Iris Recognition System for Some Clinical Applications

A Thesis

Submitted in the fulfillment of the requirements for the award of the degree of

Doctor of Philosophy

Submitted by

Atul Bansal
(Registration No. 950906023)

Under the supervision of

Dr. R. K. Sharma
Professor
Thapar University
Patiala

Dr. Ravinder Agarwal
Professor
Thapar University
Patiala

Thapar University, Patiala

Department of Electronics & Communication Engineering
Thapar University
Patiala-147004 (Punjab) India

April 2015

CERTIFICATE

I, **Atul Bansal,** hereby certify that the work which is being presented in this thesis entitled, "**IRIS RECOGNITION SYSTEM FOR SOME CLINICAL APPLICATIONS**", in fulfillment of requirements for the award of degree of **DOCTOR OF PHILOSOPHY** submitted in Electronics & Communication Engineering Department (ECED), Thapar University, Patiala, is an authentic record of my own work carried out under the supervision of **Dr. Ravinder Agarwal** (Professor, Department of Electrical & Instrumentation Engineering, Thapar University, Patiala) and **Dr. R. K. Sharma** (Professor, Department of Computer Science & Engineering, Thapar University, Patiala).

The matter presented in this thesis has not been submitted either in part or in full to any other University or Institute for the award of any degree.

Atul Bansal

(ATUL BANSAL)

Dated: October 23, 2015

Certified that the above statement made by the candidate is correct to the best of our knowledge.

R K Sharma 23.10.2015

(Dr. R. K. Sharma)
Professor,
Department of Computer Science & Engineering,
Thapar University,
Patiala-147004 (INDIA)

Regard

(Dr. Ravinder Agarwal)
Professor,
Department of Electrical & Instrumentation Engineering,
Thapar University,
Patiala-147004 (INDIA)

Dated: October 23, 2015

i

ACKNOWLEDGEMENT

Before I express my heart-felt gratitude towards all my mentors, first and foremost, I would like to thank my greatest teachers: **God and my Parents**, who have given me the power to believe in myself and my passion to pursue this study. I could never have done this without their blessings.

I would like to express my gratitude to my supervisor, **Prof. Ravinder Agarwal** for motivation, patient guidance and every-time support. I am truly very fortunate to have the opportunity to work with him. I found his guidance to be extremely valuable. I am thankful to him that he has full confidence in my ability; this work would have not been completed without his vision and encouragement.

I would like to express my warm and sincere gratitude to **Prof. R. K. Sharma**, my supervisor who always guided me as a friend. He always supported me throughout the work by motivating me from time to time. It would have been very difficult for me to come out with results without his invaluable suggestion.

I would like to express my deepest gratitude to **Mr. N. D. Agarwal**, Chancellor, GLA University, Mathura, **Prof. Jai Prakash,** Ex-VC GLA University, Mathura, for giving me opportunity to pursue my Ph.D. I would also like to thank my all-fellow colleagues and staff members of Department of Electronics & Communication Engineering, GLA University, Mathura for their encouragement and support.

My sincere thanks to **Dr. Arun Bansal**, **Dr. Poonam Agarwal** and all the subjects who helped in developing the database.

I am highly in debt to my wife **Anjali** and daughter **Ishita,** without whose unbiased sacrifice and support this work would not have been completed.

Although I have tried to express my gratitude to every person who contributed in this work directly or indirectly, there may still be someone hiding behind the veils of my forgetful part of memory. Last but not the least; I would like to thank all such souls.

(ATUL BANSAL)

SUMMARY

Today, with the increase in security threats all over the world authentication of an individual is becoming an important issue and area of interest for researchers. Over the traditional password or key based security systems biometric authentication systems are considered as very accurate and reliable. Iris Recognition System is one of them. It is very accurate system as iris images of twins or iris images of even left and right eye of same person are different. Numerous researchers have given iris recognition systems based on different feature extraction techniques. In this work, a comparative study of the existing techniques has been carried out. A simple, fast and effective statistical feature extraction based iris recognition system has been proposed and implemented. Features have been extracted in two different directions, namely, radial direction and angular direction. An attempt has been made to study the effect of number of features as well as the radial and angular resolution while normalization. Results obtained are effective, encouraging and comparable to existing techniques.

In literature, little work has been reported on clinical applications of iris recognition systems. In this thesis, clinical applications of iris recognition system have also been investigated. Three different applications, *i.e.*, to predict the gender of imposters, to predict diabetes and to predict obstructive lung disease have been considered. In security systems predicting gender of an imposter is equally important to determine the identity. Most of the work to predict the gender utilized facial images. A few studies have been reported using iris images. In the present research work, Support Vector Machine (SVM) based gender prediction model has been proposed and implemented. Results obtained show the effectiveness of system over the existing models.

Further, a non-invasive and non-contact type model, *i.e.*, a system as an aid to doctors is proposed to predict the disease from iris images. Iridology is a science to determine the status of health of an individual from iris patterns. It does not diagnose the disease. If it is known that which part of human body is not well then by doing regular exercises and following healthy habits one can delay if not prevent the occurrence of disease.

In this thesis, iris recognition algorithm has been combined with iridology to predict diabetes and obstructive lung disease.

A dataset of healthy subjects and also the subjects suffering from diabetes has been created for implementing the problem of diabetes prediction. Another dataset of healthy subjects and also the subjects with obstructive lung disease has been created. These datasets have been built with the help of I-SCAN-2 dual iris scanner of Cross Match Technologies Inc. In order to develop the models for the two clinical applications, features have been extracted using two different techniques: wavelet transform and Gabor filter. SVM based classifier has been implemented to predict the subject as healthy or suffering from diabetes or obstructive lung disease. The maximum overall accuracy of 88.3% and 89.0% has been achieved for predicting diabetes and obstructive lung disease, respectively. This accuracy has been achieved for a test data of size 80 and 100, respectively for the two applications. The proposed models are compact and handy.

CONTENTS

CERTIFICATE **i**

ACKNOWLEDGEMENT **ii**

SUMMARY **iii**

LIST OF FIGURES **viii**

LIST OF TABLES **xi**

ACRONYMS **xii**

CHAPTER 1 INTRODUCTION 1-8

1.1 Biometric Systems 1

1.2 Iris Recognition System 2

1.3 Iridology 4

1.4 Motivation of the Proposed Work 6

1.5 Objectives 7

1.6 Organization of Thesis 7

CHAPTER 2 LITERATURE SURVEY 9-19

2.1 Studies Related to Iris Recognition System 9

2.1.1 Image preprocessing 9

2.1.2 Feature extraction 15

2.1.3 Template matching 16

2.2 Studies Related to Clinical Applications 17

CHAPTER 3 IRIS RECOGNITION SYSTEM 20-48

3.1 Overview 20

3.2 Implementation of Existing Techniques 20

3.2.1 Feature extraction without normalization 21

3.2.2 Feature extraction with normalization 23

3.3 Comparative Analysis 24

3.3.1 FAR and FRR 24

3.3.2 Complexity 28

3.3.3 Memory requirement 29

3.4 Statistical Feature Extraction Based Iris Recognition System 29

3.4.1 Proposed feature extraction technique 30

3.4.2 Results and discussions 32

3.4.2.1 Feature extraction along concentric circles 33

(i) Experimentation with three statistical parameters 34

(ii) Experimentation with six statistical parameters 37

3.4.2.2 Feature extraction along angular direction 40

(i) Experimentation with three statistical parameters 40

(ii) Experimentation with six statistical parameters 43

CHAPTER 4 IRIS RECOGNITION SYSTEM FOR GENDER AND DIABETES PREDICTION 49-70

4.1 Overview 49

4.2 Predicting Gender Using Iris Images 51

4.2.1 Using statistical features and wavelet transform 52

4.2.2 Using Otsu's threshold method 55

4.3 Dataset for Predicting Gender 55

4.4 Results of Gender Prediction 55

4.5 Predicting Diabetes Using Iris Images 59

4.5.1 Using wavelet transform 61

4.5.2 Using Gabor filter 61

4.6 Dataset for Predicting Diabetes 62

4.7 Results of Diabetes Prediction 64

4.7.1 Results using wavelet transform 65

4.7.2 Results using Gabor filter 68

4.7.3 Comparison of proposed models with existing techniques 68

CHAPTER 5 PREDICTING OBSTRUCTIVE LUNG DISEASE USING IRIS IMAGES 71-78

5.1 Overview 71

5.2 Feature Extraction 72

5.3 Dataset 73

5.4 Results and Discussions 73

5.4.1 Results using wavelet transform 74

5.4.2 Results using Gabor filter 76

CHAPTER 6 CONCLUSION AND FUTURE SCOPE OF WORK 79-82

6.1 Conclusion 79

6.2 Future Scope of Work 81

LIST OF PUBLICATIONS 83

REFERENCES 84-92

LIST OF FIGURES

Figure	Caption	Page
1.1	A human eye	2
1.2	Different stages of iris recognition system	3
1.3	Iridology chart for right eye	5
1.4	Iridology chart for left eye	6
2.1	Homogeneous rubber sheet model	13
2.2	Iris image preprocessing	14
3.1	Concentric circles on iris region	21
3.2	FAR and FRR v/s Hamming distance for features extracted with normalization using *Database A*	25
3.3	FAR and FRR v/s Hamming distance for features extracted with normalization using *Database B*	25
3.4	FAR and FRR v/s Hamming distance for features extracted without normalization with 15 circles using *Database A*	26
3.5	FAR and FRR v/s Hamming distance for features extracted without normalization with 15 circles using *Database B*	26
3.6	FAR and FRR v/s Hamming distance for features extracted without normalization with 30 circles using *Database A*	27
3.7	FAR and FRR v/s Hamming distance for features extracted without normalization with 30 circles using *Database B*	27
3.8	EER for features extracted with and without normalization	28
3.9	Division of normalized iris image into horizontal and vertical groups	29
3.10	Eye images where segmentation failed	33
3.11	FAR and FRR as a function of matching threshold with variations in radial resolution for *Database A* using three parameters	34
3.12	FAR and FRR as a function of matching threshold with variations in radial resolution for *Database B* using three parameters	35
3.13	DET graph with variations in radial resolution for *Database A* using three parameters	35
3.14	DET graph with variations in radial resolution for *Database B*	36

using three parameters
3.15 FAR and FRR as a function of matching threshold with variations in radial resolution for *Database A* using six parameters 37
3.16 FAR and FRR as a function of matching threshold with variations in radial resolution for *Database B* using six parameters 38
3.17 DET graph with variations in radial resolution for *Database A* using six parameters 38
3.18 DET graph with variations in radial resolution for *Database B* using six parameters 39
3.19 FAR and FRR as a function of matching threshold with variations in angular resolution for *Database A* using three parameters 41
3.20 FAR and FRR as a function of matching threshold with variations in angular resolution for *Database B* using three parameters 41
3.21 DET graph with variations in angular resolution for *Database A* using three parameters 42
3.22 DET graph with variations in angular resolution for *Database B* using three parameters 42
3.23 FAR and FRR as a function of matching threshold with variations in angular resolution for *Database A* using six parameters 43
3.24 FAR and FRR as a function of matching threshold with variations in angular resolution for *Database B* using six parameters 44
3.25 DET graph with variations in angular resolution for *Database A* using six parameters 44
3.26 DET graph with variations in angular resolution for *Database B* using six parameters 45
4.1 Five different stages of proposed model 50
4.2 I-SCAN-2 dual iris scanner 50
4.3 Three level pyramidal decomposition 54
4.4 Two level packet decomposition 54
4.5 Accuracy of proposed gender prediction model 58
4.6 Segmenting ROI 60
4.7 Distribution of dataset 62
4.8 Age wise distribution of diabetic dataset 63

4.9	Duration of diabetes wise distribution of dataset	63
4.10	Level of diabetes control wise distribution of dataset	64
4.11	Accuracy of proposed diabetes prediction model	66
5.1	Distribution of smokers and non-smokers in subjects identified as healthy and the subjects identified as having obstructive lung disease using wavelet features	76
5.2	Distribution of smokers and non-smokers in subjects identified as healthy and the subjects identified as having obstructive lung disease using Gabor filter	77

LIST OF TABLES

Table	Title	Page
2.1	Different feature extraction algorithms for iris diagnosis system	18
3.1	Successfully segmented images	33
3.2	EER for three parameters with variations in radial resolution	36
3.3	EER for six parameters with variations in radial resolution	39
3.4	EER for three parameters with variations in angular resolution	43
3.5	EER for six parameters with variations in angular resolution	45
3.6	Comparison of proposed approach with existing techniques	47
4.1	Accuracy for different feature vectors with polynomial kernel function	56
4.2	Accuracy for different feature vectors with Gaussian kernel function	57
4.3	Accuracy for different feature vectors with RBF kernel function	57
4.4	Comparison of accuracy of gender prediction model with existing techniques	59
4.5	Overall accuracy of diabetes prediction model for different kernel functions with different feature vectors	65
4.6	Overall accuracy with different feature vectors and kernel functions for predicting diabetes	67
4.7	Comparison of accuracy of proposed diabetes prediction model with existing techniques	69
5.1	Information on the dataset	73
5.2	Overall accuracy with different feature vectors and kernel functions for predicting obstructive lung disease	75
5.3	Accuracy using wavelet features for subjects with different $\boldsymbol{\rho}$	76
5.4	Accuracy using Gabor filter for subjects with different $\boldsymbol{\rho}$	77
5.5	Comparison of accuracy of proposed obstructive lung disease prediction models with existing techniques	78

ACRONYMS

CASIA	Chinese Academy of Sciences – Institution of Automation
CCR	Correct Classification Rate
CHT	Circular Hough Transform
CT	Computer Tomography
DET	Detection Error Tradeoff
DF	Dynamic Features
DWT	Discrete Wavelet Transform
ECG	Electrocardiography
EEG	Electroencephalogram
EER	Equal Error Rate
FAR	False Acceptance Rate
FEV_1	Forced Expiratory Volume in 1 sec
FRR	False Rejection Rate
FVC	Forced Vital Capacity
GLCM	Gray Level Co-occurrence Matrix
IITD	Indian Institute of Technology, Delhi
IDDM	Insulin Dependent Diabetes Mellitus
LBP	Local Binary Pattern
RANSAC	Random Sample Consensus Algorithm
RBF	Radial Basis Function
ROI	Region of Interest
STM	Support Tensor Machine
SVM	Support Vector Machine
THFB	Triplet Half-Band Filter Bank
WED	Weighted Euclidean Distance

CHAPTER 1

INTRODUCTION

1.1 Biometric Systems

Authentication of persons using computer vision and related techniques has always been a very interesting subject of research. Over the traditional password or key-based security systems, biometric systems are considered accurate and reliable.

A biometric system provides identification about an individual based on unique features or characteristics possessed by the individual. A good number of identification systems based on behavioral characteristics such as voice (Almaadeed *et al.*, 2015; Senoussaoui *et al.*, 2014; Sino e*t al.*, 2015; Yang *et al.*, 2014), signature (Angadi and Gaur, 2014; Jayasekara *et al.*, 2006), handwriting (Xiang *et al.*, 2010), speech (Wang *et al.*, 2014; Wang *et al.*, 2015), keystroke (Giot *et al.*, 2010) and physical characteristics including face (Giot *et al.*, 2010), finger print (Cao *et al.*, 2014; Dubey *et al.*, 2007; Kumar *et al.*, 2014; Rajan and Indu, 2014) and iris (Bodade and Talbar, 2009; Boles and Boashash, 1998; Coasta and Gonzaga, 2012; Cui *et al.*, 2004; Daugman 1993; 2004) are being employed for identification of an individual.

In general, biometric systems first acquire the sample of features, such as digital image of face, a voice sample, fingerprint or an iris image. Further, by applying some mathematical transformations these features are transformed into biometric template. These templates are compared with other templates to determine the user as authentic or imposter.

To make an efficient biometric system, features selected should be unique, stable and be easily captured. In this thesis, iris recognition system and its clinical applications have been discussed.

1.2 Iris Recognition System

Iris recognition is a method of biometric authentication that uses pattern recognition techniques based on images of the iris of an individual's eye. Among various biometric systems, iris recognition is considered the most accurate and reliable biometric identification system. Iris recognition system finds applications in various security systems such as those systems at airports, in confidential sections at laboratories, ATM machines *etc*. Various researchers have proposed number of algorithms based on different feature extraction techniques. Iris recognition gives distinct phase information, which spans more than 200 degrees of freedom that makes it most accurate and reliable biometric identification system (Daugman, 1993; 2004).

Iris is a thin circular diaphragm, which lies between the cornea and lens of the human eye as shown in Figure 1.1. Iris is perforated close to its centre by a circular aperture known as pupil. The most important part of iris recognition is the identification of radius and centre of circular iris and pupil.

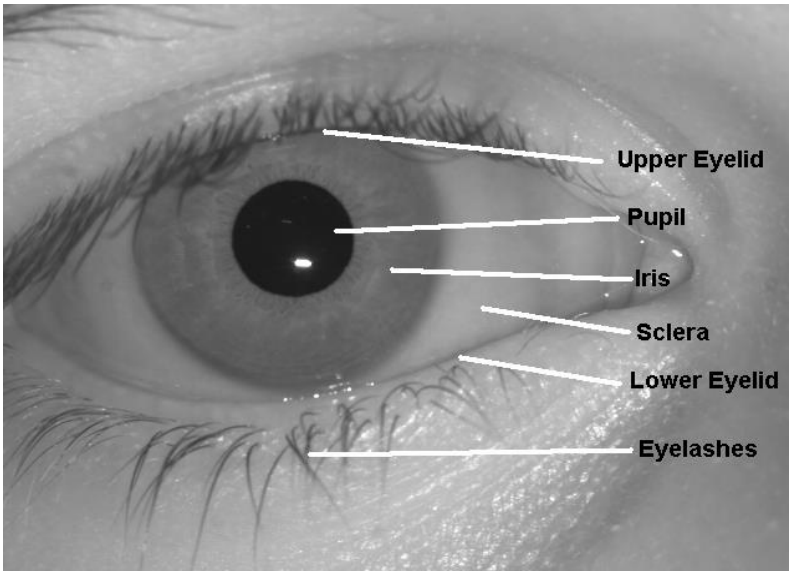


Figure 1.1: A human eye

There are four important stages of an iris recognition system as shown in Figure 1.2. These stages are discussed below, in brief.

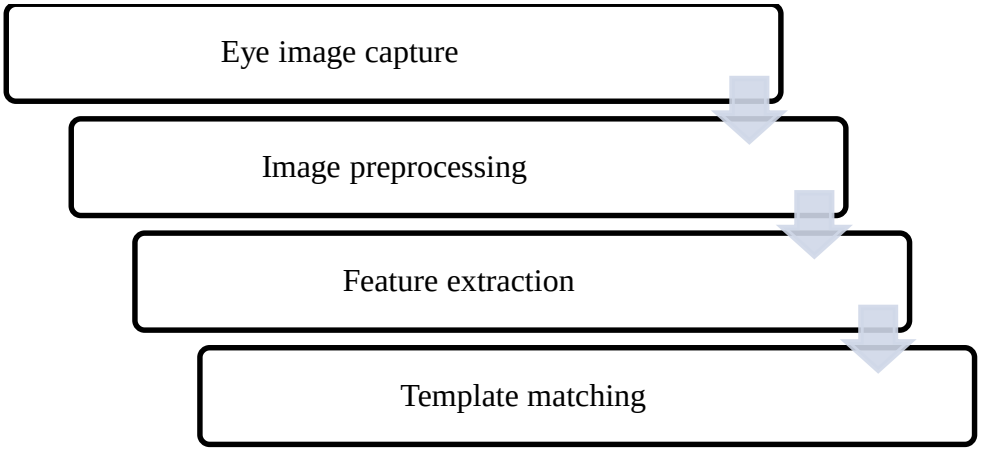


Figure 1.2: Different stages of iris recognition system

- **Eye image capture**

  The first and utmost important step of iris recognition system is capturing the eye image. Infrared sensor based cameras or slit lamps are used to acquire the eye images. Standard eye image datasets are available from Indian Institute of Technology, Delhi (IITD) (*IIT Delhi Iris Database Version 1.0*, http://web.iitd.ac.in/~biometrics/Database_Iris.htm) and Chinese Academy of Sciences – Institution of Automation (CASIA) (*CASIA Iris Image Database*, http://www.sinobiometrics.com/Databases.htm) for research purposes.

  In this thesis, *IIT Delhi Iris Database Version 1.0* is referred as *Database A* and *CASIA Iris Image Dataset-V3* as *Database B*. *Database A* consists of 2240 images acquired using JIRIS, JPC1000 digital CMOS camera in the indoor environment. The resolution of these images is 320 × 240 pixels. *Database B* consists of 2639 images acquired in the indoor environment using the camera developed at CASIA. The resolution of these images is 320 × 280.

- **Image preprocessing**

  Preprocessing refers to convert the image of eye into a form from which the desired features can be extracted and used for identification of an individual. Image preprocessing is divided into three steps - iris localization, iris normalization and image enhancement. Iris localization means to detect the inner and outer boundaries of iris, to find and remove the eyelashes of eyelids that might have covered the iris region. Iris normalization is performed to

convert the iris image from Cartesian coordinates to polar coordinates. Normalized iris image is a rectangular image with angular and radial resolutions. Normalization helps in removing the dimensional inconsistencies that arise due to variation in illumination, camera distance, angle *etc*. while capturing the image of an eye. Now, the obtained normalized image is enhanced to compensate for the low contrast, poor light source and position of light source.

- **Feature extraction**

  Feature extraction is the next important step after preprocessing. The normalized image is used to extract significant features from iris image by applying suitable transformations. These features are further encoded to make the comparisons between templates more effective.

- **Template matching**

  Recognition process is carried out using template matching. In template matching, user iris template is compared with the templates from the database using matching metric. The matching metric gives different range of values when a given iris template is compared with the other stored templates. Based upon these range of values, a decision is taken about the identity of a person, *i.e.*, the person is who they claim to be?

1.3 Iridology

Iridology, also known as iris diagnosis (Jensen, 1985; Stearn and Swanepoel, 2007), is an alternate medicine technique through which colors, patterns and characteristics of iris can be used to gather the information regarding the status of health of an individual. Iridology does not diagnose the disease.

Iris is like a map of the body. Iridology chart (http://www.bernardjensen.com/Charts_c_17.html) divides iris into mainly seven zones. Zones are further divided into finite segments where each segment represents the different organs of the body. Changes in condition of any organ of the body can be analyzed through the changes produced in these segments. It is important to extract significant features to analyze the region associated with particular organ.

Iridology practitioners compare the markings, at specific location of iris, of a subject with that of a healthy human. Any change indicates weather the subject is healthy or suffering from disease.

The history of iridology goes back to 19th century when Hungarian physician Ignatz Von Peczely noticed similar dark streaks in the iris of an owl with broken leg as he noticed in the iris of a man with broken leg. He also observed the dark streaks to grew paler and disappear as the man's leg healed. Later 1950 onwards, iridology was taken forward by Dr. Bernard Jensen, an American naturopath. He proposed the iridology charts (http://www.bernardjensen.com/Charts_c_17.html) for the researchers and practitioners. These charts are depicted in Figures 1.3 and 1.4.

