneyes," Pattern Recognition and Image Analysis, Vol. 13, No. 2, pp. 335 - 338, 2003.

[47] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression database," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 12, pp. 1615 - 1618, 2003.

[48] R. Singh, M. Vatsa, and A. Noore, "Textural feature based face recognition for single training images," IEE Electronics Letters, Vol. 41, No. 11, pp. 23 - 24, 2005.

[49] S. K. Singh, D. S. Chauhan, M. Vatsa, and R.Singh, "A robust skin color based face detection algorithm," Tamkang Journal of Science and Engineering, Vol. 6, No. 4, pp. 227 - 234, 2003.

[50] V. Starovoitov, D. D. Samal, and B. Sankur, "Matching of faces in camera images and document photographs," Proceedings of the IEEE International Conference on Acoustic Speech and Signal Processing, Vol. 4, pp. 2349 - 2352, 2000.

[51] X. Tang, and X. Wang, "Face sketch recognition," IEEE Transactions on Circuits, Systems and Video Technology, Vol. 14, pp. 50 - 57, 2004.

[52] M. Turk, and A. Pentland, "Eigenfaces for recognition" Journal of Cognitive Neuroscience, Vol. 3, pp. 72 - 86, 1991.

[53] S. Venkatesh, and R. A. Owens, "Energy feature detection scheme," Proceedings of the International Conference on Image Processing, pp. 553 - 557, 1989.

[54] M. Webster, and R. de Valois, "Relationship between spatial-frequency and orientation tuning of striate-cortex cells," Journal of the Optical Society of America, Vol. 2, pp. 1124 - 1132, 1985.

[55] L. Wiskott, J. M. Fellous, N. Krger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 775 - 779, 1997.

[56] M. H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: a survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24 No. 1, pp. 34 - 58, 2002.

[57] D. Zhang, S. Chen and Z.-H. Zhou, "Enhanced $(PC)^2A$ for face recognition with one training image per person," Pattern Recognition Letters, Vol. 25, No. 10, pp. 1173 - 1181, 2004.

[58] W. Zhao, R. Chellappa, P. Phillips, and A. Rosenfeld, "Face recognition: a literature survey," ACM Computing Survey, Vol. 35, No. 4, pp. 399 - 458, 2003.

[59] http://www.iqbiometrix.com/products_faces_40.html

[60] http://amp.ece.cmu.edu/projects/FaceAuthentication/download.htm

[61] http://www.equinoxsensors.com/products/HID.html