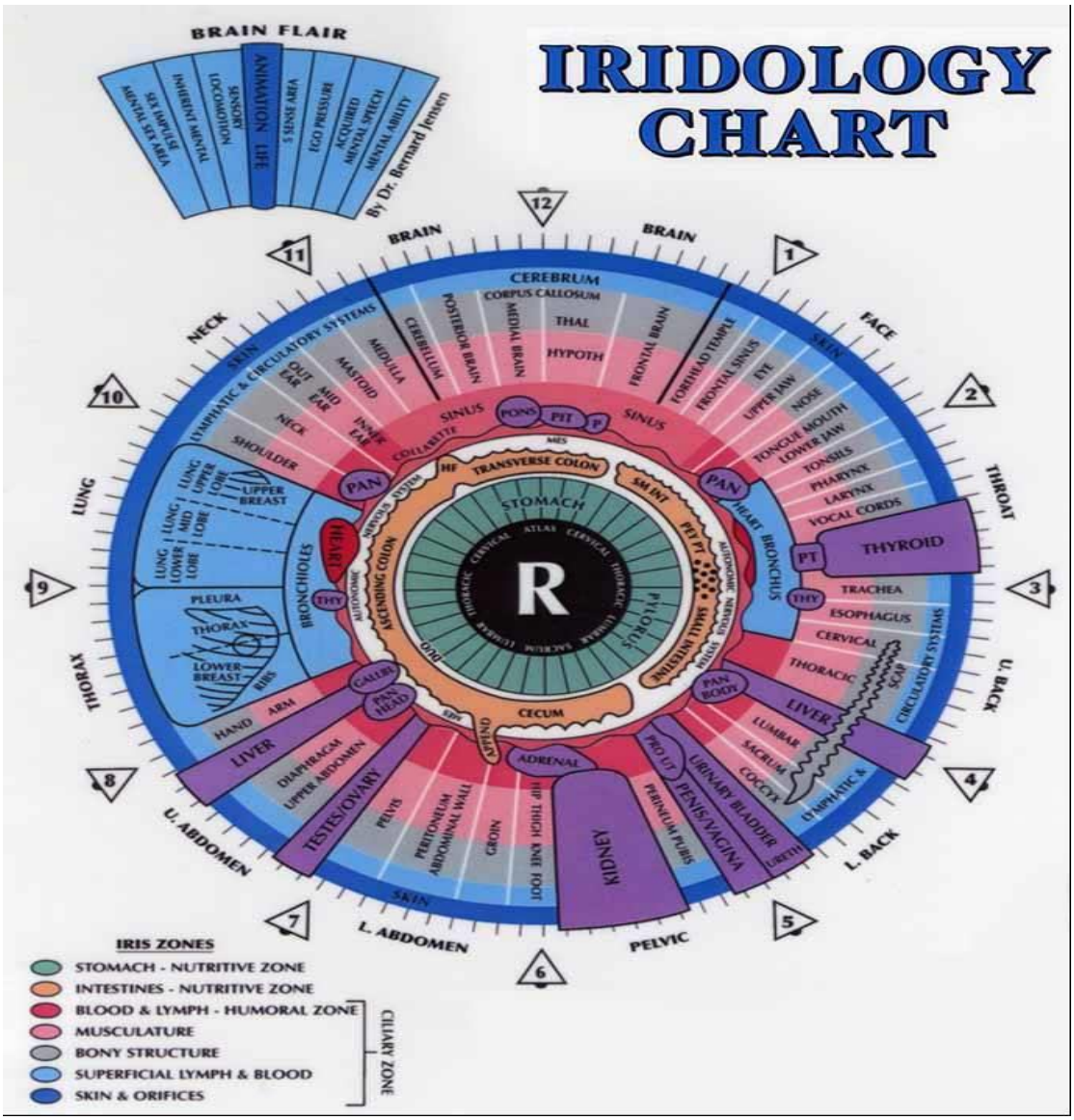


Figure 1.3: Iridology chart for right eye

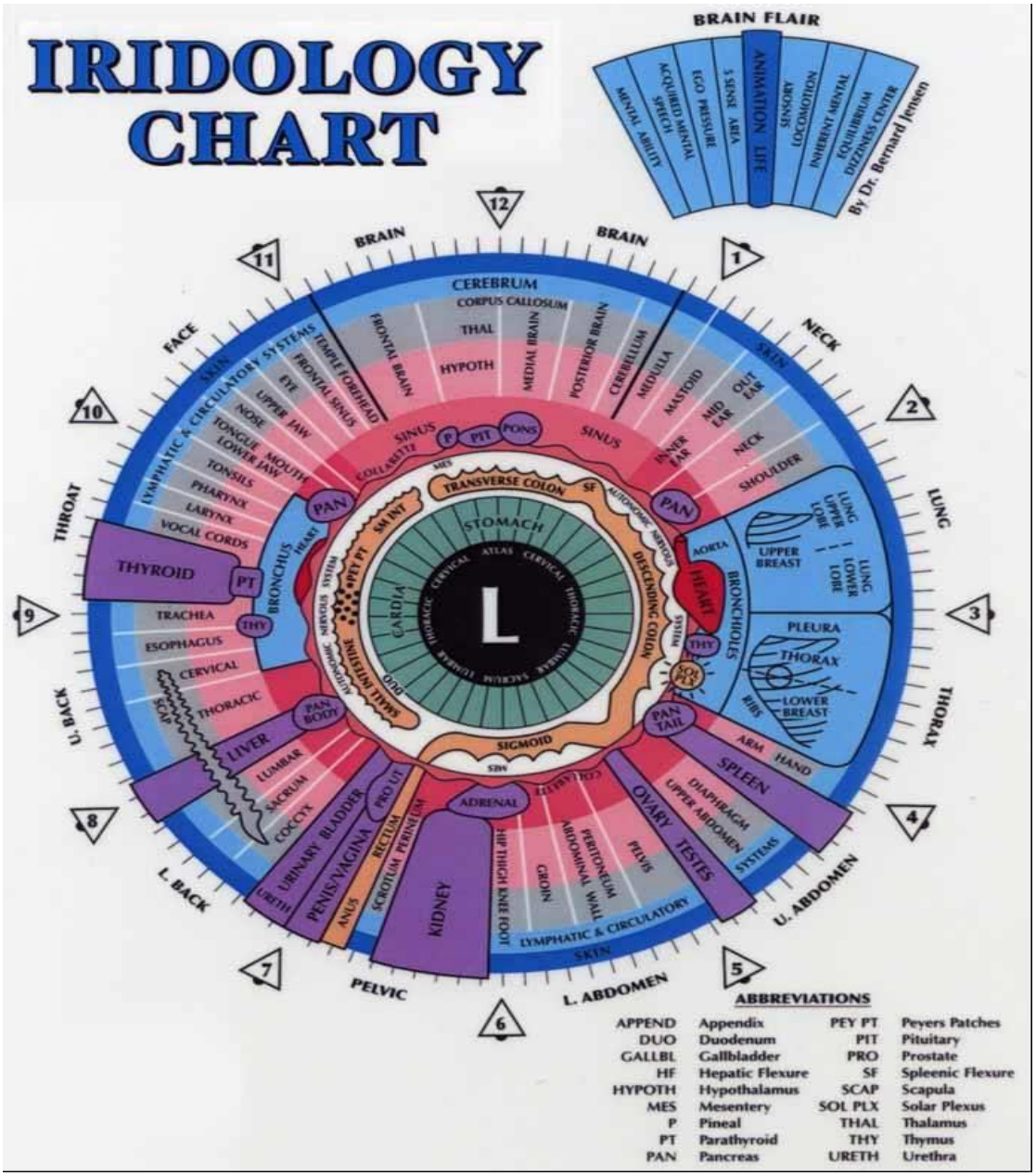


Figure 1.4: Iridology chart for left eye

1.4 Motivation of the Proposed Work

Researchers have proposed a number of iris recognition systems based upon different feature extraction techniques (Bodade and Talbar, 2009; Boles and Boashash, 1998; Daugman 1993; 2004; Wildes *et al.*, 1994; Wildes, 1997). The following gaps in the iris recognition systems have been identified. We have worked on one of these gaps in the proposed piece of work.

i. Segmentation of iris and pupil area from the human eye leads to the loss of significant information.

ii. Iris recognition systems have different accuracy for iris patterns of the people from different region as they have some different features in their iris patterns.

iii. Iris recognition systems have been mainly used for human identification and security purposes. These systems have not been used for clinical applications to predict the diseases. An effort has been made in this study for proposing a model in order to predict the diseases based on iris patterns.

1.5 Objectives

The University research board approved the following objectives for this work.

i. To implement the existing algorithms for iris recognition system.
ii. To propose new features, for iris recognition.
iii. To design and implement algorithm for the feature encoding and template matching based upon Support Vector Machine (SVM) for iris recognition and to establish its efficiency over the existing algorithms.
iv. To propose a model for study and analysis of iris patterns as a medical diagnostic tool for some clinical applications.

1.6 Organization of Thesis

This thesis has been organized into six chapters. In **Chapter 1**, we briefly defined biometric systems, iris recognition system and iridology. The three stages of iris recognition system have been discussed in brief. The chapter describes objectives and motivation of this research.

In **Chapter 2**, the various studies related to iris recognition system and different stages of iris recognition system have been presented. Literature related to image preprocessing, literature related to feature extraction and literature related to template matching have been discussed. The work related to clinical applications of iris recognition system has also been presented.

Chapter 3 provides the details of implementation of two existing iris recognition algorithms. This chapter further elaborates the proposed statistical feature extraction based iris recognition system. Features have been extracted along two different directions, namely, radial and angular directions. Results have been computed in two different directions with different set of features. The implementation of the system and its comparison with the existing algorithms is presented.

Chapter 4 describes the clinical applications of iris recognition system. Two significant applications, predicting gender and predicting diabetes, are presented. The main objective of iris recognition system is to identify the user as an authentic or an imposter. The system does not reveal about imposter's gender or ethnicity. In this chapter, SVM based gender prediction of an imposter has been proposed and implemented. Next, a health examination system to predict diabetes is proposed and implemented. Results have been computed by extracting features using two different techniques.

In **Chapter 5**, another application of iris recognition system to predict the obstructive lung disease is described. Reduced lungs capacity effects the respiration and leads to large number of obstructive lung disease. The proposed model predicts the obstructive lung disease using iris images. Implementation of the proposed model with two different feature extraction techniques is described in this chapter.

Finally, **Chapter 6** concludes the study by presenting the various inferences from different experiments. This chapter also presents the brief idea to further enhance the work carried out in this thesis.

CHAPTER 2

LITERATURE SURVEY

In this chapter, various works and studies related to iris recognition system have been discussed. The studies related to different stages of iris recognition system have also been presented. The work related to clinical applications of iris recognition system is also surveyed in this chapter.

2.1 Studies Related to Iris Recognition System

As discussed in the last chapter iris recognition system consists of three phases, namely, image preprocessing, feature extraction, template matching. An attempt has been made in this chapter to present the related literature on image preprocessing, literature on feature extraction and finally literature on template matching.

2.1.1 Image preprocessing

Image preprocessing is the preliminary stage of iris recognition system. The purpose of preprocessing is to isolate the iris region from an eye image. In this step, noise in the iris region due to reflection, illumination and occlusion because of eyelids or eyelashes is also minimized. A good number of algorithms have been proposed by different researchers for three stages of preprocessing: iris localization, iris normalization and iris image enhancement. The studies related to three stages of preprocessing are described below:

A. Iris localization

It has been noticed that a lot of work has been done to find the inner and outer boundaries of an iris. This process called as iris localization is an important step in the preprocessing of an iris image. In iris localization the inner and outer boundaries of an iris are approximated as circles. It is worth mentioning that the centre of iris and pupil

are generally different and locating these centres becomes an important process. In the following, the works carried out for iris localization are presented.

Most of the commercially used iris recognition systems use a very effective integro-differential operator proposed by Daugman (1993; 2004) to locate the inner and outer boundaries of iris as well as upper and lower eyelids. The complete operator behaves as a circular edge detector searching iteratively for the maximum contour integral derivative at successively finer scales of analysis through three-parameter space of centre coordinates and radius defining a path of contour integration. The operator is very accurate and fast in computation. Masek (2003) argued that the approach does not work fine when there is noise in the eye image, such as noise from reflections, since the algorithm works only on a local scale.

Hough transform is a standard computer vision algorithm used to determine the parameters of geometrical shapes in an image such as lines or circles. Different researchers employed Circular Hough Transform (CHT) (Kong and Zhang, 2001; Masek, 2003; Ma *et al.*, 2002; Wildes *et al.*, 1994; Tisse *et al.*, 2002). In this method, an edge map is generated by calculating the first order derivatives of intensity values in an eye image. From the edge map, votes are cast in Hough space for the parameters of circles passing through each edge point. These parameters are the centre coordinates and the radius which are sufficient to describe any circle. This method has been employed by a large number of researchers but there are certain problems associated with Hough transform such as it is computationally intensive that makes it unsuitable for the real time applications. Secondly, it requires threshold values to be chosen for edge detection which may result in critical edge points being removed and thus resulting in failure to detect circles or arcs.

Iris localization has been done by utilizing active contour model by some of the researchers (Ritter, 1999). The contour is defined as a set of vertices connected as a simple closed curve. Active contour responds to pre-set internal and external forces by deforming internally or moving across an image until equilibrium is reached. Two opposing forces change the position of number of points contained by the contour. The average radius and centre of contour obtained are the parameters of iris boundary. The efficiency of active contour model is poor if performed on edge images rather than variance image as proposed by Ritter (1999).

Bisection method (Sung *et al.*, 2004) is another technique in which edge detection technique is applied to detect the information about edges in the iris image. Here, centre of pupil is considered as the reference point to detect the inner and outer boundaries of iris image. For every two points on the same edge component, perpendicular line is drawn to the centre point. The centre of pupil is the centre point with maximum number of line intersections. Now, a virtual circle is drawn with reference to the centre of pupil and radius of circle is being increased within certain range. The circles with maximum number of edge points are considered as inner and outer boundaries of the iris. The performance of both active contour model and bisection method is affected by the non-uniform illuminations. As a result, the iris inner boundary cannot be localized accurately.

Teo and Ewe (2005) and Grabowski (2006) employed black hole search method to compute the centre and area of pupil. This approach is based upon the fact that pupil is considered as darkest region in the image. Initially, a threshold is defined to identify the dark areas in the iris image. These dark areas are called as black holes. The area of pupil is the total number of those black holes within the region. From the formulae of circle area, the radius of pupil can be computed. This method is not suitable for dark images as the dark iris area can be detected instead of the area of the pupil.

Chou *et al.* (2010) proposed IRIS-4 method for effective image segmentation. The basic idea of the IRIS-4 method is to design a classifier to remove non-target edge pixels using four-spectral measurements before a curve model is fitted to the edge pixels using Random Sample Consensus Algorithm (RANSAC). Iris boundary candidates are computed based on the canny edge detector. Here, Images in the training set should contain human eyes of different skin and iris colors for the method to work well.

After localization, the important step is to detect and remove the eyelids and eyelashes that may cover the iris areas. Numerous techniques have been employed in recent past. Masek (2003) employed line Hough transform while, Kang and Park (2007) employed parabolic Hough transform to isolate eyelids. Zhang *et al*. (2006)'s and Huang *et al*. (2004)'s algorithms use only the region of the normalized image near the pupil for reducing the effect caused by eyelid occlusion. Xu *et al*. (2006), Bachoo and Tapamo (2005) divided the iris image into blocks to isolate the eyelids. A good number of

researchers have proposed various eyelash detection techniques based upon the pixel intensities (Huang *et al.*, 2004; Kang and Park, 2007; Zhang *et al.*, 2006).

B. Iris normalization

Once iris region is successfully segmented from an eye image, next step is to transform the iris region to the fixed dimensions. The constant dimension of iris region is vital to eliminate the noise due to pupil dilation.

Daugman (1993; 2004), devised the homogeneous rubber sheet model. This model remaps each point within the iris region from Cartesian coordinates $(x,\ y)$ to polar coordinates $(r,\ \theta)$ where r is on the interval $(0,\ 1)$ and θ is angle $(0,\ 2\pi)$. The centre of the pupil is considered as the reference point and radial vectors pass through the iris region. Number of data points selected along each radial line is defined as the radial resolution and the number of radial lines going around the iris region is defined as the angular resolution. Fixed number of radial data points are taken at a particular angle. The normalized pattern is now created by backtracking to find the Cartesian coordinates of data points from the radial and angular position in the normalized pattern as

$$I\ (x\ (r,\theta), y\ (r,\theta)\ )\ \rightarrow\ I\ (r,\theta) \tag{2.1}$$

with

$$x\ (r,\theta)\ =\ (1\ -\ r)\ x_p\ (\theta)\ +\ r x_l(\theta) \tag{2.2}$$

and

$$y\ (r,\theta)\ =\ (1\ -\ r)\ y_p(\theta)\ +\ r y_l\ (\theta) \tag{2.3}$$

where $I(x,\ y)$ is iris region image, $(x,\ y)$ are the original Cartesian coordinates, $(r,\ \theta)$ are corresponding normalized polar coordinates, and $(x_p,\ y_p\ )$ & $(x_l,\ y_l)$ are the coordinates of the pupil and iris boundaries along θ direction.

This normalization process constitutes the steps of homogeneous rubber sheet model. This process gives a 2-*D* array with horizontal and vertical dimensions determined by angular and radial resolutions, respectively, as shown in Figure 2.1.

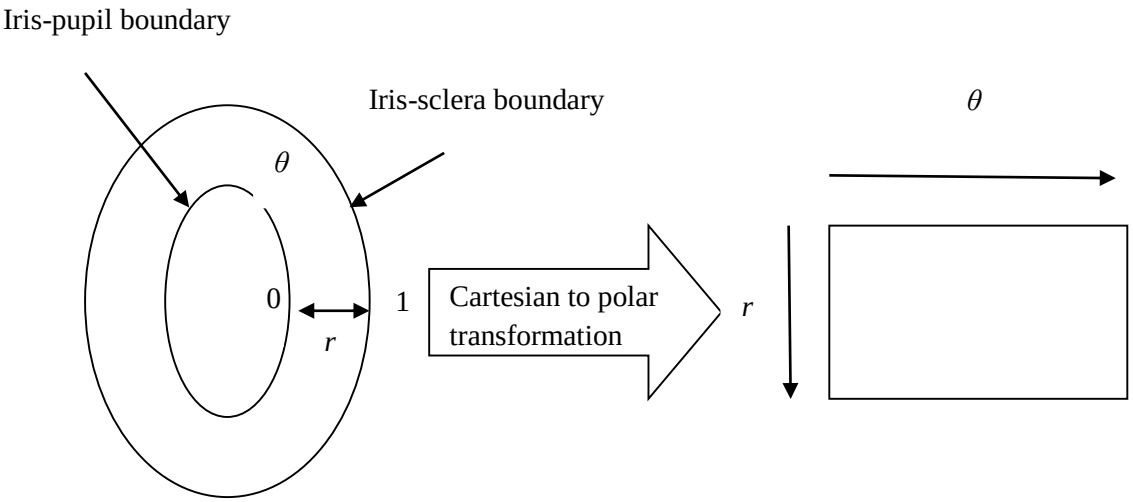


Figure 2.1: Homogeneous rubber sheet model

Wildes *et al.* (1994) utilized image registration technique for normalizing the iris image. This technique geometrically warps a newly acquired $I_a(x, y)$ image, into alignment with a selected database $I_d(x, y)$ image. The image intensity values of the new image are made to be close to those of corresponding points in the reference image.

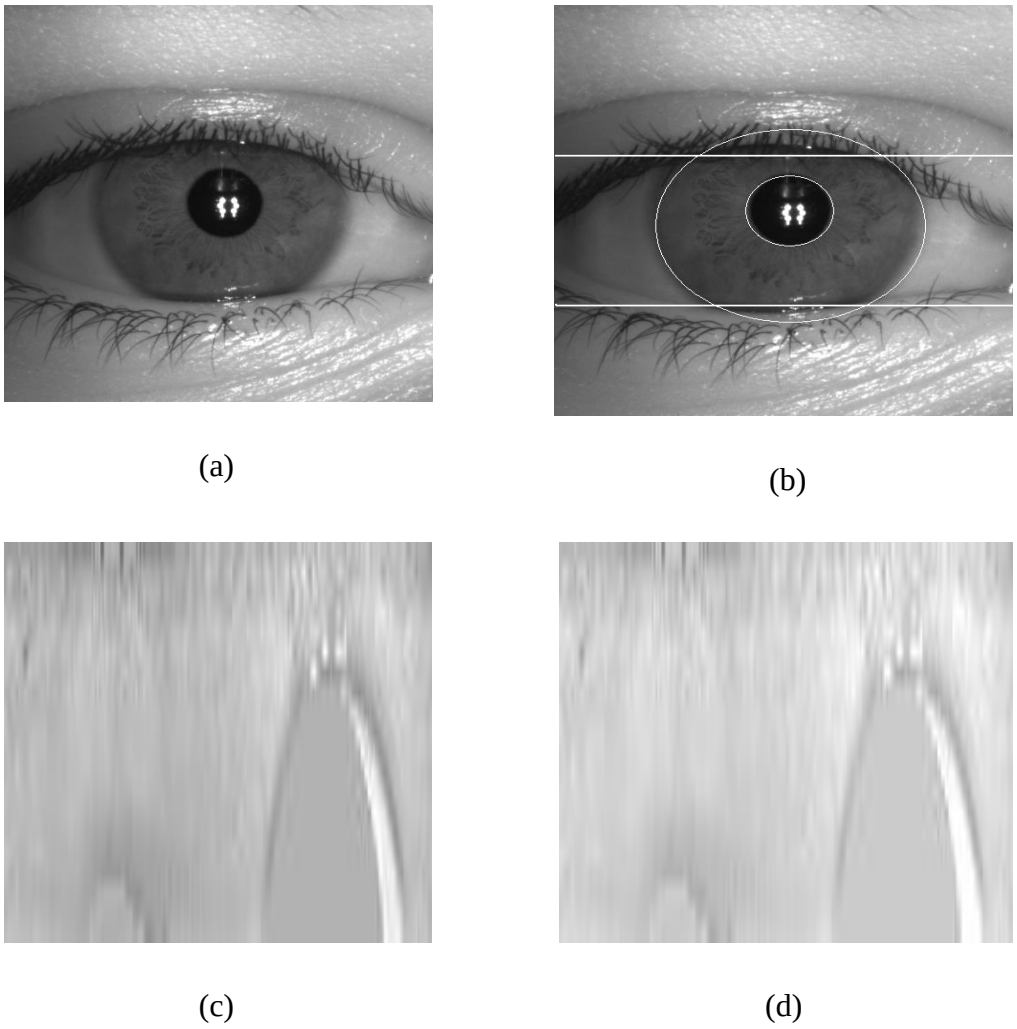
Boles and Boashash (1998) used the concept of virtual circles. Here, iris images are first scaled to have constant diameter so that when comparing two images, one is considered as the reference image. Once the two iris images have the same dimensions, features are extracted from the iris region by storing the intensity values along virtual concentric circles, with origin at the centre of the pupil. In this approach, rotational invariance is attained.

C. Image enhancement

To compensate the low contrast, poor light source and the position of light source the enhancement of the obtained normalized image is very important. Local histogram analysis (Zhu *et al.*, 2000) and simple threshold operation (Huang *et al.*, 2004) can be utilized to reduce the reflection noise.

In this thesis, for image preprocessing, initially RGB to gray transformation of iris image is carried out for the colored images. Next, the inner and outer boundaries of iris are ascertained using CHT. After segmenting iris region from an eye image, one has to

normalize the iris image which is done using Daugman's homogeneous rubber sheet model. To reduce the reflection noise local histogram analysis (Zhu *et al.*, 2000) and simple threshold operation (Huang *et al.*, 2004) have been implemented. The process of iris image preprocessing is illustrated in Figure 2.2.



(a) (b)

(c) (d)

Figure 2.2: Iris image preprocessing (a) original image (b) localized image (c) normalized image (d) enhanced image

2.1.2 Feature extraction

Various algorithms have been proposed and employed by a large number of researchers to extract the significant features from the normalized iris image.

Daugman (1993; 2004) has proposed a novel method of feature extraction using 2-*D* Gabor filter. Gabor filter's impulse response is defined by the multiplication of harmonic function with Gaussian function. The phase information present in the iris image is extracted using Gabor filter as it does not depend on extraneous factors like illumination, contrast and position of camera. Field (1987) has proposed a log-Gabor filter for feature extraction, which is Gaussian on logarithmic scale. The advantage of log filter is in removing the zero frequency component obtained in case of Gabor filter.

Wavelet transform localizes features in both spatial and frequency domain with varying window size. Bodade and Talbar (2009), Xiaofu and Pengfei (2008), Wen-Shiung *et al.* (2005), Wang and Xie (2006), Rahul *et al.* (2011), Sharma *et al.* (2014) and Rydgren *et al.* (2004) employed different wavelet functions and filter banks for the feature extraction. Every filter is tuned for fixed resolution with each wavelet defined by scaling functions. The output of filters is then encoded to generate a compact biometric template.

Hilbert transform is computationally effective and this has been explored by researchers in feature extraction. Here, an analytical image is constructed with the help of original image and its Hilbert transform (Tisse *et al.*, 2002). The filtering is performed in the Fourier domain using pure real filters. Emergent frequency and instantaneous phase is computed from the analytic image. Feature vector is encoded by thresholding the emergent frequency and the instantaneous phase.

Recently, other advanced techniques such as Gray Level Co-occurrence Matrix (GLCM) based Haralick features (Sundaram and Dhara, 2011), Local Binary Pattern (LBP) (He *et al.*, 2011), Triplet Half-Band Filter Bank (THFB) (Rahulkar and Holambe, 2012), Hybrid finer directional wavelet filter bank (Rahulkar *et al.*, 2013) and Dynamic Features (DF) (Costa and Gonzaga, 2012) have also been used in iris recognition. Sundaram and Dhara (2011) processed the normalized image by 2-*D* Haar wavelet and computed GLCM based Haralick features from the low frequency data. He *et al*. (2011) generated iris feature code by implementing chunked encoding method

based on statistical information from an iris's LBP image. Rahulkar and Holambe (2012) designed a triplet 2-*D* biorthogonal wavelet basis for iris feature extraction. Costa and Gonzaga (2012) extracted DF from iris image. Their methodology extracted information about the ways the human eye reacts to light, and used this information for biometric recognition purposes. Ko *et al.* (2006) used mean of pixel values for extracting features from iris image. Kyaw (2009) utilized statistical feature extraction technique on the iris image without implementing normalization process. He proposed a statistical technique to extract features from segmented image by considering virtual circles on iris image. A difficulty with his technique is that the image is not normalized and thus the system may not work well when there is an inconsistency in the size of the iris due to varying illumination, varying image distance *etc*.

2.1.3 Template matching

The last step in iris recognition system is the matching of individual iris code with that of iris code from the database. Iris recognition is carried out using template matching. Hamming distance is one of the most popular template matching techniques. The Hamming distance gives a measure of dissimilar number of bits between two bit patterns. By Hamming distance of two bit patterns, a decision is taken whether the two patterns were generated from different irises or from the same one. An individual iris region contains features with high degrees of freedom, each iris region will produce a bit-pattern which is independent of the pattern produced by another iris. On the other hand, two iris codes produced from the same iris will be highly correlated. Hamming distance is the matching metric employed by different researchers (Daugman, 2004; Ko *et al.*, 2006; Kyaw, 2009). The advantage of Hamming distance is its fast matching speed because templates are in binary format.

The Weighted Euclidean Distance (WED) can also be used to compare two templates, especially if the template is composed of integer values. WED gives a measure of similarity between the collections of values among two templates. This metric is employed by Zhu *et al.* (2000). An iris template is matched with all templates in the database. The two templates are said to be matched when the WED is minimum.

Normalized correlation between acquired and database representation for goodness of match has been employed by Wildes (1997). Normalized correlation is advantageous over standard correlation, since it is able to account for local variations in image

intensity that corrupt the standard correlation calculation but it is computationally not much effective.

Several researchers (Roy and Bhattacharya, 2007; 2008) employed SVM for the iris recognition. SVM is a relatively new learning machine technique, which is based on the principle of structural risk minimization. SVM is a binary classifier that optimally separates the two classes and was proposed by Cortes and Vapnik (1995). There are two important aspects in the development of SVM as classifier. The first aspect is determination of the optimal hyper plane, which will optimally separate the two classes, and the other aspect is transformation of non-linearly separable classification problem into linearly separable problem. Burges (1998), and Cristianini and Shawe (2000) provided in-depth information on SVM.

2.2 Studies Related to Clinical Applications

Iris recognition system along with clinical iridology has been employed to determine the health of an individual. Iridologists compare the specific markings in subject's iris at particular location with that of healthy human iris. Based upon comparison results they predict the disease in a particular organ. Hence, for an automated diagnosis system it becomes important to extract significant features from an iris. Biomedical signal processing for the development of diagnostic models has always been an interesting area of research. Burke and Nasor (2004) employed wavelet analysis for Electrocardiograph (ECG) signals. Nova *et al*. (2014) employed Gabor filter to extract features from Computer Tomography (CT) images for the development of emphysema classification framework. Gaidhane *et al.* (2012) employed Gerschgorin circle theorem based similarity measure technique for registration of medical images. Kumar and Anand (2006) processed the Electroencephalogram (EEG) signal for monitoring the depth of anesthesia.

Several researchers (Lai and Chiu, 2010; Othman and Prabuwono, 2010; Stearn and Swanepoel, 2007) have utilized iris images to determine the health status of an individual.

Table 2.1 illustrates the feature extraction algorithms employed by different researchers for automated iris diagnosis models. This table also contains the dataset considered by these researchers and the accuracy achieved by them in their work.

Table 2.1: Different feature extraction algorithms for iris diagnosis system

S. No.	Author(s)	Disease	Feature Extraction Algorithm	Sample Size	Accuracy (in %)
1.	Wibawa and Purnomo (2006)	Detecting broken tissues of pancreas	Minimum filter	34	94.0
2.	Ma and Li (2007)	Nerve system	Gabor filter	44	86.4
3.	Ma and Li (2007)	Alimentary canal	Gabor filter	53	84.9
4.	Ramlee and Ranjit (2009)	Cholesterol	Otsu's threshold method	30	--
5.	Lesmana *et al*. (2011)	Pancreas disorder	GLCM	50	83.3
6.	Sivasankar *et al.* (2012)	Pulmonary diseases	Fuzzy C-means clustering and gray level analysis	32	84.3

Another significant application of iris recognition system is to predict gender of an imposter from iris images. For security applications, it is equally important to know the gender of an imposter. Human attributes such as ethnicity and gender can be determined using distinct features of an iris. Khan *et al*. (2011) presented a critical evaluation and comparative study of different techniques used for gender classification. Qiu *et al.* (2006; 2007) proposed a model to determine ethnicity from iris images. The model classifies an individual as an Asian or non-Asian. They used *k*-means clustering algorithm and Gabor filter bank to obtain commonly occurring fundamental texture elements. An accuracy of 88.3% on test set has been reported using SVM classifier. Thomas *et al*. (2007) claimed to be first to predict gender from iris images. They used ND_IRIS software for segmentation and normalization, and extracted texture features from real component of log-Gabor-filtered normalized iris images. They also utilized certain geometrical features such as area of pupil, difference in the centre of iris and pupil, and difference in the area of iris and pupil. The model provided an accuracy of 80.0% using C4.5 classifier. Lagree and Bowyer (2011) predicted both ethnicity and gender from iris textures. They claimed an accuracy of more than 90.0% for ethnicity determination and 62.0% for gender determination.

As mentioned in Table 2.1, Wibawa and Purnomo (2006) and Lesmana *et al*. (2011) have diagnosed the pancreas status of subjects using iridology. The status of pancreas indicates whether the subject is affected by diabetes or not. This concept has been used to predict diabetes in a subject using iris recognition system. Sivasankar *et al.* (2012) have reported the use of iris images to predict pulmonary diseases. We have taken this concept to predict obstructive lung disease using iridology.

Chapter Summary

In this chapter, literature related to iris recognition system has been surveyed. Survey is conducted on three stages of iris recognition system, namely, preprocessing, feature extraction and template matching. An attempt has been made to discuss different algorithms of these three stages, reported in the literature. In literature little work has been reported on clinical applications of iris recognition system. In this chapter, different feature extraction algorithms employed by researchers in predicting diseases using iridology and image processing techniques have been tabulated. Studies related to prediction of gender using iris images have also been discussed.

CHAPTER 3

IRIS RECOGNITION SYSTEM

3.1 Overview

Automated security of information and authentication of persons have invariably been an interesting subject of research. Biometric systems for authentication are based on features obtained from one's face (Giot *et al.*, 2010), finger (Cao *et al.*, 2014), voice (Senoussaoui *et al.*, 2014) and/or iris (Daugman, 1993; 2004). Iris recognition system is widely used in high security areas. As discussed in Chapter 2, number of researchers have proposed various algorithms for feature extraction from iris. One of these algorithms uses statistical features for iris recognition system. In this chapter, two different types of feature extraction techniques proposed by Ko *et al*. (2006) and Kyaw (2009) have been implemented and compared. These techniques differ in the process of normalization. Here, an attempt has been made to compare these two techniques in terms of False Acceptance Rate (FAR), False Rejection Rate (FRR), memory requirement and algorithmic complexity.

A simple, fast and effective statistical features based iris recognition system has also been proposed in this chapter. System performance has been measured under different set of features.

3.2 Implementation of Existing Techniques

Different stages of iris recognition system, namely, image preprocessing, feature extraction and template matching have been described in Chapter 1. Two different types of feature extraction techniques, explaining cumulative sum based change analysis (Ko *et al.*, 2006) and explaining correlation between adjacent pixels (Kyaw, 2009) have been implemented and compared. For both techniques, image preprocessing is carried out as described in Chapter 2. These two techniques differ in the process of normalization. Ko *et al*. (2006) have utilized statistical feature extraction technique on

the normalized iris image whereas Kyaw (2009) extracted the features without normalizing the iris image. In the following sub-sections, these techniques are explained. Hamming distance (Daugman, 1993; 2004; Ko *et al.*, 2006; Kyaw, 2009) based template matching has been implemented for determining the user as authentic or imposter.

3.2.1 Feature extraction without normalization

In this technique, with center of pupil as reference number of concentric circles are drawn on the iris region along which statistical feature are extracted as shown in Figure 3.1. Statistical features, including, mean, median, mode and standard deviation have been considered in this study. These statistical features are described below, in brief, as taken from Kyaw (2009).

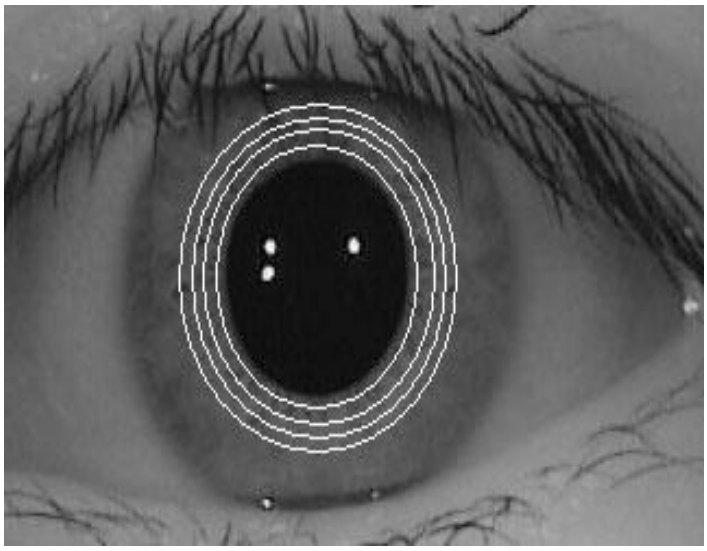


Figure 3.1: Concentric circles on iris region

- *Mean*

Mean (also called as expected value) of a statistical distribution followed by a discrete random variable, is obtained by summing the product of variable's value and its corresponding probability, over all possible values of variable as defined by the distribution. As illustrated by Kyaw (2009), this can be calculated by using the following equation for all circles in an iris image.

$$\bar{X}^c = \frac{1}{N^c}\sum_{i=1}^{N^c} X_i^c \quad , c = \overline{1, C} \tag{3.1}$$

where C is number of circles in the segmented iris, X_i^c is intensity (gradient) value of i^{th} pixel of the c^{th} circle and N^c is number of pixels along the c^{th}circle.

- *Median*

Median is the central tendency obtained by arranging the values in descending (or ascending) order and then selecting one in the middle. If total number of values in the sample is even, then median is the mean of two middle numbers. Median is a useful statistical quantity in cases where the distribution has very large extreme values that would otherwise skew the data.

- *Mode*

Mode refers to a number in the list that occurs most often.

- *Standard deviation*

Standard deviation is a measure of spread or dispersion of a set of data, *i.e.*, values are spread out for larger standard deviation. One can calculate standard deviation using

$$s = \sqrt{\frac{1}{N^c - 1}\sum_{i=1}^{N^c}(X_i^c - \overline{X_i^c})^2}\,, \; c = \overline{1, C} \tag{3.2}$$

where C is number of circles in the segmented iris, X_i^c is intensity (gradient) value of i^{th} pixel of the c^{th} circle, $\overline{X_i^c}$ is the mean of the c^{th} circle and N^c is number of pixels along the c^{th}circle.

Using these features, an image can be regarded as a feature vector (F_c) that is stored in the database for identification process.

$$F_c = (\overline{X^c}, Md^c, M^c, s^c), \; c = \overline{1, C} \tag{3.3}$$

where C is number of circles in the segmented iris, $\overline{X^c}$ is mean of the c^{th}circle, Md^c is median of the c^{th} circle, M^c is mode of the c^{th} circle and s^c is standard deviation of the c^{th} circle.

Iris code or Boolean vector is formed by thresholding the difference of features of adjacent circles. The value of code is taken as 1 if the difference between the features is positive and it is taken as 0 if the difference between the features is non-positive.

In this thesis, we have extended the work of Kyaw (2009) by conducting experiments with different number of circles. An attempt has also been made to study the effect of number of circles on system performance.

3.2.2 Feature extraction with normalization

Feature extraction from normalized iris pattern requires segmented iris image to be transformed to polar coordinates from Cartesian coordinates. For this homogeneous rubber sheet model (Daugman, 1993; 2004) is employed. Normalized iris image is used for feature extraction in case of cumulative sum based change analysis method. It is important to analyze the changes of grey values of iris patterns and extract features from iris image. Cumulative sums are calculated in a simple manner and do not require much processing. Algorithm for feature extraction using cumulative sum based change analysis as proposed by Ko *et al.* (2006) involves following steps:

Step 1. Divide normalized iris image into basic cell regions for calculating cumulative sums. (one cell region is defined as a set of $m \times n$ pixels, and an average grey value is used as a representative value of a basic cell region to calculate the cumulative sum)

Step 2. Basic cell regions are grouped into horizontal and vertical direction (five basic regions are grouped)

Step 3. Calculate cumulative sums over the each group as illustrated in (3.5).

The cumulative sums are calculated by taking mean representative values (say X_1, X_2, X_3, X_4 and X_5) of each cell regions within a group.

- First calculate the average X

$$X = (X_1 + X_2 + X_3 + X_4 + X_5)/5 \tag{3.4}$$

- Calculate cumulative sum from 0; $S_0 = 0$

- Calculate the cumulative sums by adding the difference between current value and the average to the previous sum, *i.e.*,

$$S_i = S_{i-1} + (X_i - X) \text{ for } i = 1, 2, 3, 4, 5 \qquad (3.5)$$

Step 4. Generate iris feature codes.

Iris feature code or Boolean vector is formed by comparing the cumulative sum of adjacent groups. If slope of cumulative sum change is upward, the iris code is set to one otherwise to zero.

3.3 Comparative Analysis

MatLAB code for both methods is generated and tested on 500 images of *Database A* and 500 images of *Database B.* Comparative study with and without normalization is carried out in terms of FAR, FRR, memory requirement and algorithmic complexity.

3.3.1 FAR and FRR

Performance analysis of the two methods is done by calculating FAR and FRR. FAR is computed as the ratio of number of false acceptance to total number of attempts and FRR is computed as the ratio of number of false rejections to total number of attempts.

Figures 3.2 and 3.3 depict the percentage variation in FAR and FRR with respect to Hamming distance for feature extraction with normalization for two different sets of databases. For *Database A*, the two error rates are equal at a value of 1.5% and this gives the optimal Hamming distance as 154 and for *Database B*, the Equal Error Rate (EER) has been found to be 2.7% at Hamming distance 154.

As discussed earlier, the experiments have been conducted to extend the work by Kyaw (2009). We have conducted the experiments by taking four values of number of concentric circles: 15, 20, 25 and 30 on iris image. However, results have been reported for number of circles as 15 and 30 only since the improvement in the performance of the systems with 20 and 25 concentric circles was not significant. However, it has been noticed that increase in number of circles improves the performance of iris recognition system.

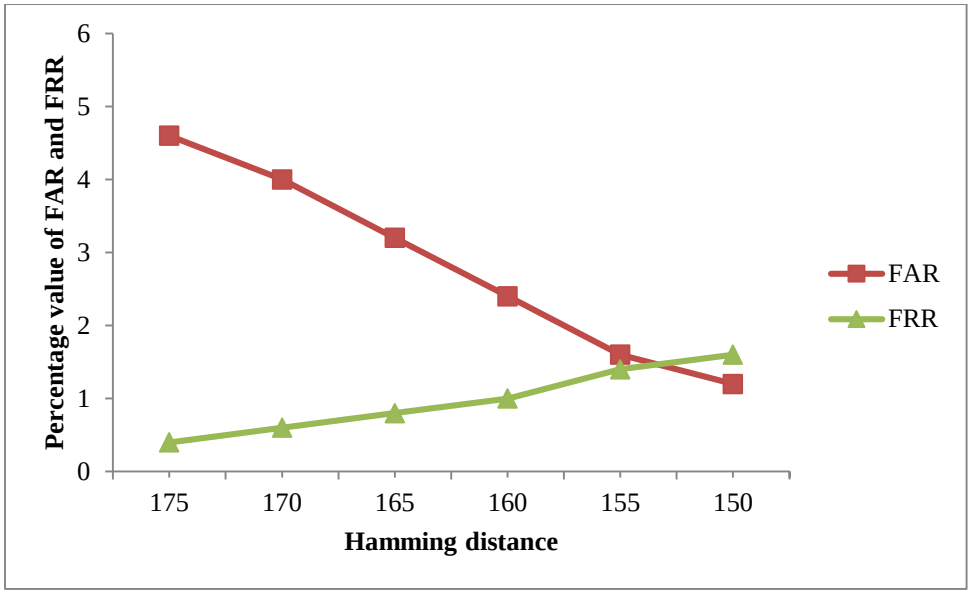


Figure 3.2: FAR and FRR v/s Hamming distance for features extracted with normalization using *Database A*

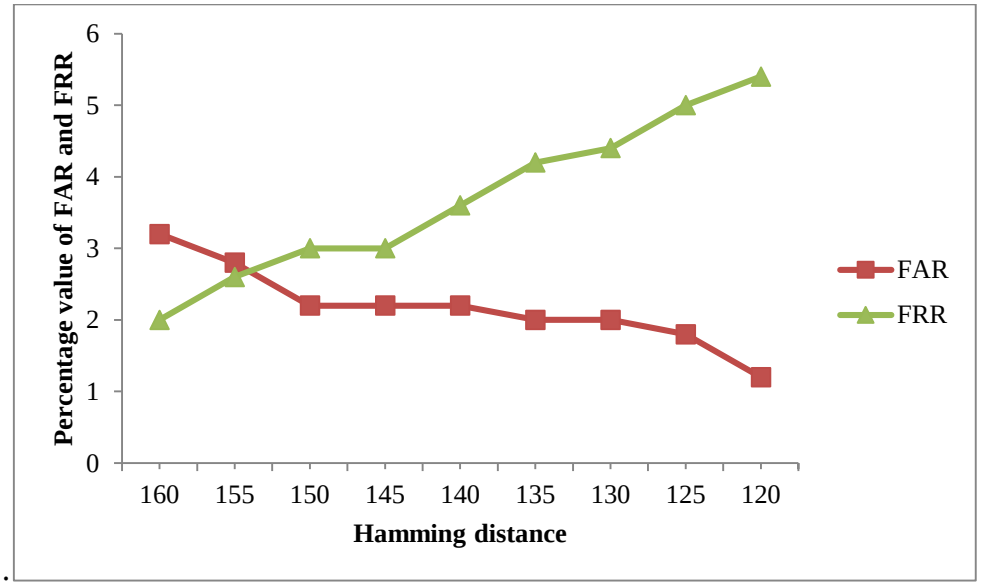


Figure 3.3: FAR and FRR v/s Hamming distance for features extracted with normalization using *Database B*

Figures 3.4 and 3.5 show the results for feature extraction without normalization for two different set of databases when number of circles drawn on iris images is 15. In this experiment, EER has been found to be 19.7% for *Database A* at Hamming distance 14.2 and EER is 22.7% for *Database B* at Hamming distance 11.6. The number of circles drawn on the iris images is increased to 30 and the results are recomputed as

shown in Figures 3.6 and 3.7. It has been observed that with the increase in number of circles EER becomes 19.3% for *Database A* at Hamming distance 36.5 and EER is 22.0% for *Database B* at Hamming distance 29.7.

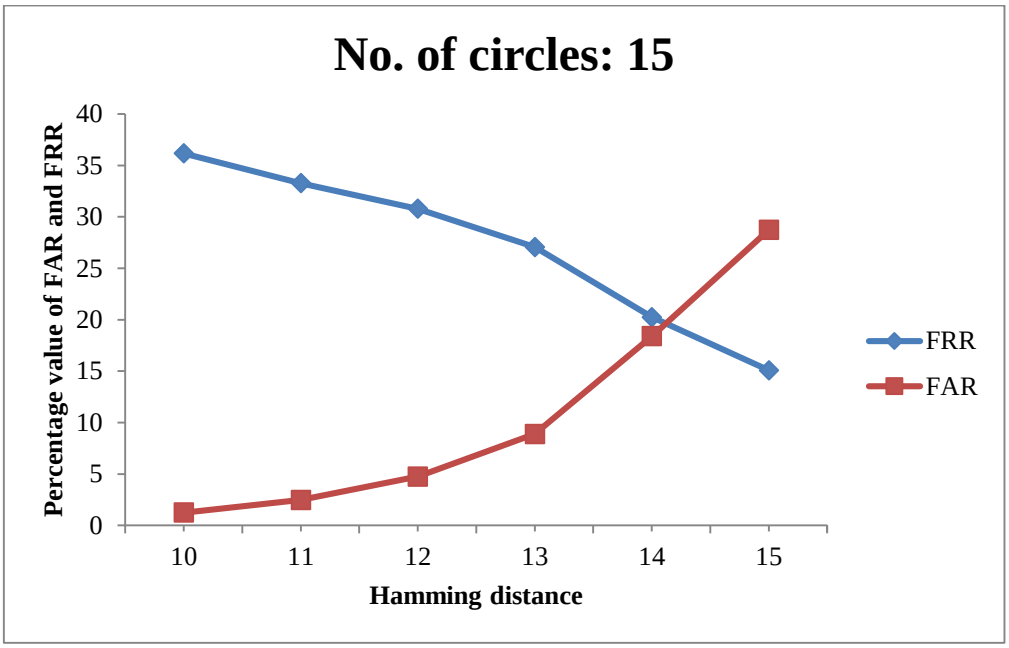


Figure 3.4: FAR and FRR v/s Hamming distance for features extracted without normalization with 15 circles using *Database A*

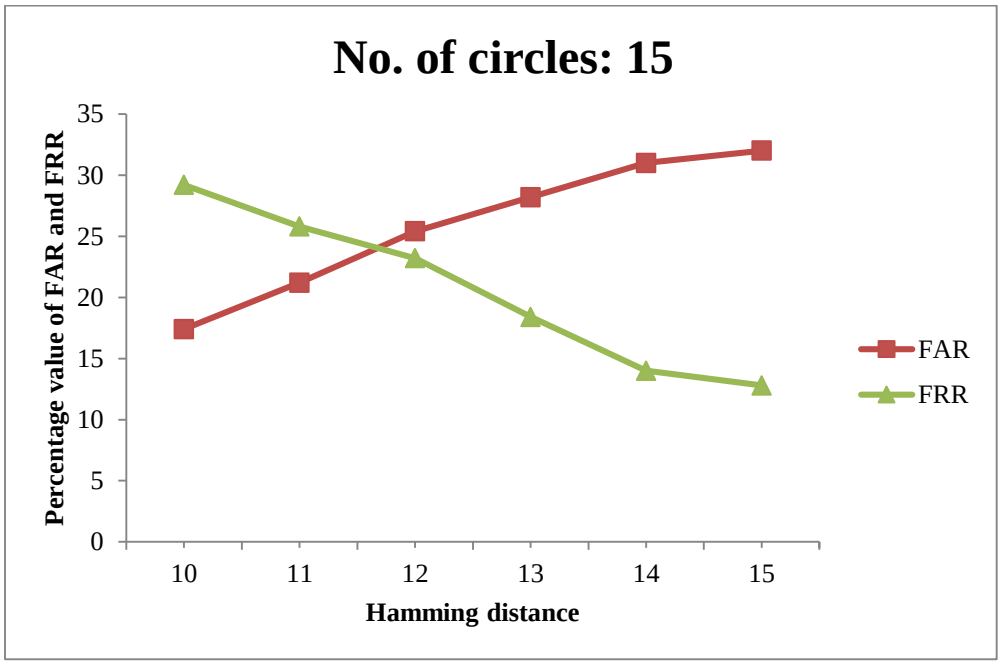


Figure 3.5: FAR and FRR v/s Hamming distance for features extracted without normalization with 15 circles using *Database B*

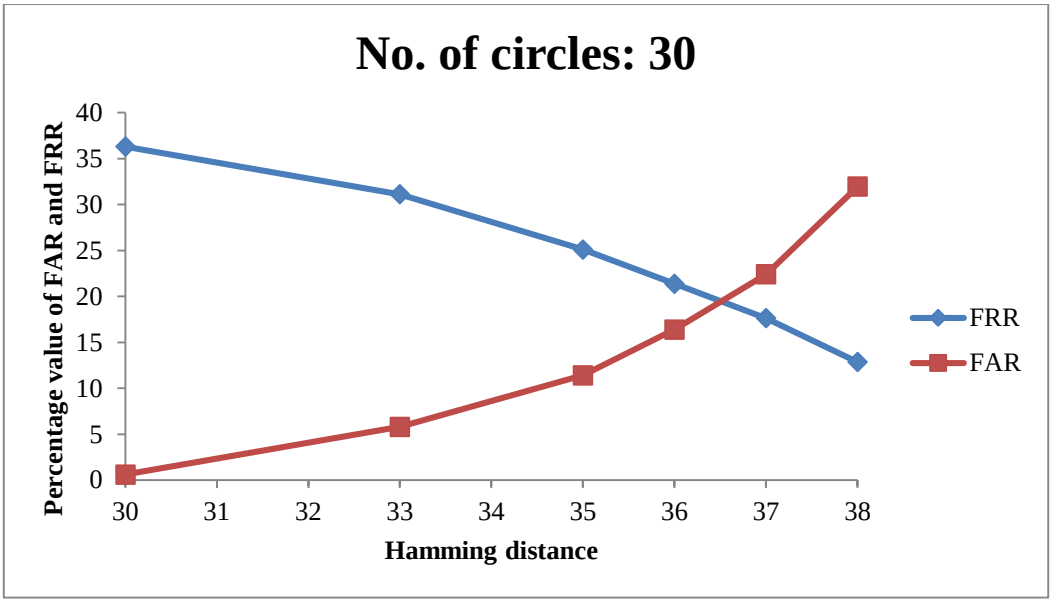


Figure 3.6: FAR and FRR v/s Hamming distance for features extracted without normalization with 30 circles using *Database A*

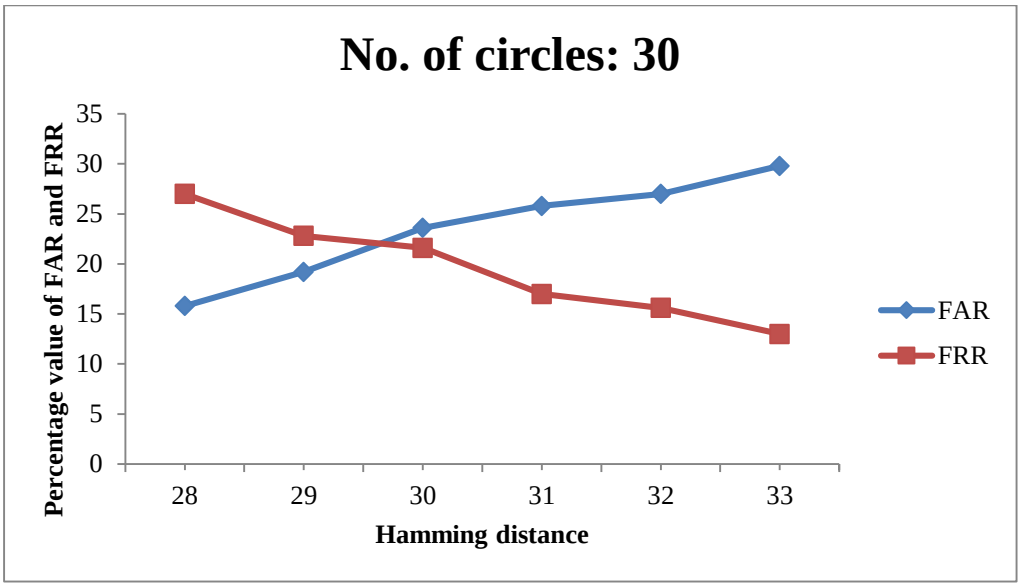


Figure 3.7: FAR and FRR v/s Hamming for features extracted without normalization with 30 circles using *Database B*

Results discussed above clearly indicate that for both available standard databases iris recognition system with features extracted on normalized iris image has much better performance as compared to one with features extracted on iris images without normalization. FAR and FRR are having much less values for feature vectors with normalization as compared to feature vectors without normalization. Figure 3.8 shows

the comparison of EER for features extracted with and without normalization. From this figure, it is evident that the value of EER is very less for feature vectors obtained with normalization when compared to feature vectors acquired without normalization. Further, as number of circles is increased to extract the features from the iris images without normalization the performance improves but still method with normalization shows better result.

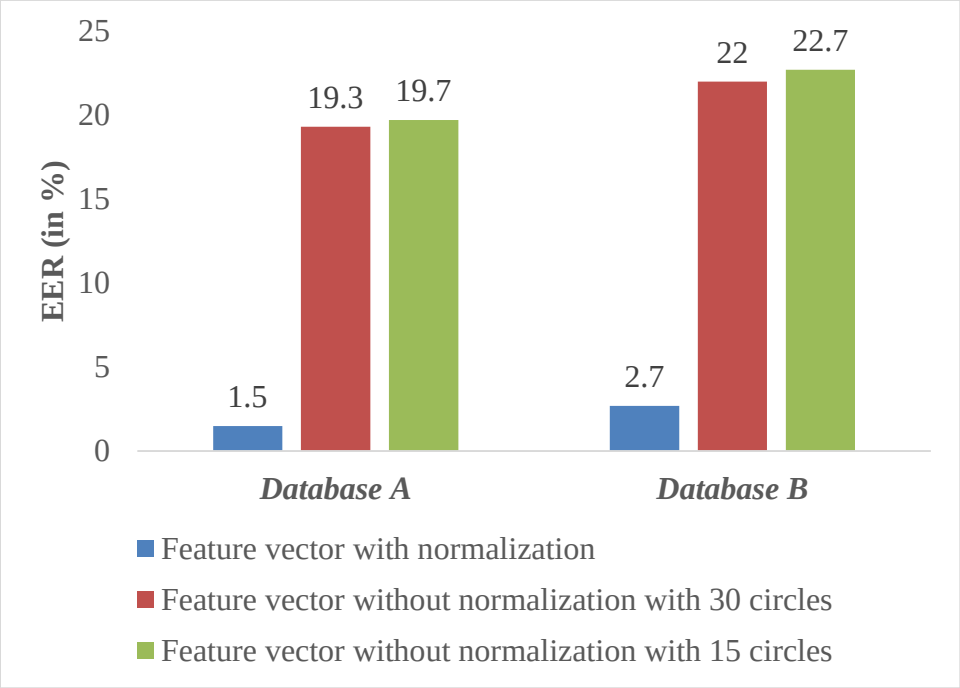


Figure 3.8: EER for features extracted with and without normalization

3.3.2 Complexity

In this work, we have compared the two techniques based on their computational complexity as it is an important issue. Algorithmic complexity of method with normalization is of the order $O(E * HG * R * D * B)$ or $O(E * VG * C * D * B)$ where E is number of images to be tested, HG is number of horizontal groups, R is number of rows considered for horizontal code, D is number of images stored in the database, B is number of bits in the feature vector, VG is number of vertical groups, C is number of columns considered for vertical code. Whereas, algorithmic complexity of method without normalization is of the order $O(E * NC * D * B)$ where E is number of images to be tested, NC is number of circles drawn on iris image, D is number of images stored in the database, B is number of bits in the feature vector. It can be noted

that under the assumption *NC* < *HG* * *R*, algorithmic complexity of the method with normalization is more than algorithmic complexity of method without normalization.

3.3.3 Memory requirement

The iris code length (number of bits) of an image in case of method without normalization is *number of features* × *number of circles* and in case of method with normalization, it is *number of horizontal groups* × *number of vertical groups.* In this work, for a normalized image of size 200 × 1000 one horizontal / vertical group consists of five basic cell regions of size 5 × 5. Iris code length for method without normalization is 120 (4 × 30) and for method with normalization is 320 (40 × 8) as depicted in Figure 3.9. Hence, memory required to store iris templates for method with normalization is more than method without normalization.

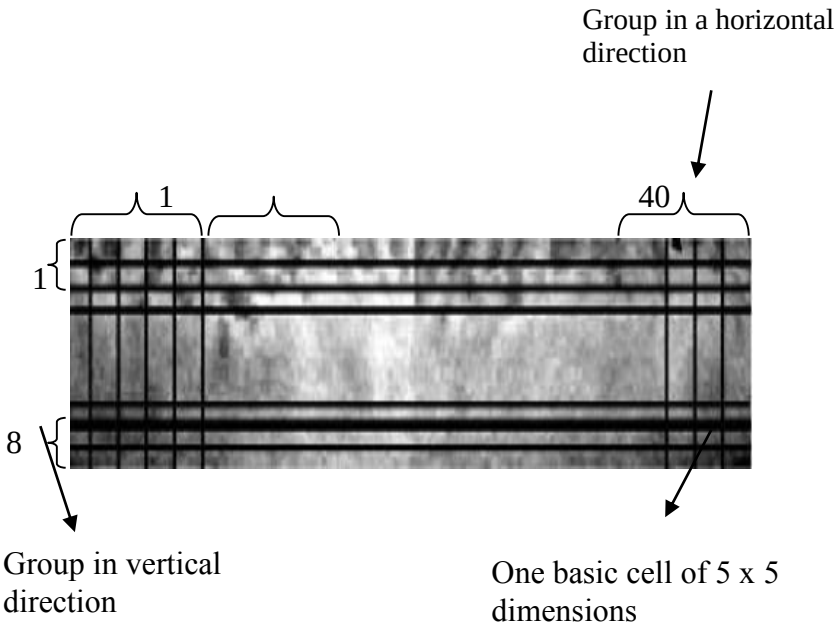


Figure 3.9: Division of normalized iris image into horizontal and vertical groups

3.4 Statistical Feature Extraction Based Iris Recognition System

As mentioned earlier, iris recognition systems have been proposed by numerous researchers using different feature extraction techniques for accurate and reliable biometric authentication. In this section, a statistical feature extraction technique based on relationship between adjacent pixels, has been proposed and implemented.

3.4.1 Proposed feature extraction technique

Once a normalized iris image is obtained after preprocessing of an eye image as illustrated in Figure 2.2, statistical approach has been used for feature extraction. Each row of normalized iris image is equivalent to a virtual circle drawn on the iris region. The features of an iris can be extracted along the concentric circles as well as along the angular direction in the iris region.

Here, two sets of experiments have been performed. In the first experiment, statistical features have been computed along each row of 2-*D* normalized array. While in second experiment, statistical features have been calculated along each column of 2-*D* normalized array. The effect of number of statistical features as well as radial and angular resolutions while normalization, on the performance of proposed iris recognition system, has also been analyzed. The combinations of radial and angular resolutions considered in the present work are, (50, 1000), (100, 1000), (150, 1000) and (200, 1000) while experimenting with radial resolution and (200, 250), (200, 500), (200, 750) and (200, 1000) while experimenting with angular resolution.

For both experiments, three parameters, namely, mean, median and standard deviation have initially been used. The system performance has been recomputed using six parameters, namely, mean, median, standard deviation, skewness, kurtosis and co-efficient of variation. Mean, median and standard deviation have been described in Section 3.2.1. Skewness, kurtosis and coefficient of variation are described below.

- *Skewness* (S_k)

  Skewness is a measure of the asymmetry of the data around the sample mean. If skewness is negative, the data are spread out more to the left of the mean than to the right. If skewness is positive, the data are spread out more to the right. The skewness for different rows is taken as,

$$S_k^r = \frac{\frac{1}{N^r}\sum_{i=1}^{N^r}(X_i^r - \bar{X}^r)^3}{s^{r3}},\ r = 1, 2, 3, \ldots, R \tag{3.6}$$

  where R is number of rows in the normalized iris image, X_i^r is intensity value of i^{th} pixel of the r^{th} row, $\bar{X}^r$ is the mean of the r^{th} row, s^r is the standard deviation of the r^{th} row.

- *Kurtosis* (ku)

Kurtosis is a measure of outlier-prone of a distribution. The kurtosis of the normal distribution is 3. Distributions that are more outlier-prone than the normal distribution have kurtosis greater than 3; distributions that are less outlier-prone have kurtosis less than 3. The kurtosis for different rows is defined as:

$$ku^r = \frac{\frac{1}{N^r}\sum_{i=1}^{N^r}(X_i^r - \bar{X}^r)^4}{s^{r\,4}}, \quad r = 1, 2, 3, \dots, R \tag{3.7}$$

where R is number of rows in the normalized iris image, X_i^r is intensity value of i^{th} pixel of the r^{th} row, $\bar{X}^r$ is the mean of the r^{th} row, s^r is the standard deviation of the r^{th} row.

- *Co-efficient of variation (CV)*

Co-efficient of variation is the percentage variation in the mean, standard deviation being considered as the total variation in the mean. Co-efficient of variation has been calculated using equation,

$$CV^r = \frac{s^r}{\bar{X}^r} \times 100, \quad r = 1, 2, 3, \dots, R \tag{3.8}$$

where R is number of rows in the normalized iris image, $\bar{X}^r$ is the mean of the r^{th} row, s^r is the standard deviation of the r^{th} row.

In first experiment, statistical features are computed along each row. The computation gives a set of feature vectors (F_r) for an image. This set is stored in the database for identification process. This set is denoted as:

$$F_r = (\bar{X}^r,\ Md^r,\ s^r,\ S_k^r,\ ku^r,\ CV^r),\ r = 1, 2, 3, \dots, R \tag{3.9}$$

where R is the number of rows in normalized iris image, $\bar{X}^r$ is mean of the r^{th} row, Md^r is median of the r^{th} row, s^r is standard deviation of the r^{th} row, S_k^r is skewness of the r^{th} row, ku^r is kurtosis of the r^{th} row, and CV^r is co-efficient of variation of the r^{th} row.

In second experiment, statistical features are computed along each column. This computation gives a set of feature vectors (F_c) for an image. This set is stored in the database for identification process. This set is denoted as:

$$F_c = (\bar{X}^c,\ Md^c,\ s^c,\ S_k^c,\ ku^c,\ CV^c),\ c = 1, 2, 3, \dots, C \tag{3.10}$$

where C is the number of columns in normalized iris image, $\bar{X}^c$ is mean of the c^{th} column, Md^c is median of the c^{th} column, s^c is standard deviation of the c^{th} column, S_k^c is skewness of the c^{th} column, ku^c is kurtosis of the c^{th} column, and CV^c is co-efficient of variation of the c^{th} column.

In this work, similarity of two iris codes is obtained using Hamming distance. Hamming distance requires feature vectors to be converted into binary format. The binary vector in the first experiment is formed by taking the difference between features of the adjacent rows and then thresholding the difference to a binary number, whereas, the binary vector in the second experiment is formed by taking the difference between features of the adjacent columns and then thresholding the difference to a binary number. In both experiments, 1-bit per feature for truncation is needed. Although, segmentation of eyelids and removal of eyelashes have been considered but rotational noise has not been removed. Therefore, shifting of templates has not been considered. Hamming distance measures the number of dissimilar bits between two binary vectors and this distance is zero when two vectors are from same iris image. A distance metric of Hamming distance between binary vectors of test image and iris templates stored in the database is generated. Now, minimum of the distance metric is compared with a matching threshold to decide the user as an authentic or an imposter. If the minimum Hamming distance is less than the matching threshold, the two templates are from same iris image and if it is larger than matching threshold, the subject is considered as an imposter. Performance of proposed iris recognition system has been measured by recording FAR and FRR at different matching thresholds in the distance metric.

3.4.2 Results and discussions

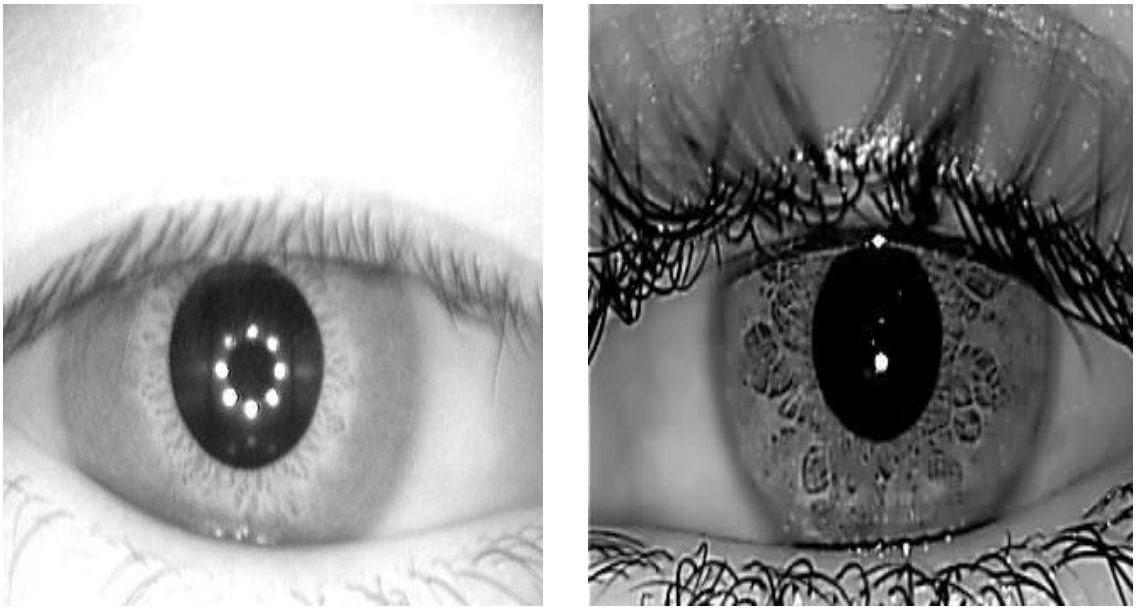
Experiments have been conducted on two databases, namely, *Database A* consisting of 2240 iris images acquired from 224 subjects and interval subset of *Database B* consisting of 2639 iris images acquired from 249 subjects. Preprocessing, feature extraction and recognition processes have been implemented on these images using image processing module of MatLAB 7.1 on Intel Core2 Duo 1.80 GHz processor with 1 GB RAM. Preprocessing is carried out as explained in Chapter 2.

The proposed system has been tested only on the successfully segmented iris images. Table 3.1 shows the number of iris images segmented successfully for both the databases. The results in this table are based on visual inspection.

Table 3.1: Successfully segmented images

Database	**Number of Images Considered**	**Number of Images Segmented Successfully**	**Result (in %)**
Database A	2240	2192	97.9
Database B	2639	2589	98.1

The segmentation process could not be applied successfully on the eye images having higher occlusion / reflection. Two of the eye images where iris could not be segmented successfully are given in Figures 3.10 (a) and 3.10 (b)



(a) (b)

Figure 3.10: Eye images where segmentation failed

3.4.2.1 Feature extraction along concentric circles

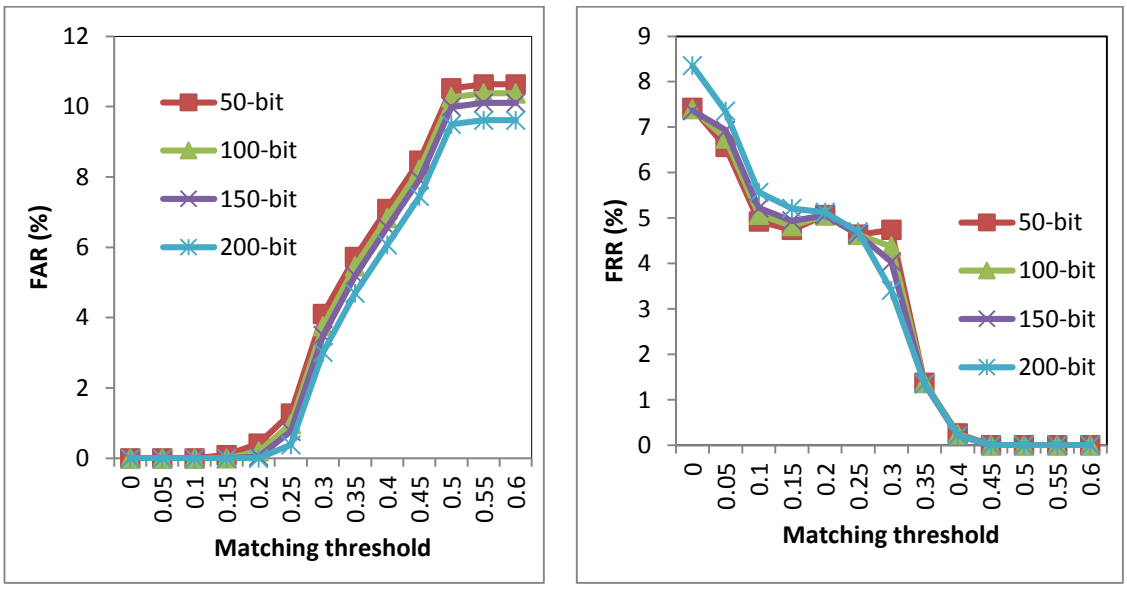
In this experiment, features have been extracted along each row. Number of rows in the normalized image represents radial resolution in normalization process that

corresponds to length of one feature. Feature vector as discussed earlier is formed by combining different features.

(i) Experimentation with three statistical parameters

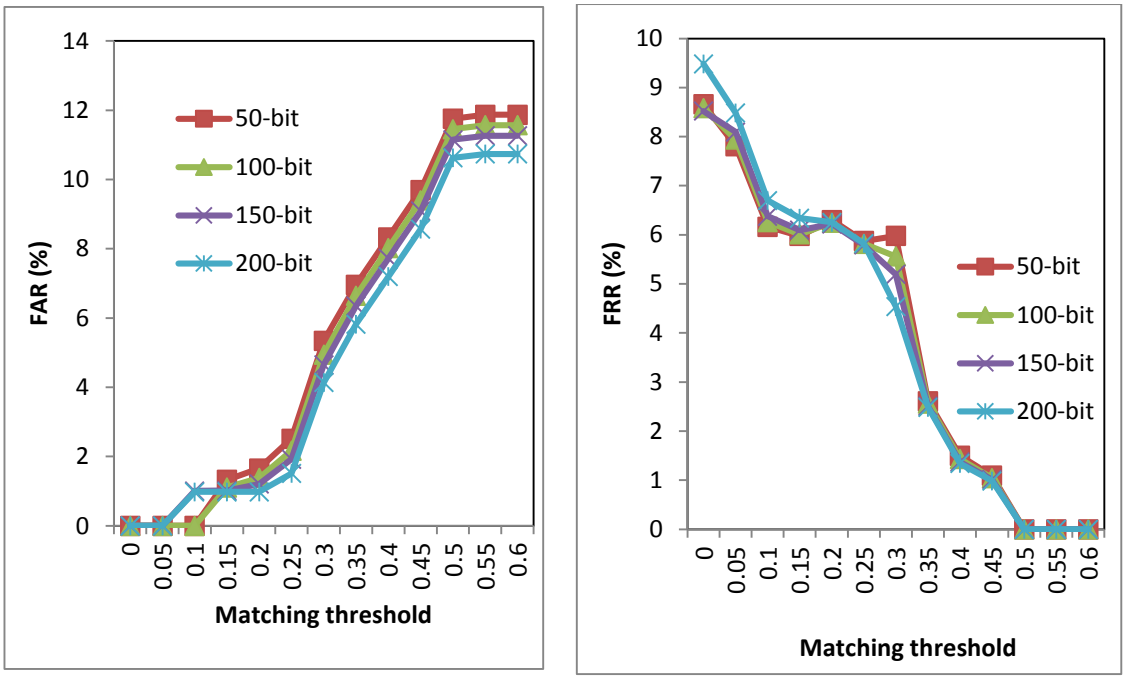
Iris recognition system performance has initially been measured by considering three statistical parameters, namely, mean, median and standard deviation. FAR and FRR have been computed for 50, 100, 150 and 200 rows in 2-*D* normalized iris image for the iris images in the two databases. Length of feature vector for each feature is therefore 50-bit, 100-bit, 150-bit and 200-bit. Figures 3.11 (a) and 3.11 (b) show variations in FAR, FRR for feature vector with different sizes for *Database A* and Figures 3.12 (a) and 3.12 (b) show variations in FAR, and FRR for feature vector with different sizes for *Database B*.

Figures 3.13 and 3.14 depict Detection Error Tradeoff (DET) graphs for *Database A* and *Database B*, respectively. DET is the graphical representation of FRR as a function of FAR. In the first set of experiments, while considering three parameters one can note from Figures 3.11 (a), 3.11 (b), 3.12 (a) and 3.12 (b) that increase in the threshold value of Hamming distance, from 0 to 0.6 (with an increment of 0.05), decreases FRR and increases FAR of the system for both databases. Two errors become equal at threshold value of 0.31.



(a) (b)

Figure 3.11: FAR and FRR as a function of matching threshold with variations in radial resolution for *Database A* using three parameters (a) FAR and (b) FRR



(a) (b)

Figure 3.12: FAR and FRR as a function of matching threshold with variations in radial resolution for *Database B* using three parameters (a) FAR and (b) FRR

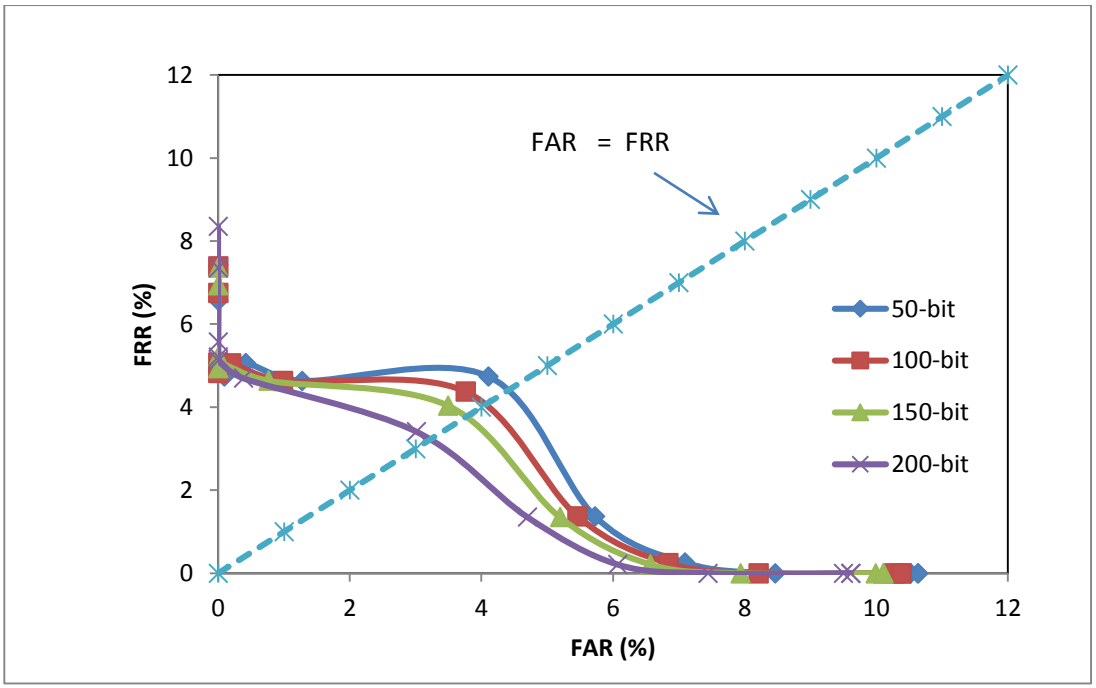


Figure 3.13: DET graph with variations in radial resolution for *Database A* using three parameters

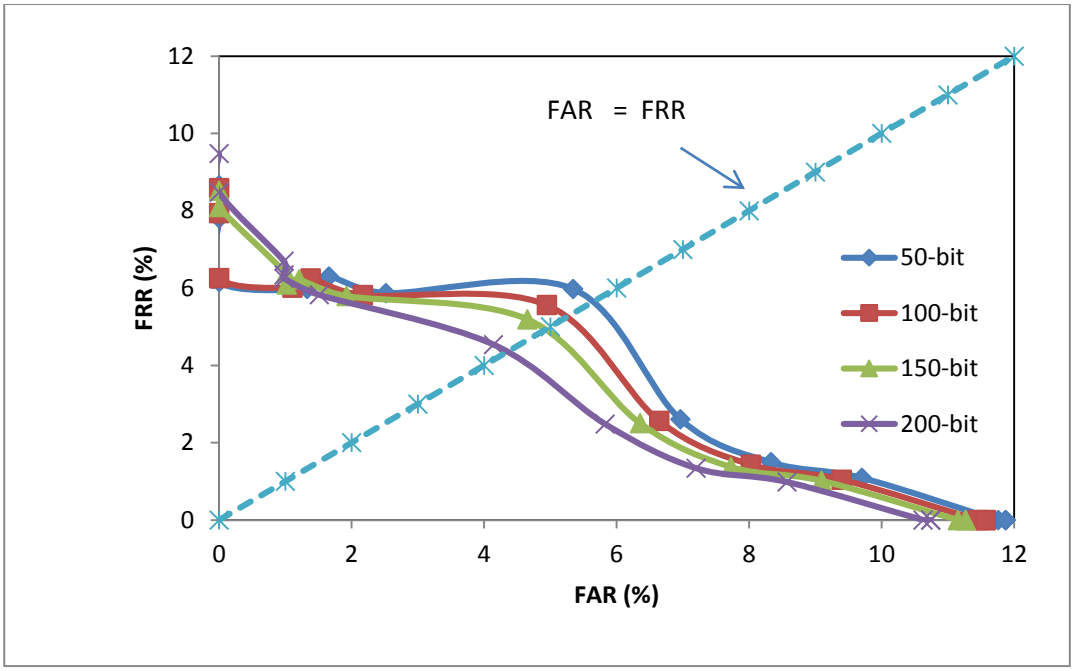


Figure 3.14: DET graph with variations in radial resolution for *Database B* using three parameters

On increasing the size of binary feature vector from 50-bit to 200-bit, FAR decreases and FRR increases at a given value of matching threshold. It is important to note that EER decreases with increase in feature vector size. Table 3.2 shows the comparison of EER for feature vectors with different sizes.

Table 3.2: EER for three parameters with variations in radial resolution

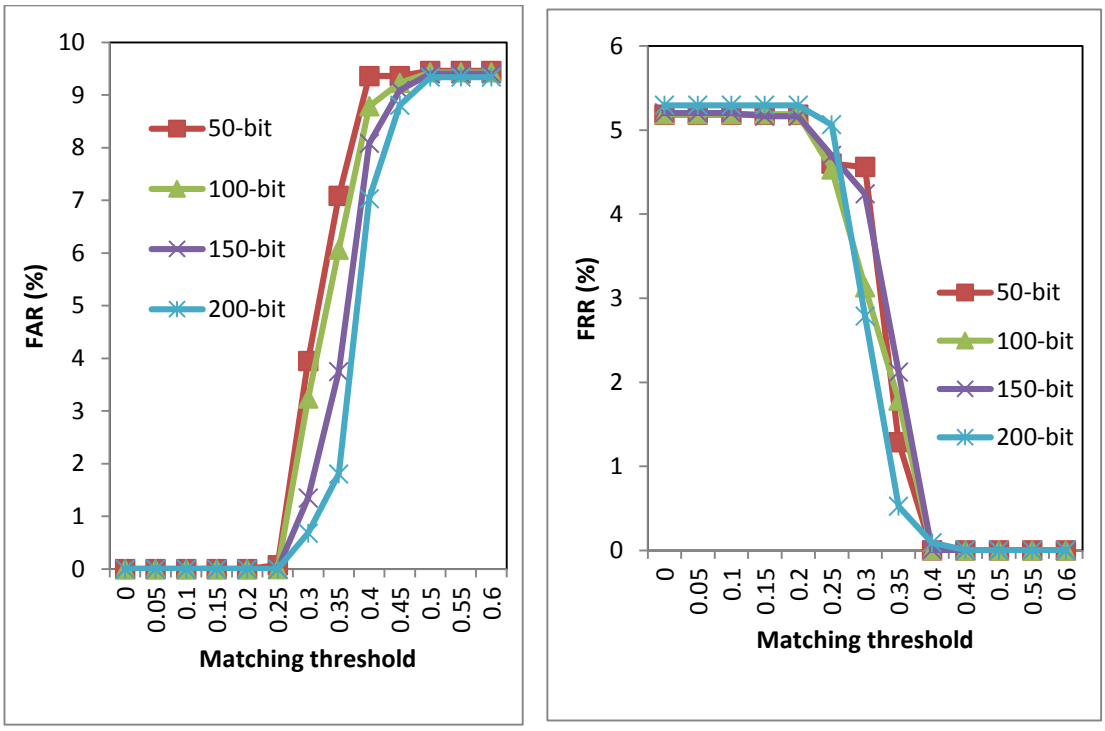
Radial Resolution	***Database A*** **EER (in %)**	***Database B*** **EER (in %)**
50	4.3	5.5
100	3.9	5.2
150	3.7	4.9
200	3.2	4.3

Results given in Table 3.2 suggest that when radial resolution is increased from 50 to 200, EER decreases from 4.3% to 3.2% for *Database A* and EER decreases from 5.5% to 4.3% for *Database B*. Hence, for a given set of statistical features, increasing radial resolution with normalization improves the performance of iris recognition system.

(ii) Experimentation with six statistical parameters

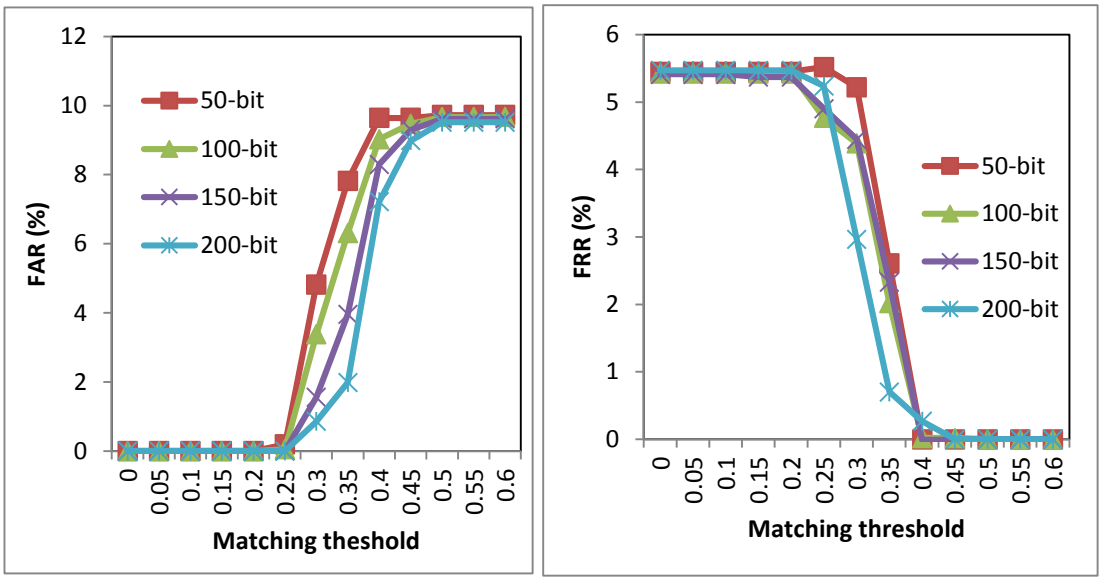
In this experiment, six statistical parameters, namely, mean, median, standard deviation, skewness, kurtosis and coefficient of variation have been considered. FAR and FRR have been recomputed for 50, 100, 150 and 200 rows in 2-*D* normalized iris image. Figures 3.15 (a) and 3.15 (b) show variations in FAR and FRR for feature vectors with different sizes for *Database A* and Figures 3.16 (a) and 3.16 (b) show variations in FAR and FRR for feature vectors with different sizes for *Database B*.

Figures 3.17 and 3.18 show DET graphs for *Database A* and *Database B*, respectively. One can infer from Figures 3.15 (a), 3.15 (b), 3.16 (a) and 3.16 (b) that increase in matching threshold from 0 to 0.6 (with an increment of 0.05) has similar effect on the variations in FAR and FRR as with three statistical parameters.



(a) (b)

Figure 3.15: FAR and FRR as a function of matching threshold with variations in radial resolution for *Database A* using six parameters (a) FAR and (b) FRR



(a) (b)

Figure 3.16: FAR and FRR as a function of matching threshold with variations in radial resolution for *Database B* using six parameters (a) FAR and (b) FRR

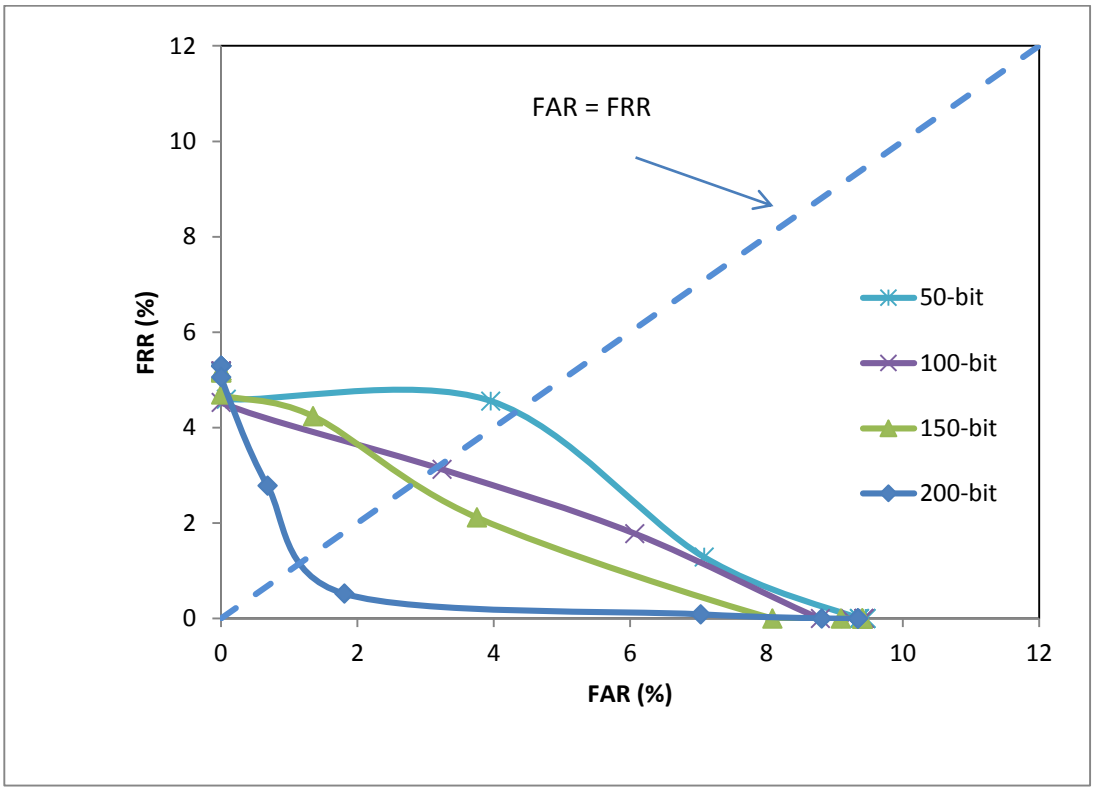


Figure 3.17: DET graph with variations in radial resolution for *Database A* using six parameters

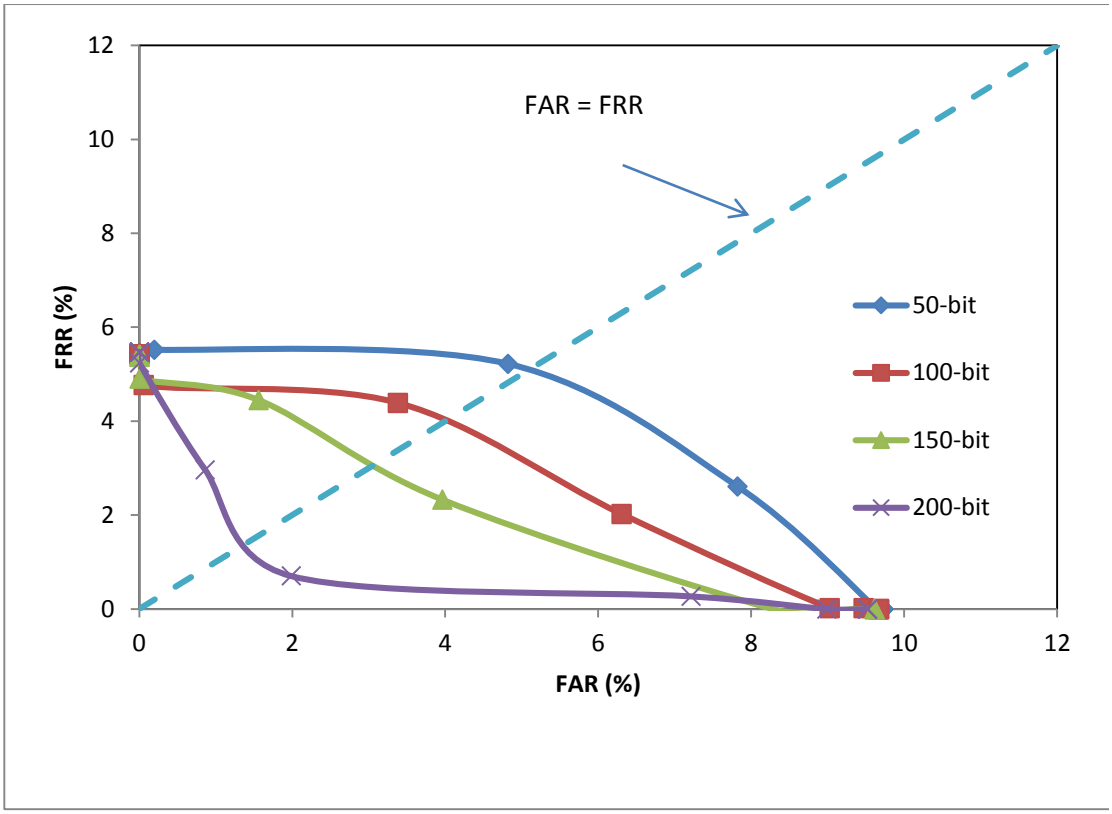


Figure 3.18: DET graph with variations in radial resolution for *Database B* using six parameters

Using DET graphs shown in Figures 3.17 and 3.18, EER for two databases is computed at different radial resolutions while normalization as given in Table 3.3. When radial resolution is increased, the EER decreases from 4.3% to 1.3% for *Database A* and this decreases from 5.0% to 1.5% for *Database B*.

Table 3.3: EER for six parameters with variations in radial resolution

Radial Resolution	***Database A*** **EER (in %)**	***Database B*** **EER (in %)**
50	4.3	5.0
100	3.1	3.9
150	2.8	3.0
200	1.3	1.5

These two experiments based on different sets of statistical features show that there is a significant improvement in FAR, FRR and EER as the number of statistical parameters for feature extraction from iris images is increased from three to six. The proposed statistical feature extraction based iris recognition system along concentric circles/radial direction is simple and effective with EER of 1.3% for *Database A* and of 1.5% for *Database B*.

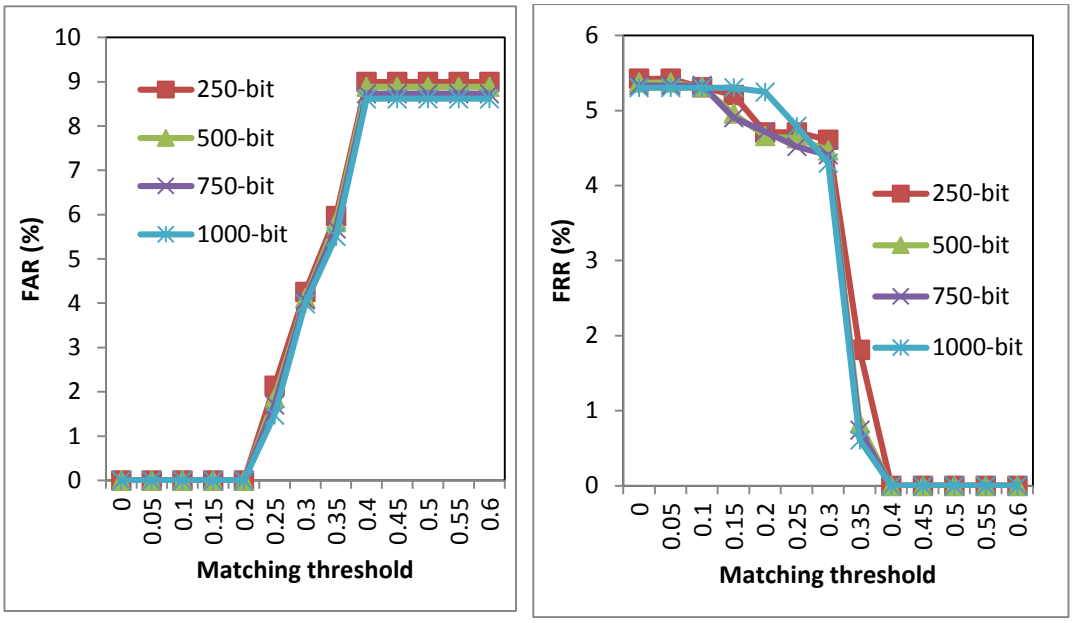
3.4.2.2 Feature extraction along angular direction

In second experiment, statistical features have been computed along the radial vectors drawn in the iris region extending from pupil-iris boundary to iris-sclera boundary. In 2-*D* normalized iris image, each column corresponds to a radial vector drawn on iris region. Here, features have been extracted along each column of the normalized iris image. Number of columns in the normalized image represents angular resolution in normalization process that corresponds to length of one feature. Feature vector, as discussed earlier, is obtained by combining different features.

(i) Experimentation with three statistical parameters

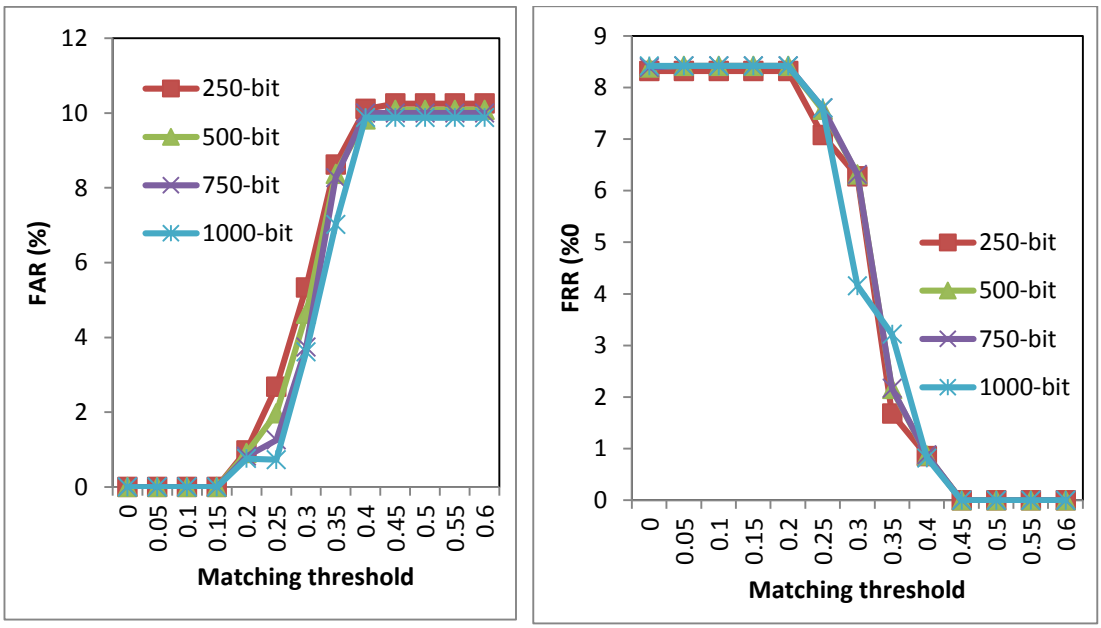
In this set of experiment three parameters, namely, mean, median and standard deviation have been considered on the lines similar to feature extraction along radial direction. Experiment has again been conducted on two different sets of iris databases. FAR and FRR have been computed for 250, 500, 750 and 1000 columns in 2-*D* normalized iris image. Length of feature vector for each feature is thus 250-bit, 500-bit, 750-bit and 1000-bit. Figures 3.19 (a) and 3.19 (b) show variations in FAR and FRR for feature vector with different sizes for *Database A* and Figures 3.20 (a) and 3.20 (b) show variations in FAR and FRR for feature vector with different sizes for *Database B*.

Figures 3.21 and 3.22 show DET graphs for *Database A* and *Database B*, respectively. For three parameters, one can note from Figures 3.19 (a), 3.19 (b), 3.20 (a) and 3.20 (b) that the variations in FAR and FRR of iris recognition system is similar to the variations observed in the first experiment.



(a) (b)

Figure 3.19: FAR and FRR as a function of matching threshold with variations in angular resolution for *Database A* using three parameters (a) FAR and (b) FRR



(a) (b)

Figure 3.20: FAR and FRR as a function of matching threshold with variations in angular resolution for *Database B* using three parameters (a) FAR and (b) FRR

On increasing the size of binary feature vector from 250-bit to 1000-bit, FAR decreases and FRR increases at a given value of matching threshold. Table 3.4 contains the comparison of EER for feature vector with different sizes.

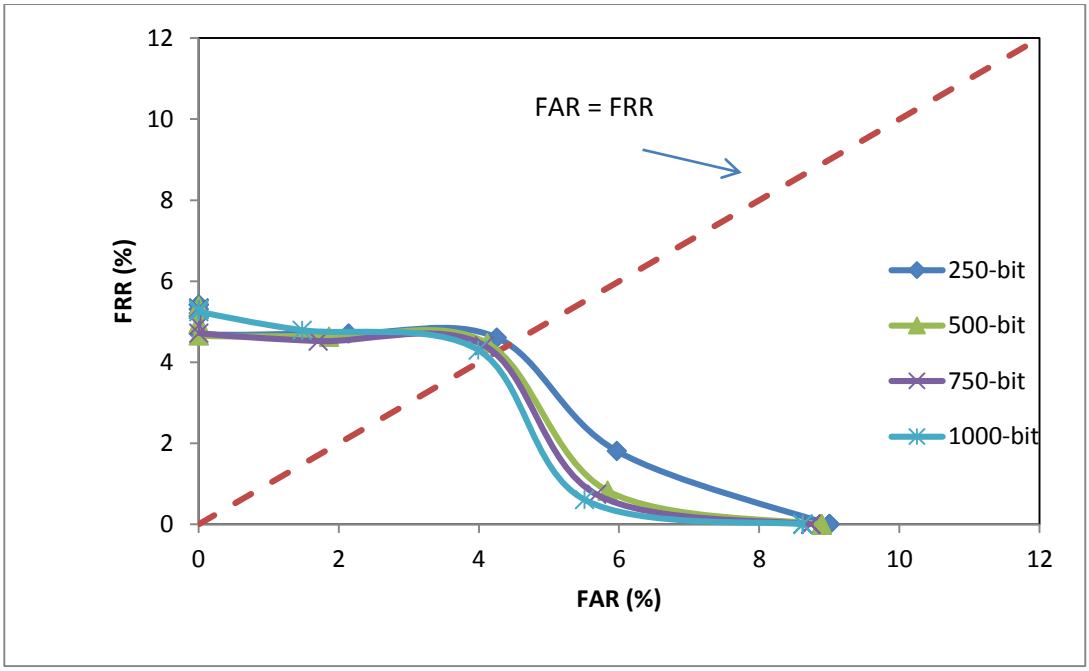


Figure 3.21: DET graph with variations in angular resolution for *Database A* using three parameters

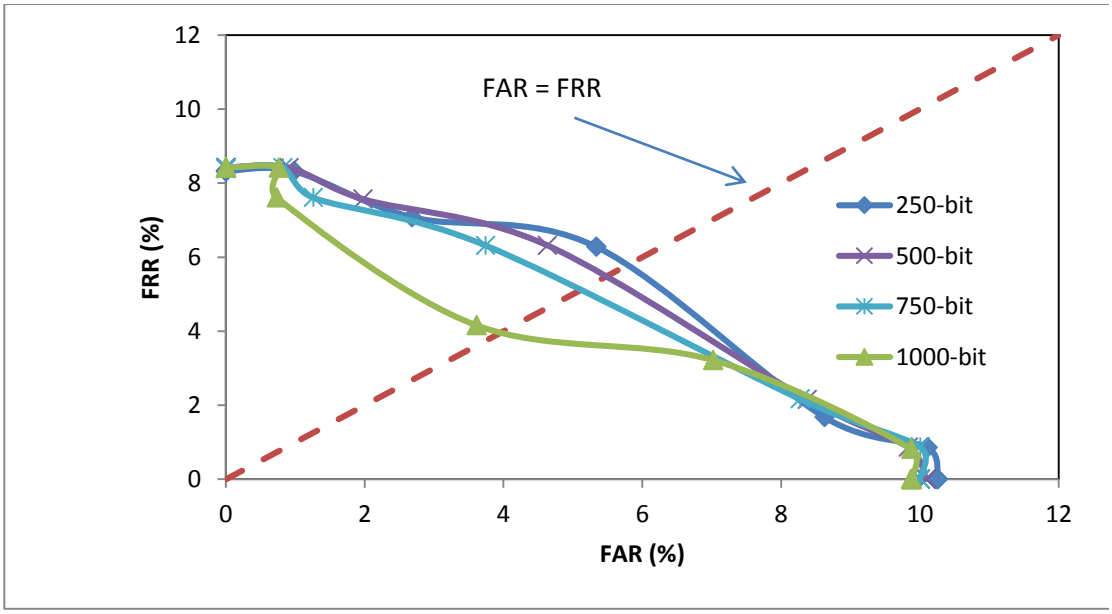


Figure 3.22: DET graph with variations in angular resolution for *Database B* using three parameters

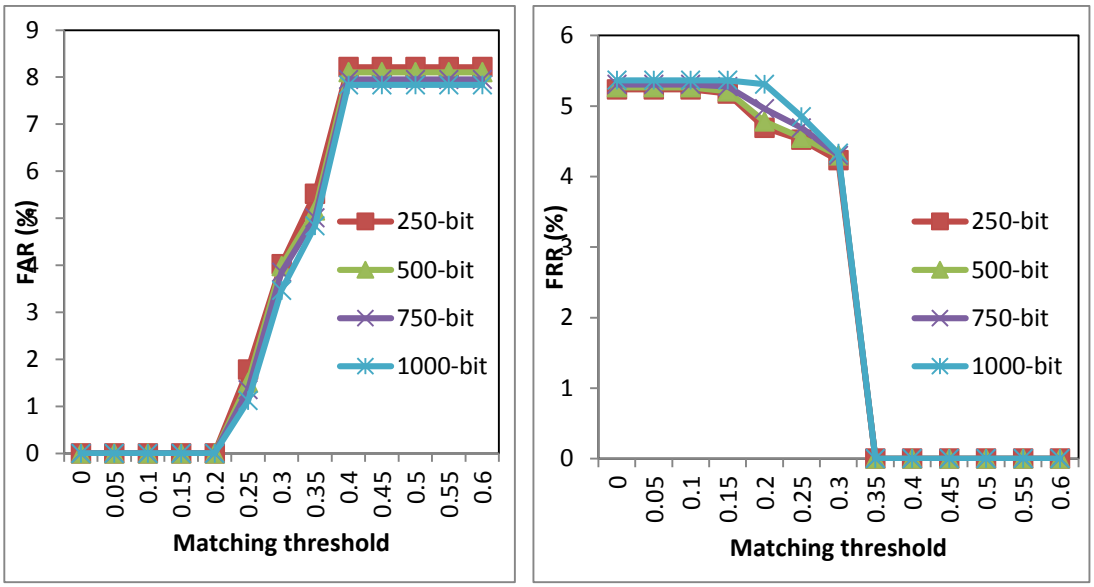
Results given in Table 3.4 suggest that if angular resolution is increased from 250-bit to 1000-bit EER decreases from 4.3% to 4.0% for *Database A* and decreases from 5.7% to 4.0% for *Database B*. As such, for a given set of statistical features, increasing angular resolution with normalization improves the performance.

Table 3.4: EER for three parameters with variations in angular resolution

Angular Resolution	*Database A* EER (in %)	*Database B* EER (in %)
250	4.3	5.7
500	4.2	5.4
750	4.1	5.0
1000	4.0	4.0

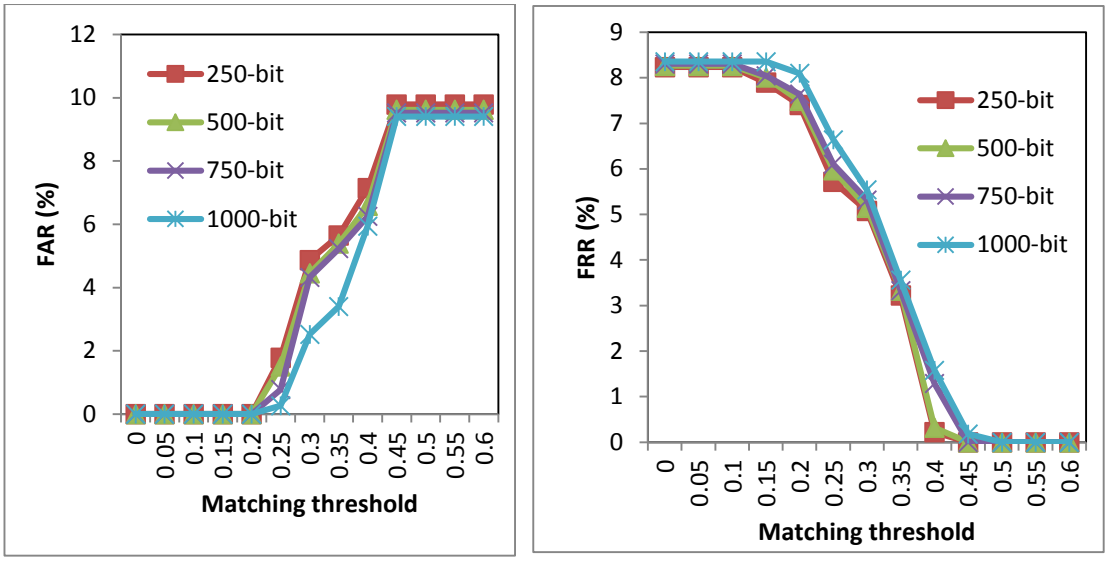
(ii) Experimentation with six statistical parameters

Iris recognition system performance is next measured by considering six statistical parameters. FAR and FRR have again been computed for 250, 500, 750 and 1000 columns in 2-*D* normalized iris image. Figures 3.23 (a) and 3.23 (b) show variations in FAR and FRR for feature vector with different sizes for *Database A* and Figures 3.24 (a) and 3.24 (b) show variations in FAR and FRR for feature vector with different sizes for *Database B.*



(a) (b)

Figure 3.23: FAR and FRR as a function of matching threshold with variations in angular resolution for *Database A* using six parameters (a) FAR and (b) FRR



(a) (b)

Figure 3.24: FAR and FRR as a function of matching threshold with variations in angular resolution for *Database B* using six parameters (a) FAR and (b) FRR

It can be noted from Figures 3.23 (a), 3.23 (b), 3.24 (a) and 3.24 (b) that trend of variations in FAR and FRR for six parameters along angular direction is same as depicted in earlier experiments. EER is again computed for two databases at different angular resolutions while normalization and is given in Table 3.5.

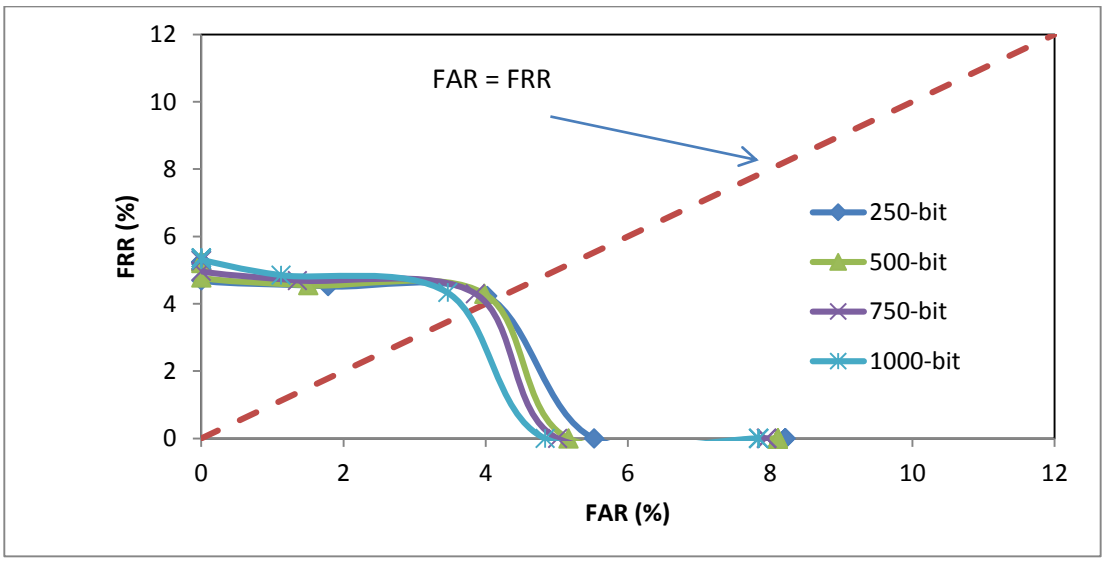


Figure 3.25: DET graph with variations in angular resolution for *Database A* using six parameters

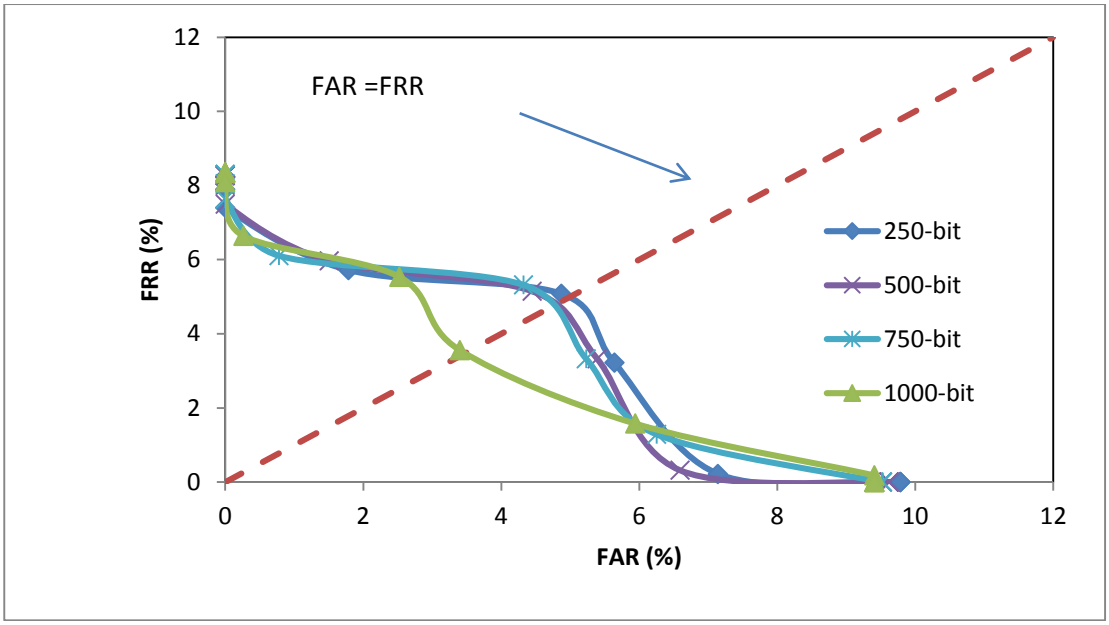


Figure 3.26: DET graph with variations in angular resolution for *Database B* using six parameters

When angular resolution is increased, EER decreases from 4.0% to 3.6% for *Database A* and this decreases from 4.9% to 3.4% for *Database B* as shown in Figures 3.25 and 3.26.

Table 3.5: EER for six parameters with variations in angular resolution

Angular Resolution	***Database A*** **EER (in %)**	***Database B*** **EER (in %)**
250	4.0	4.9
500	4.0	4.7
750	3.9	4.6
1000	3.6	3.4

These two experiments based on different sets of statistical features again emphasize that there is an improvement in FAR, FRR and EER, as the number of statistical parameters for feature extraction from iris images is increased from three to six. Proposed statistical feature extraction based iris recognition system along angular direction achieves an EER of 3.6% for *Database A* and of 3.4% for *Database B*.

It is worth mentioning here that performance of the proposed system improves with increased number of features but the database becomes large to handle. Increasing angular/radial resolution also improves the system performance, but at the expense of large size of feature vector set. In this work, experiments have also been conducted by taking radial resolution as 400 and angular resolution as 2000. It has been found that the EER improves by a maximum of 0.02% when the radial resolution is taken as 400 and angular resolution as 2000 for different experiments.

We implemented iris recognition system using statistical feature extraction technique proposed by Ko *et al*. (2006) and Kyaw (2009) in the same environment. We also implemented 1-*D* log-Gabor wavelet filter method (Masek, 2003) for feature extraction in the same environment.

In the proposed technique the length of iris code (number of bits) is *number of features* $\times$ *radial resolution* in case of feature extraction along concentric circles and it is *number of features* $\times$ *angular resolution* in case of feature extraction along the angular direction. Therefore, maximum length of iris code obtained for experiment along radial direction is $6 \times 200 = 1200$ bits and that along angular direction is $6 \times 1000 = 6000$ bits. While, the length of iris code in case of 1-*D* log-Gabor wavelet filters technique is $2 \times$ *radial resolution* $\times$ *angular resolution* (Masek, 2003). Therefore, even the normalized iris pattern of size 20×240 would produce iris code of length 9600 bits. Proposed approach creates a compact 150-byte (radial direction) or 750-byte (angular direction) size template as compared to 1200-byte size template in case of 1-*D* log-Gabor wavelet filter technique, which allows for efficient storage and comparison of irises.

Table 3.6 contains the comparison of proposed approach with these algorithms in terms of most effective EER, iris code length and iris code creation time. Experiments with different feature extraction algorithms have been carried out in the same computing environment. The time taken by the computing device from first stage of iris recognition system, *i.e.*, preprocessing to generation of binary iris template/code is taken as iris code creation time. Table 3.6 shows that the performance of proposed iris recognition system based on statistical feature extraction technique is comparable, effective and encouraging.

Table 3.6 reveals that for *Database B* proposed approach along radial direction gives minimum EER and for *Database A* EER of proposed approach along radial direction is valuable but slightly less than 1-*D* log-Gabor wavelet technique. Here, it is worth mentioning that the proposed statistical feature extraction based iris recognition creates a compact iris code, requires less memory to store the database and takes less time to compute iris code as compared to other mentioned techniques.

Table 3.6: Comparison of proposed approach with existing techniques

Feature Extraction Technique	***Database A* EER (in %)**	***Database B* EER (in %)**	**Iris Code Length (bits)**	**Iris Code Creation Time (ms)**
Statistical without normalization	19.3	22.0	120	210
Cumulative sum based change analysis	1.5	2.7	320	204
1-*D* log-Gabor wavelet	1.2	1.5	9600	204
Proposed approach (angular direction)	3.6	3.4	6000	178
Proposed approach (radial direction)	1.3	1.5	1200	134

Chapter Summary

In this chapter, implementation and comparative analysis of iris recognition systems with and without normalization is carried out, which shows that both methods have their own advantages and disadvantages. The algorithmic complexity and memory requirement of iris recognition system without normalization is less than method with normalization. However, the iris recognition system with normalization has significantly improved EER of 1.5% and 2.7% as compared to 19.3% and 22.0% for *Database A* and *Database B*, respectively.

Further, a statistical feature extraction based iris recognition system has also been proposed and implemented. It is demonstrated that statistical features can be computed along radial and angular directions. It is observed that system performance improves with increased number of statistical features. Results also show that increased radial or angular resolution while normalization in place improves the accuracy of iris recognition system. The most effective error rate achieved in the experiments

conducted in this work is 1.3% when 200-bit feature vector along radial direction is considered and is 3.4% when 1000-bit feature vector along angular direction is considered. Results also show that statistical feature extraction based system creates compact iris code and takes less time for feature extraction when compared with other systems present in literature.

CHAPTER 4

IRIS RECOGNITION SYSTEM FOR GENDER AND DIABETES PREDICTION

4.1 Overview

Iris recognition system can be utilized for clinical applications other than its conventional application in authentication of an individual. A good number of researchers (Lagree and Bowyer, 2011; Lai and Chiu, 2010; Lesmana *et al.*, 2011; Ma and Li, 2007; Othman and Prabuwono, 2010; Ramlee and Ranjit, 2009; Stearn and Swanepoel, 2007; Wibawa and Purnomo, 2006) have utilized iris images to determine the health status of an individual. In Chapter 3, implementation of iris recognition system has been discussed. In this chapter, two clinical applications of iris recognition system to predict the gender of imposter and to predict diabetes in human beings are proposed.

Iris recognition system in combination with iridology can be used to design an automated computer model to determine the status of health of an individual. There are five stages in iris recognition based proposed model as shown in Figure 4.1. These stages have been discussed below, in brief.

- ***Iris image acquisition***

  In this work, I-SCAN-2 dual iris scanner of Cross Match Technologies Inc. (http://www.crossmatch.com/i-scan-2.php) has been used for capturing the eye image. Front view of I-SCAN-2 is shown in Figure 4.2. It uses near-infrared illumination (700-900 nm) and produces 480×480 bmp images. Image of both eyes are captured during single scan.

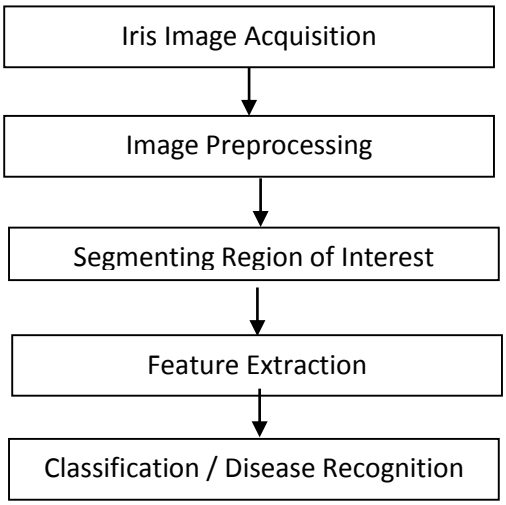


Figure 4.1: Five different stages of proposed model

Figure 4.2: I-SCAN-2 dual iris scanner

- ***Image preprocessing***

  Image preprocessing, as discussed in Section 2.1.1, has been used to obtain the iris region from an eye image and to remove the noise due to eyelids/eyelashes and reflection.

- ***Segmenting region of interest***

  Iris is like a map of the body. Changes in the condition of any organ of the body can be analyzed through the changes produced in the segments described by iridology charts (http://www.bernardjensen.com/Charts_c_17.html). It is important to extract significant features to analyze from the region associated with the particular organ. Hence, specific Region of Interest (ROI) is segmented from the normalized iris image.

- ***Feature extraction***

  Feature extraction (Zahedi and Zahara, 2013; Zahara, 2012) is most important task in applications employing image processing. Iris texture consists of large amount of blood vessels and nerves. Depending upon the health of an individual certain changes like holes, fractures, hyper pigmentations, radii solaris, nerve rings, lines and color of iris (Ma and Li, 2007; Othman and Prabuwono, 2010; Lai and Chiu, 2010) are observed in the iris texture. As discussed in Chapter 2, various algorithms have been proposed by different researchers for extracting significant features from an iris image. A feature vector created from these features is used for classification.

- ***Classification / Disease recognition***

  To classify a healthy / unhealthy subject (or recognize a disease), SVM has been used as a classifier in this model. In SVM, a hyper-plane is constructed as the decision surface in such a way that the margin of separation between positive and negative examples is maximized (Cristianini and Shawe, 2000; Haykin, 2004). SVM employs kernel based learning algorithm. The effectiveness of SVM depends upon the selection of the kernel and the kernel parameters. Some of the kernel functions are linear, polynomial and Radial Basis Function (RBF).

4.2 Predicting Gender Using Iris Images

Human beings can recognize gender very easily as compared to a machine. Gender classification is an important problem of computer vision. Gender recognition is useful in the area of forensic sciences, security systems and many more. Recently, some researchers (Chu *et al.*, 2010; Han *et al.*, 2009; Nazir *et al.*, 2010; Rai and Khanna, 2010) have made the most of facial images to classify gender whereas a few studies (Lagree and Bowyer, 2011; Thomas *et al.*, 2007) of this nature have utilized the properties of iris images to predict the gender with success rate of 62.0% and 82.0% respectively.

Gender classification/prediction using iris image model involves five steps as shown in Figure 4.1. Here, iris image acquisition is carried out using I-SCAN-2 dual iris scanner. Further, image preprocessing is carried out in the manner similar to iris recognition system and explained in Chapter 2. For gender prediction, whole normalized iris image is taken. Hence, there is no need for segmenting ROI. Gender prediction model has

been implemented using two different feature extraction techniques, namely, combination of statistical features and wavelet transform (Model-I), and Otsu's threshold method (Model-II). The two techniques have been discussed briefly in the following sub-sections. Thereafter, SVM based classifier has been developed to predict the gender of subject as male or female.

4.2.1 Using statistical features and wavelet transform

In this work, feature vector for an iris image is obtained by combining two different feature extraction techniques, namely, statistical and wavelet transform.

- ***Statistical features***

  Statistical features have been computed along two different directions, *i.e.*, angular direction and radial direction. Top-most row of 2-*D* array represents a circular ring near to iris-pupil boundary and bottom-most row represents a circular ring near to iris-sclera boundary. Moving from top to bottom row indicates moving along radial direction. Statistical features computed along each row correspond to features computed along virtual circles. Similarly, statistical features have been computed along each column of 2-*D* array that correspond to angular direction. Statistical features that have been considered in this work are mean, median and standard deviation. Statistical features computed in this work are different from the features computed by Thomas *et al*. (2007). They computed mean, standard deviation and variance on real part of a log-Gabor-filtered normalized iris image whereas in this research work, statistical features have been computed using pixel values of normalized iris image.

  The statistical computation gives a feature vector *F* for an image, which is denoted as:

$$F = (\bar{X}_r, Md_r, s_r, \bar{X}_c, Md_c, s_c) \tag{4.1}$$

  where $\bar{X}_r$ is a row vector of mean computed along each row, Md_r is a row vector of median computed along each row, s_r is a row vector of standard deviation computed along each row, $\bar{X}_c$ is a row vector of mean computed along each column, Md_c is a row vector of median computed along each column, s_c is a row vector of standard deviation computed along each column.

Size of one statistical feature computed along each row would be 100 and size of one statistical feature computed along each column would be 500 for a 100×500 sized normalized iris image. Thus, size of statistical feature vector *F* for an iris image obtained is 1800.

- ***Wavelet transform***

  As already discussed, another set of features is extracted from an iris image using 2-*D* Discrete Wavelet Transform (DWT). DWT decomposes an image into four sub-sampled images. The original image of size $N \times N$ is split into four sub-images each of size $N/2 \times N/2$ and contains information from different frequency components. After decomposition four sub-sampled images are approximation (LL), horizontal (HL), vertical (LH) and diagonal (HH). It can be noted that

  (i) HH is an image that has been high passed in the horizontal and vertical directions.

  (ii) LH is an image that has been low passed in the vertical and high passed in the horizontal direction.

  (iii) HL is an image that has been high passed in the vertical and low passed in the horizontal direction.

  (iv) LL means low pass filtered in both directions.

  Using DWT an image can be decomposed more than once. There are mainly two ways for decomposition, namely, pyramidal and packet decomposition. In case of pyramidal decomposition, further decompositions are applied only to the approximation (LL) sub-band. At each level, the approximation sub-band is further decomposed whereas, in case of packet decomposition, the decomposition is not limited to LL sub-band only rather it allows further decomposition of all sub-bands at each level. Figure 4.3 depicts three level pyramidal decomposition and Figure 4.4 shows two level packet decomposition.

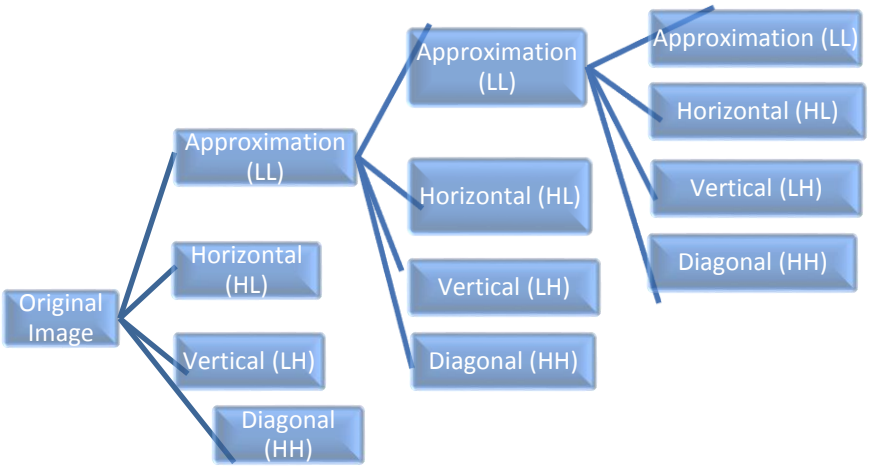


Figure 4.3: Three level pyramidal decomposition

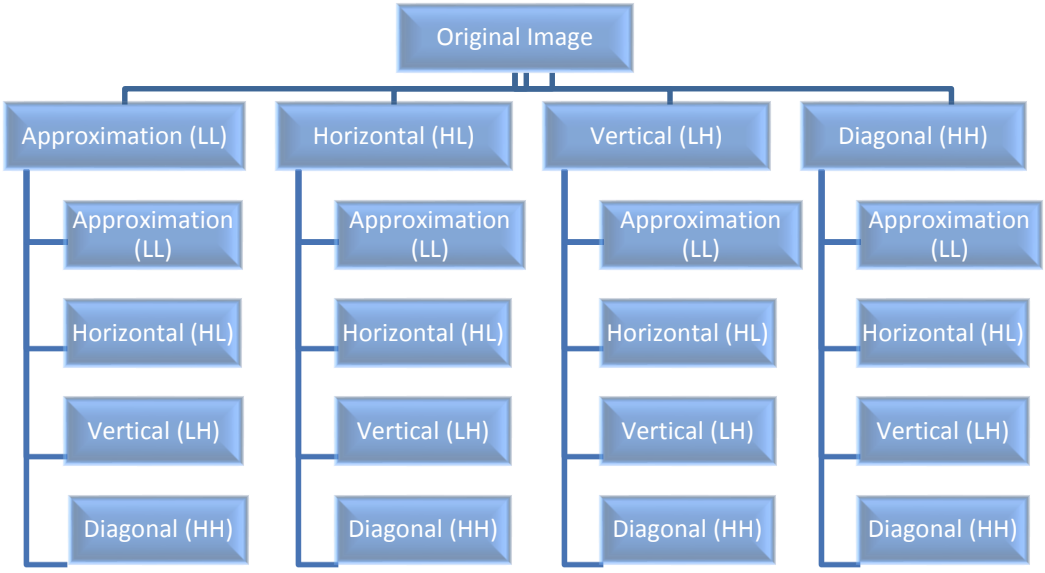


Figure 4.4: Two level packet decomposition

One can utilize different types of possible wavelet basis function for feature extraction. Here, three level pyramidal decomposition using Daubechies db2 wavelet basis function has been considered for extracting features.

The wavelet function for Daubechies db2 is given as:

$$\psi(x) = \sqrt{2}\sum_{k=0}^{3}(-1)^k h_{3-k}\varphi(2x-k) \tag{4.2}$$

where $\varphi(x)$ is a scaling function given as

$$\varphi(x) = \sqrt{2}\sum_{k=0}^{3} h_k \varphi(2x-k) \tag{4.3}$$

and $h_0, h_1, h_2\ and\ h_3$ are filter coefficients satisfying

$$h_0 + h_2 = \frac{1}{\sqrt{2}} = h_1 + h_3 \tag{4.4}$$

and $$\sum_{k=2l}^{3+2l} h_k h_{k-2l} = \begin{cases} 1 & for\ l = 0 \\ 0 & for\ l = 1 \end{cases} \tag{4.5}$$

Feature vector may be created by considering different possible combinations of LL, LH and HL components obtained at different level of decomposition. Here, an attempt has been made to study the effect of different combinations of LL, HL and LH components obtained after third level of decomposition on system performance. Therefore, with 100×500 sized normalized iris image the length of feature vector is 819/1638/2457 considering 1/2/3 sub-band components.

The size of the overall feature vector created by combining both statistical features and wavelet features is, therefore, 2619/3438/4257 when 1/2/3 sub-bands are considered. These features have been further used for predicting the gender of subjects using SVM classifier.

4.2.2 Using Otsu's threshold method

In this work, Otsu's threshold (Otsu, 1979) is selected automatically from gray level histogram of the normalized iris image. Based upon the set of training images from the dataset an optimum threshold is determined. Now, test image's Otsu threshold is computed and compared with the optimum threshold for predicting gender.

4.3 Dataset for Predicting Gender

A dataset of 400 iris images is created for the application of gender prediction. Iris images of 200 subjects (100 Male and 100 Female) are captured. For experimentation, both left and right iris images are considered. The subjects are of age group 17-35.

4.4 Results of Gender Prediction

SVM based gender prediction model has been implemented using image processing module of MatLAB 7.1. In Model-I, as discussed earlier, two different feature extraction techniques, namely, statistical and wavelet transform have been combined to

generate the feature vector. To validate the experiments, 10-fold cross validation technique has been employed where complete dataset is divided into 10 equally sized groups. Each group is used for testing once while the other nine groups are used to train the classifier. Hence, there are 10 iterations for training the SVM and testing. The accuracy of the proposed model is measured using Correct Classification Rate (CCR) independently for all iterations. The overall accuracy of the proposed model is the mean of accuracies obtained from individual iterations. Experiments have been conducted by selecting three different kernel functions, namely, polynomial (with degree 2), Gaussian and RBF for SVM.

Tables 4.1, 4.2 and 4.3 illustrate the accuracy of the proposed gender classification model for polynomial, Gaussian and RBF kernel functions, respectively. Overall accuracy and accuracy of individual iteration for a specific feature vector is given. It is observed from Table 4.1 that CCR varies from 79.0% to 85.8% for polynomial kernel function. The maximum overall accuracy of the model is 83.7% for feature vector generated as a combination of LL and HL sub-band wavelet coefficients along with statistical features. Similarly, it is seen from Table 4.2 that CCR of the proposed gender classification model varies from 80.0% to 87.4% with Gaussian kernel function. The maximum overall accuracy obtained is 85.6% for feature vector generated by combining statistical features with LL and HL sub-band components.

Table 4.1: Accuracy for different feature vectors with polynomial kernel function

Feature Vector/Iteration	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Overall (in %)**
Statistical + LL	80.2	79.0	81.0	84.0	82.8	82.8	83.7	81.1	80.8	83.6	81.9
Statistical +HL	79.8	79.6	80.0	81.4	79.8	80.6	81.8	82.6	82.8	81.6	81.0
Statistical +LH	80.0	80.2	81.1	82.6	80.2	82.0	82.4	83.6	83.2	82.8	81.8
Statistical +LL+HL	82.6	81.8	82.5	84.6	82.5	83.2	84.6	85.2	85.0	85.8	**83.7**
Statistical +LL+LH	81.2	82.0	81.6	82.2	82.2	81.2	82.2	81.0	81.6	82.8	81.8
Statistical +HL+LH	80.6	80.2	79.9	81.0	80.2	80.6	81.6	82.2	82.4	82.0	81.0
Statistical +LL+HL+LH	79.8	80.1	79.8	80.8	81.0	82.2	82.8	81.2	80.8	82.8	81.1

Table 4.2: Accuracy for different feature vectors with Gaussian kernel function

Feature Vector/Iteration	1	2	3	4	5	6	7	8	9	10	Overall (in %)
Statistical + LL	82.2	81.2	83.2	85.8	84.8	85.0	85.8	83.6	83.0	85.6	84.0
Statistical +HL	81.6	81.0	82.4	83.8	81.6	82.6	83.6	85.0	84.6	83.6	82.9
Statistical +LH	82.2	80.8	83.8	84.6	82.0	84.2	84.2	85.8	85.4	83.6	83.6
Statistical +LL+HL	84.6	84.0	84.8	86.2	84.2	85.0	86.8	87.4	86.8	87.0	**85.6**
Statistical +LL+LH	83.1	84.4	83.8	84.6	84.0	83.4	84.4	82.8	83.2	84.2	83.7
Statistical +HL+LH	81.8	82.0	82.0	83.4	82.0	82.8	83.8	84.2	84.6	83.8	83.0
Statistical +LL+HL+LH	81.6	81.8	80.0	82.0	83.2	84.6	85.0	82.8	83.0	84.6	82.8

With RBF kernel function, the system accuracy varies from 79.8% to 86.8% and the maximum overall accuracy obtained is 84.4% for feature vector generated by combining statistical features with LL and HL sub-band components as shown in Table 4.3.

Table 4.3: Accuracy for different feature vectors with RBF kernel function

Feature Vector/Iteration	1	2	3	4	5	6	7	8	9	10	Overall (in %)
Statistical + LL	81.8	79.8	82.8	85.0	83.6	84.2	85.8	83.2	82.2	83.6	83.2
Statistical +HL	81.0	80.8	81.6	83.2	81.2	80.8	83.6	83.8	83.2	82.0	82.1
Statistical +LH	81.8	80.2	82.4	83.6	81.6	82.8	84.2	84.0	83.8	82.2	82.6
Statistical +LL+HL	83.2	83.0	83.0	84.6	83.6	83.6	86.8	85.2	85.2	86.2	**84.4**
Statistical +LL+LH	82.2	83.6	83.6	83.0	83.0	83.0	84.4	81.0	82.8	83.4	83.0
Statistical +HL+LH	80.4	81.8	81.4	83.2	81.2	82.6	83.8	83.2	84.0	82.6	82.4
Statistical +LL+HL+LH	80.8	80.6	79.8	81.2	82.8	84.0	84.2	81.8	82.8	84.0	82.2

The overall accuracy of the proposed model for different length feature vector with distinct kernel functions is shown in Figure 4.5 in bar chart form. From this it is evident that maximum efficiency of 85.6% is obtained for Gaussian kernel function with feature vector generated by combining LL and HL components of decomposed image along with statistical features.

System performance has also been measured in terms of specificity (1 - False Positive Rate) and sensitivity (1 – False Negative Rate) for each possible combination of different feature vectors and different kernel function. The most effective specificity of the proposed model is 0.93 and the most effective sensitivity of the proposed model is 0.94.

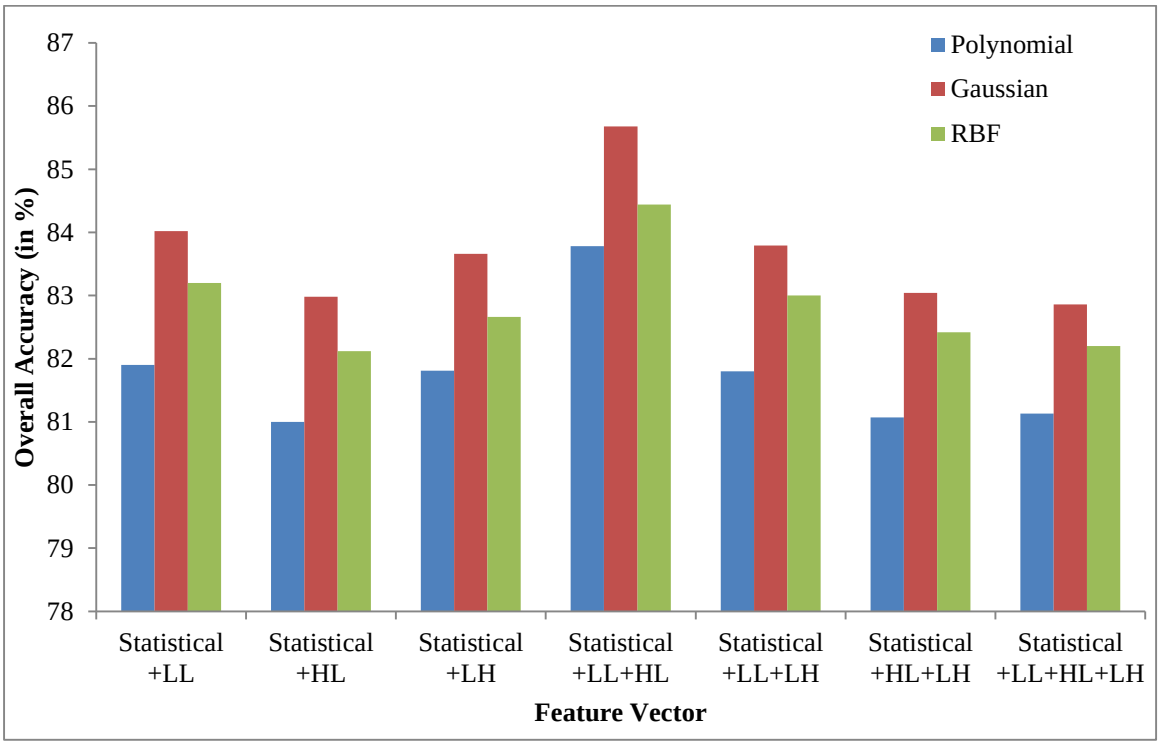


Figure 4.5: Accuracy of proposed gender prediction model

In Model-II, the gender prediction model has been implemented using Otsu's threshold method (Otsu, 1979). As discussed earlier, a threshold from gray level histogram of normalized iris image is computed. Using the set of training images an optimum threshold of 128 is obtained. Next, the test image threshold is computed and compared with 128. If value is greater than 128 the subject is considered as female and if less than 128, it is considered as male. An overall accuracy of 68.0% is achieved with this model.

The system performance is lower than the model proposed above using statistical and wavelet features.

Table 4.4 illustrates the comparison of proposed iris recognition system based gender prediction model with existing iris image based gender prediction models. Results show that the proposed model is effective and has better accuracy.

Table 4.4: Comparison of accuracy of gender prediction model with existing techniques

S. No.	Feature Extraction Technique	Classifier	Number of Subjects	Accuracy (in %)
1.	Geometric and log-Gabor filter (Thomas *et al.*, 2007)	C4.5 decision tree	300	80.0
2.	Texture filters (Lagree and Bowyer, 2011)	SVM	60	62.0
3.	Proposed Model-I (using wavelets and statistical)	SVM	200	85.6
4.	Proposed Model-II (using Otsu's threshold)	SVM	200	68.0

4.5 Predicting Diabetes Using Iris Images

Diabetes is a group of metabolic diseases in which a person has high blood sugar, because the body does not produce enough insulin or cells do not respond to the insulin produced. There are mainly three types of diabetes, namely, Type-I, Type-II and gestational diabetes (http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0002194). In Type-I diabetes, human body produces negligibly small insulin. It is diagnosed in children, teens or even in adults. It is often referred as Insulin Dependent Diabetes Mellitus (IDDM). In Type-II diabetes, human body is not able to use the produced insulin efficiently. It is non-insulin dependent. It is often diagnosed in adults mainly after the age of 40. However, nowadays it is also diagnosed in teens and young adults because of eating habits and high obesity rates. Gestational diabetes normally occurs in pregnant women, who have never been diabetic before but have a high blood glucose level during pregnancy. Normal symptoms of high blood sugar include blurred vision, excess thirst, frequent urge to urine, fatigue, hunger and weight loss. Pancreas is associated with diabetes mellitus due to its function of insulin secretion. The pancreas

is a fish-shaped, grayish pink gland about 5 inches long that stretches across back of the abdomen, behind the stomach. The state of pancreas extracted from iris images have been used in this work for developing a diabetes prediction model.

The diabetes prediction model is similar to the one shown in Figure 4.1. Here, a dataset of iris images of healthy subjects and one suffering from diabetes is created using I-SCAN-2 dual iris scanner. The preprocessing stage is implemented similar to the one in gender prediction, *i.e.,* using CHT, homogeneous rubber sheet model and histogram equalization. Next, important step is segmenting ROI. The pancreas is divided into head, body and tail. Iridology charts (http://www.bernardjensen.com/ charts_c_17.html) show head of the pancreas in right eye between 7 and 8 o'clock. These charts also depict the body of pancreas in left eye between 7 and 8 o'clock, and tail of pancreas between 4 and 5 o'clock in the same eye, as shown in Figure 1.3 and Figure 1.4. The cells that control blood sugar level, called islets, are concentrated somewhat in the tail of pancreas. In this work, experiments have been conducted by considering left eye only. ROI of size 200 × 90 is segmented out from the normalized iris image as shown in Figure 4.6.

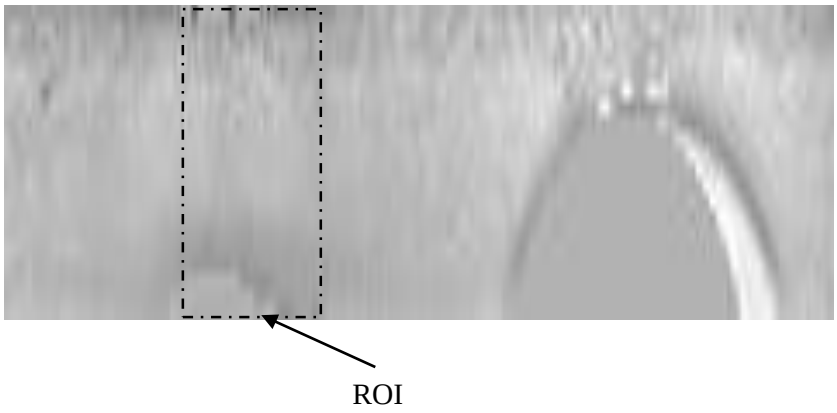


Figure 4.6: Segmenting ROI

Once the ROI is separated from the iris region, the significant features are extracted by applying two different feature extraction techniques, namely, wavelet transform and Gabor filter.

4.5.1 Using wavelet transform

Features from the normalized iris image are obtained by applying 2-*D* DWT. In this work, two different approaches using wavelet transform have been proposed and implemented.

- **Approach 1**

  In this case, the pyramidal decomposition of 2-*D* normalized iris image up to third level using Daubechies db2 wavelet has been implemented. An attempt has been made to study the effect of different wavelet coefficients by generating feature vector using different combinations of LL, HL and LH components.

- **Approach 2**

  In this approach, again the pyramidal decomposition with Daubechies db2 wavelet function is implemented for feature extraction. Here, an attempt has been made to study the effect of level of decomposition. Feature vector is obtained by considering only LL component at each level of decomposition, *i.e.*, LL1 for level one decomposition, LL2 for level two decomposition and so on up to fifth level. At the same time, combination of LL components of different levels is also considered.

4.5.2 Using Gabor filter

In this work, iris texture features have been extracted using 2-*D* Gabor filter. The 2-*D* Gabor filter in spatial domain is expressed as:

$$\left.\begin{aligned} h(x,y) &= h'(x,y)\exp(j2\pi\xi) \\ h'(x,y) &= \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \\ \xi &= x.cos(\theta) + y.\sin(\theta) \end{aligned}\right\} \quad (4.6)$$

where $(x,\ y)$ is coordinates of spatial region, $\sigma$is the standard deviation w represents frequency component and θ is direction parameter of filter. Iris feature vector is created by combining 12 texture energy features computed along four different directions and with three different frequencies. Initially, for each filter channel $(w_i,\ \theta_j)$ $(i = 1,2,3)\ and\ (j = 1,2,3,4)$ the segmented pancreas region of iris image $I(x,y)$ is

convolved with Gabor filter $h(x, y)$ to obtain an image $S(x, y)$in frequency domain. The energy for each channel is then computed as:

$$e(w_i, \theta_i) = \sqrt{Re(S)^2 + Im(S)^2} \tag{4.7}$$

Where $Re\ (S)$ and $Im\ (S)$ are the real and imaginary part of image $S(x, y)$. The feature vector so obtained is as shown in (4.8).

$$F = \begin{bmatrix} e(w_1, \theta_1), e(w_1, \theta_2), e(w_1, \theta_3), e(w_1, \theta_4), \dots \\ \dots, e(w_3, \theta_1), e(w_3, \theta_2), e(w_3, \theta_3), e(w_3, \theta_4) \end{bmatrix} \tag{4.8}$$

4.6 Dataset for Predicting Diabetes

A dataset of iris images of healthy people and those suffering from diabetes has been created. Iris image dataset of 80 subjects; 40 diabetic (25 Male, 15 Female) and 40 healthy people (22 Male, 18 Female) has been used as depicted in Figure 4.7.

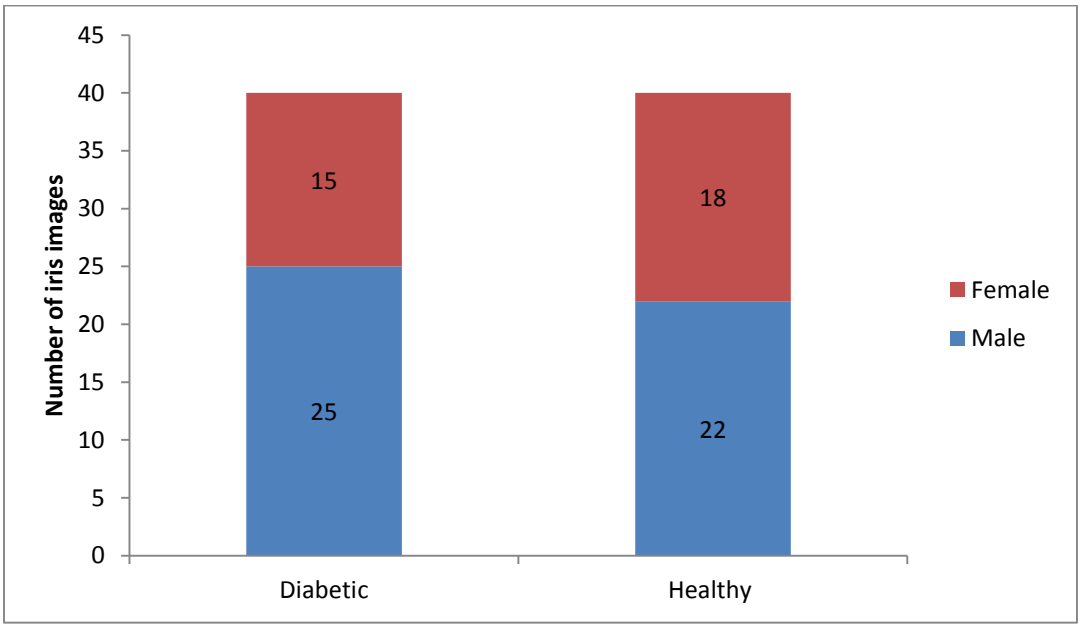


Figure 4.7: Distribution of dataset

Care has been taken to consider the different age group people while taking the images of diabetic people. Figure 4.8 illustrates that 45.0% (18 out of 40) of total diabetic people are of the age group 21-40 years while 37.5% (15 out of 40) are of age group 41-60 years and remaining 17.5% (7 out of 40) people are from 61-80 years age group. An informed written consent of each person has been taken for this research.

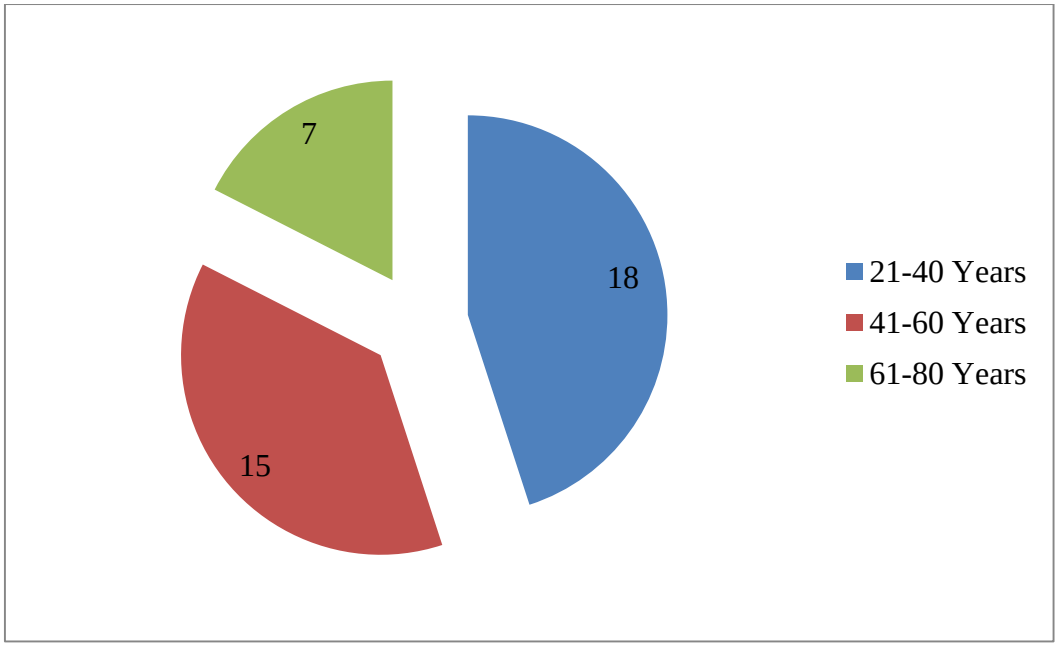


Figure 4.8: Age wise distribution of diabetic dataset

Only diagnosed cases of Type-II diabetes have been considered for the study. Figure 4.9 depicts the duration of diabetes wise distribution of subjects. The subjects vary from 2 months of diabetes to 20 years of diabetes. The mean duration of diabetes was 7.4 years.

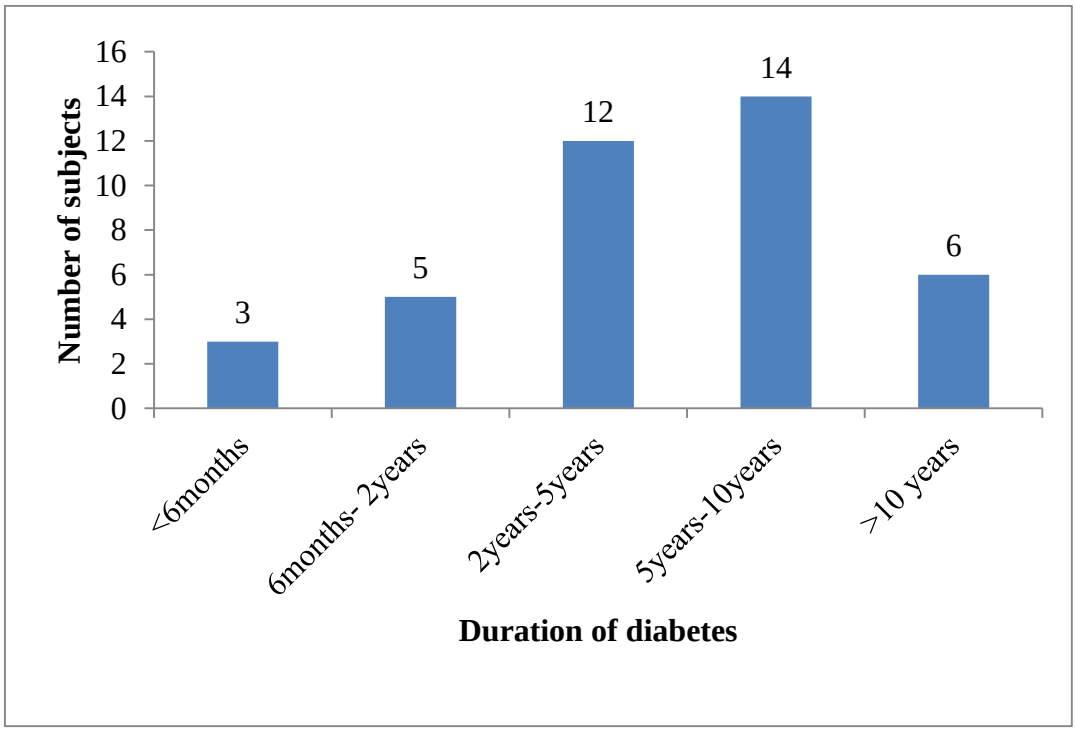


Figure 4.9: Duration of diabetes wise distribution of dataset

Subjects have once a week to monthly frequency of measurement of blood glucose level. Figure 4.10 illustrates the level of diabetes control in subjects. Out of 40 subjects, only 3 subjects had good control diabetes with proper medication, diet and exercise. The blood glucose level of subjects ranges from 60mg/dL to 160 mg/dL (fasting) and 106mg/dL to 360 mg/dL (pp). Only 15 subjects knew their *Hba1c* value and it was in the range of 6 to 8.5.

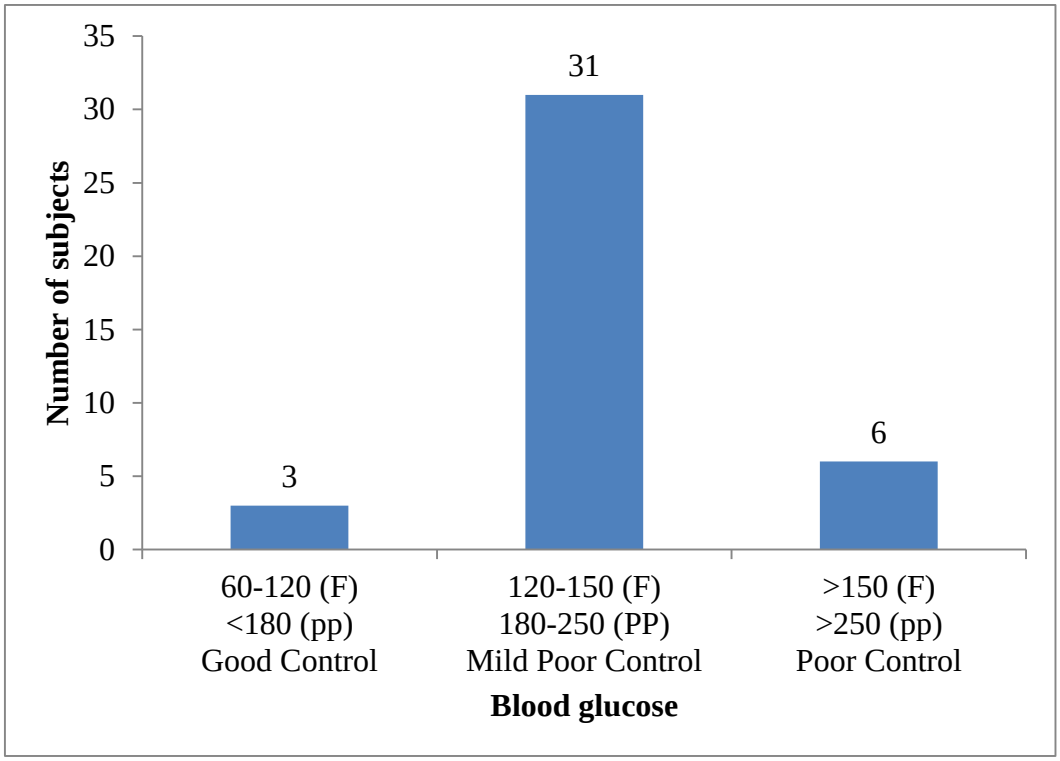


Figure 4.10: Level of diabetes control wise distribution of dataset

Out of 40 diabetic subjects, one subject had a cardiovascular disease (gone through angioplasty); five subjects had a problem of hypertension; two had one eye surgery while others do not have any comorbid condition in them.

4.7 Results of Diabetes Prediction

Iris recognition based method to predict diabetes has been proposed and implemented using two different techniques, namely, wavelet transform and Gabor filter. In both approaches, classification of subjects into healthy one or one with diabetes is done using SVM.

4.7.1 Results using wavelet transform

Two approaches have been proposed using wavelets. In approach 1, an attempt has been made to study the effect of different wavelet coefficients on system accuracy. Feature vector has been generated using different combinations of LL, HL and LH components. The effect of kernel function on accuracy of the proposed model is also studied. Experiments have been conducted by selecting three different kernel functions, namely, polynomial, Gaussian and RBF for SVM.

Experiments have been validated using 4-fold cross validation technique. The overall accuracy is the mean of four iterations. Table 4.5 depicts the comparison of overall accuracy obtained for the iris recognition based diabetes prediction model for different combinations of kernel function and feature vectors.

Table 4.5: Overall accuracy of diabetes prediction model for different kernel functions with different feature vectors

Feature Vector	Gaussian (in %)	RBF (in %)	Polynomial (2nd degree) (in %)	Polynomial (3rd degree) (in %)
LL	81.3	81.3	76.3	75.5
HL	76.3	81.3	75.0	77.5
LH	76.3	80.0	75.0	75.0
LL + HL	86.3	**87.5**	81.3	77.5
LL + LH	85.0	86.3	81.3	76.3
HL + LH	81.3	85.0	80.0	78.6
LL + HL + LH	81.3	81.3	77.5	78.6

The accuracy of the proposed model is also shown in Figure 4.11 in bar chart form. From this, it is evident that maximum accuracy of 87.5% is obtained for RBF kernel function with feature vector generated by combining LL and HL components of decomposed image.

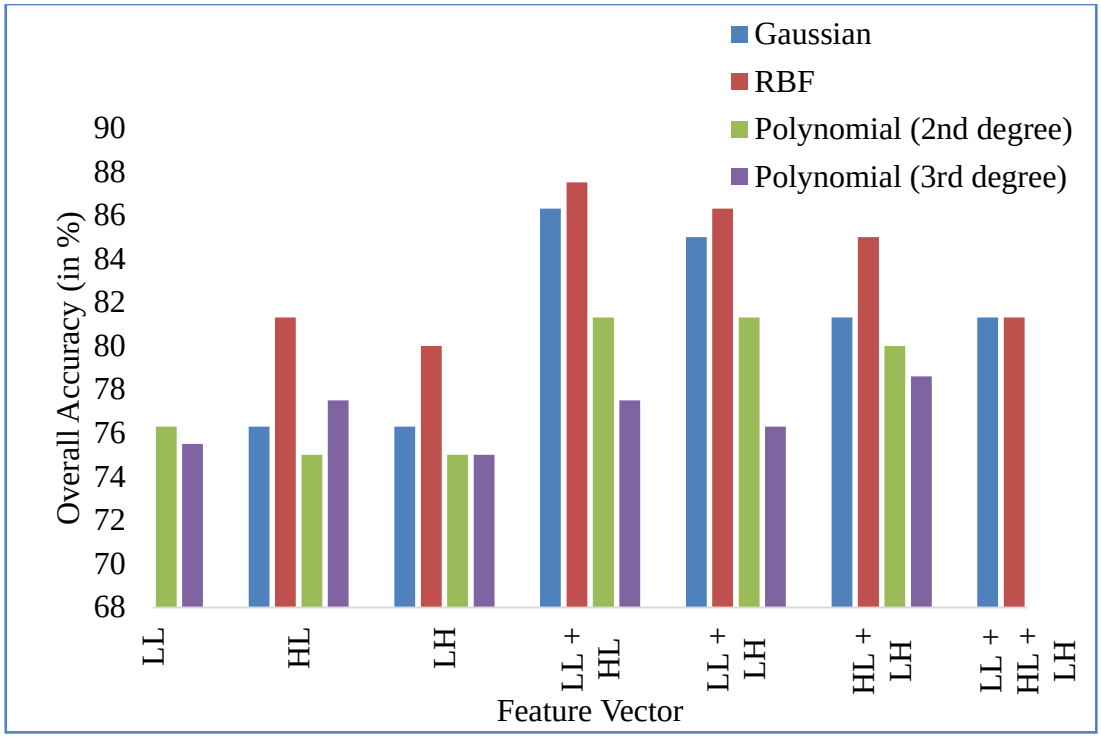


Figure 4.11: Accuracy of proposed diabetes prediction model

In approach 2, an attempt has been made to study the effect of level of decomposition on the system performance. Experiments have been conducted by considering three different kernel functions, namely, polynomial, RBF and sigmoid. Again, 4-fold cross validation technique has been applied to validate the experiments. The overall accuracy of the system is computed by taking mean of four iterations.

To study the effect of level of decomposition, feature vector is generated by considering either only LL component obtained after each level of decomposition or various possible combinations of LL components obtained after each level of decomposition. Total number of possible combinations, considering single LL component and combinations of two, three, four or five LL components, is 31.

Table 4.6 illustrates the accuracy of proposed method for predicting diabetes. From this table, one can infer that as the level of decomposition is increased from level 1 to level 5 the accuracy of the proposed system increases for a particular kernel function of SVM. Results have been obtained for level 6 and level 7 decomposition also. It has been observed that further decomposition beyond level 5 does not produce improvement in accuracy.

Table 4.6: Overall accuracy with different feature vectors and kernel functions for predicting diabetes

kernel Function / Feature Vector	Polynomial (2nd degree) (in %)	Polynomial (3rd degree) (in %)	RBF (in %)	Sigmoid (in %)
LL1	75.3	73.5	80.3	80.5
LL2	75.5	74.8	80.5	81.3
LL3	76.3	75.5	81.3	81.3
LL4	78.0	78.5	83.3	81.8
LL5	80.3	80.3	84.0	83.0
LL1 + LL2	77.5	80.3	80.5	80.3
LL1 + LL3	78.0	80.3	80.5	80.3
LL1 + LL4	78.0	81.0	81.3	81.0
LL1 + LL5	79.3	81.0	81.3	81.0
LL2 + LL3	76.0	79.3	80.5	80.3
LL2 + LL4	75.3	77.5	80.5	80.3
LL2 + LL5	75.3	79.3	81.0	80.0
LL3 + LL4	76.0	80.0	80.3	80.5
LL3 + LL5	76.3	81.3	80.3	80.5
LL4 + LL5	76.3	81.3	81.5	80.8
LL1 + LL2 + LL3	80.0	82.0	84.0	82.0
LL1 + LL2 + LL4	80.3	82.3	85.3	82.3
LL1 + LL2 + LL5	80.3	82.3	85.3	83.3
LL1 + LL3 + LL4	80.3	82.3	85.0	83.0
LL1 + LL3 + LL5	81.0	82.3	85.0	83.0
LL1 + LL4 + LL5	81.3	82.5	85.3	83.3
LL2 + LL3 + LL4	80.0	81.0	84.0	82.0
LL2 + LL3 + LL5	80.0	81.3	84.3	82.5
LL2 + LL4 + LL5	80.3	81.3	84.0	82.0
LL3 + LL4 +LL5	81.3	81.3	84.0	82.0
LL1 + LL2 + LL3 + LL4	83.3	83.0	87.3	84.3
LL1 + LL2 + LL3 + LL5	83.5	83.3	87.0	84.5
LL1 + LL2 + LL4 + LL5	84.0	83.0	87.0	84.3
LL1 + LL3 + LL4 + LL5	83.3	83.3	87.0	84.5
LL2 + LL3 + LL4 + LL5	83.0	83.5	86.8	84.0
LL1 + LL2 + LL3 + LL4 + LL5	84.3	84.5	**88.3**	85.3

Table 4.6 reveals that the performance of the system improves as the feature vector is considered as combination of LL components. The maximum accuracy is achieved by considering RBF kernel function for a particular feature vector. It is evident from the results that combination of LL1, LL2, LL3, LL4 and LL5 along with RBF function gives the best performance. The maximum accuracy of 88.3% is achieved in predicting diabetes.

4.7.2 Results using Gabor filter

In this approach, segmented pancreas region from 2-*D* normalized iris images is convolved with Gabor filter to extract the significant features. The feature vector is obtained as computed in (4.8). In this approach, overall maximum accuracy obtained with RBF kernel function is 82.1%.

4.7.3 Comparison of proposed models with existing techniques

Table 4.7 illustrates the comparison of proposed iris recognition based models to predict diabetes with existing irido-diagnostic models. The obtained accuracy with higher sample size establishes the effectiveness of the system. The overall maximum specificity of the proposed model to predict diabetes is 0.94 and maximum sensitivity is 0.95.

It has been observed that proposed model does not work well in case when patient is having operated eyes. Two subjects with operated eye and having diabetes have been identified as healthy. Another significant observation from this study is that a person having controlled diabetes for last 3 years with proper medicines, diet and exercise has also been identified as healthy.

Table 4.7: Comparison of accuracy of proposed diabetes prediction model with existing techniques

S. No.	Disease	Feature Extraction Technique	Classifier	No. of Samples	Accuracy (in %)
1.	Detecting broken tissues in pancreas (Wibawa and Purnomo, 2006)	Minimum filter	Visual inspection	34	94.0
2.	Alimentary canal (Ma and Li, 2007)	Gabor filter	SVM	53	84.9
3.	Nerve system (Ma and Li, 2007)	Gabor filter	SVM	44	86.4
4.	Hearing loss (Stearn and Swanepoel, 2007)	Digital image enhancement	Visual inspection	53	70.0
5.	Pancreas disorder (Lesmana *et al.*, 2011)	GLCM	Neighborhood based modified backpropagation using adaptive learning parameters (ANMBP)	50	83.3
6.	Proposed model	Gabor filter	SVM	80	82.1
7.	Proposed model	Wavelet transform approach 1	SVM	80	87.5
8.	Proposed model	Wavelet transform approach 2	SVM	80	88.3

Chapter Summary

In this chapter, SVM based gender prediction model using iris images, that combines statistical features with 2-*D* DWT based features, has been proposed and implemented. Experiments have been conducted to study the effect of different combinations of LL, HL and LH components, obtained after third level of decomposition, along with same statistical features on system performance. An attempt has also been made to study the effect of kernel function on accuracy of the proposed model. Maximum accuracy of

85.6% obtained in this work is encouraging and demonstrates improved gender classification model as compared to the models proposed earlier.

A system as an aid to doctors to determine the diabetes using iris images has been proposed and implemented in this chapter. The significant features are extracted from a specific location in the iris image using 2-*D* wavelet tree decomposition and using Gabor filter. It has been found that the combination of coefficients as well as level of decomposition significantly affects the performance of system. SVM based binary classifier has been implemented to classify the subjects. The overall accuracy of 88.3%, maximum specificity 0.94 and maximum sensitivity 0.95 to determine diabetes with higher sample size is quite effective and encouraging.

CHAPTER 5

PREDICTING OBSTRUCTIVE LUNG DISEASE USING IRIS IMAGES

5.1 Overview

In Chapter 4, iris recognition system based models to predict gender, and to predict diabetes have been discussed. In this chapter, one more application of iris recognition system for predicting obstructive lung disease is proposed and implemented.

In human beings, lungs are the most essential respiratory organs. Their weakness affects respiration and leads to various obstructive lung disease such as bronchitis, asthma or even lung cancer. Spirometry is the widely used non-invasive test to assess how well lungs function and is helpful in diagnosing the obstructive lung disease (Pierce, 2005). It measures how much air one inhales, how much one exhales and how quickly one can inhale and exhale. Smoking is one of the reasons for developing obstructive lung disease. The probability of smokers developing obstructive lung disease is higher when compared to non-smokers (Lundback *et al.*, 2013; Vestbo, 2013; Yuh-chin and Huang, 2012; Min *et al.*, 2008).

In this chapter, experiments have been carried out on iris images to predict obstructive lung disease. These days computer aided diagnostic tools are becoming an important area of research in medical science (Kumar and Anand, 2006). Many researchers (Smith *et al.*, 2013; Wang *et al.*, 2010; Sivasankar *et al.*, 2012) have proposed various methods and models related to lung diseases. The objective of this work is to develop an application for predicting whether a subject is suffering from obstructive lung disease or not using one's iris images. This can help doctors further in taking decisions. For this purpose, the SVM based iris recognition system has been combined with iridology for implementing the proposed model.

The proposed model involves five stages as illustrated in Figure 4.1. As discussed earlier, I-SCAN-2 dual iris scanner has been used to capture iris images. Once the normalized iris image is obtained after preprocessing as shown in Figure 2.2 the ROI can be segmented using iridology charts.

As per the iridology charts (Figures 1.3 and 1.4), lungs of human body appear between 2 o'clock and 3 o'clock in left eye; and between 9 o'clock and 10 o'clock in right eye. To predict obstructive lung disease, ROI, *i.e.*, lungs region is segmented from the normalized iris image. In present work, the size of ROI is taken as 100×60. After segmenting ROI from iris image, the features can be extracted.

5.2 Feature Extraction

Two different models have been proposed and implemented for predicting obstructive lung disease in this work. In Model-I, the salient features from the segmented lung region of normalized iris image are extracted using 2-*D* wavelet tree and in Model-II, feature extraction is carried out using Gabor filter.

- ***Using wavelet transform***

  Here, 2-*D* DWT has been used to extract significant features from iris images. Features are extracted in a manner similar to the Approach 2 of wavelet transform proposed for method to predict diabetes in Chapter 4. Again, an attempt has also been made to study the effect of decomposition levels. Feature vector is obtained either by considering approximation coefficients after each level of decomposition or by possible combinations of these approximation coefficients. Results have been reported up to fifth level of decomposition. In this research work, experiments have been carried out using Daubechie's db2 wavelet function.

- ***Using Gabor filter***

  In this model, segmented lungs region from 2-*D* normalized iris images is convolved with Gabor filter to extract features. The feature vector is obtained as discussed in Chapter 4. The feature vector obtained is used for classifying the subjects into healthy one or one having obstructive lung disease.

5.3 Dataset

Iris images from 100 subjects have been captured using I-SCAN-2 dual iris scanner of Cross Match Technologies Inc. Spirometry test has been used to diagnose the obstructive lung disease in these subjects. This test uses two parameters, namely, Forced Expiratory Volume in 1 second (FEV_1) and Forced Vital Capacity (FVC). If the ratio of these parameters ($\boldsymbol{\rho}$ = FEV_1 / FVC) is less than 0.8 for a subject then one will have a high probability of suffering from obstructive lung disease. We have conducted the spirometry test on all 100 subjects considered in this study. Based upon spirometry test ratio, dataset has been classified into healthy subjects and the subjects suffering from obstructive lung disease. Table 5.1 contains the information on the subjects considered in this work.

Table 5.1: Information on the dataset

Health Status of Subjects	$\boldsymbol{\rho}$	**Total Number of Subjects**	**Number of Smokers**	**Number of Non-Smokers**
Suffering from obstructive lung disease	≤ 0.7	25	25	0
	$0.7 < \rho < 0.8$	24	22	2
Healthy	≥ 0.8	51	3	48

A questionnaire related to the health details, such as whether cough with sputum, smoker's cough, shortness of breath, history of TB, blood in sputum, was completed by the subjects. An informed written consent of each person has been taken for this research.

5.4 Results and Discussions

Experiments have been conducted using image processing toolbox of MatLAB 7.1. The two models proposed in this chapter have been developed using the dataset of 100 subjects. Experiments have been conducted by considering three different kernel functions for SVM, namely, polynomial, RBF and sigmoid.

5.4.1 Results using wavelet transform

Segmented lung region is first decomposed with 2-*D* wavelet transform. The effect of level of decomposition on system performance is evaluated by considering either only LL component obtained after each level of decomposition or by considering various possible combinations of LL components obtained after each level of decomposition. Total number of possible combinations, considering single LL component and combinations of two, three, four or five LL components comes out to be 31.

Table 5.2 depicts the accuracy of proposed method for predicting obstructive lung disease using 10-fold cross validation technique. This table shows that as the level of decomposition increases, accuracy of the proposed system increases. It has also been observed that further decomposition beyond level 5 does not lead to improvement in accuracy.

It has also been observed that performance of the system improves as the size of feature vector is increased by considering the combinations of LL coefficients. It is clear from the results that the combination of LL1, LL2, LL3, LL4 and LL5 along with RBF kernel function gives the maximum accuracy of 89.0% in predicting obstructive lung disease.

Table 5.3 illustrates the accuracy of proposed system with feature vector taken as a combination of LL1, LL2, LL3, LL4 and LL5 for subjects with different spirometry test ratio.

Figure 5.1 depicts that 95.3% of the subjects identified as having obstructive lung disease are from the category of smokers and 97.8% of the subjects identified as healthy are from the category of non-smokers.

Table 5.2: Overall accuracy with different feature vectors and kernel functions for predicting obstructive lung disease

kernel Function / Feature Vector	Polynomial (2nd degree) (in %)	Polynomial (3rd degree) (in %)	RBF (in %)	Sigmoid (in %)
LL1	77.0	76.0	80.0	80.0
LL2	77.0	77.0	81.0	81.0
LL3	78.0	77.0	82.0	81.0
LL4	79.0	79.0	82.0	82.0
LL5	80.0	82.0	83.0	83.0
LL1 + LL2	78.0	81.0	82.0	80.0
LL1 + LL3	78.0	81.0	82.0	80.0
LL1 + LL4	78.0	82.0	82.0	82.0
LL1 + LL5	80.0	83.0	82.0	82.0
LL2 + LL3	80.0	80.0	82.0	80.0
LL2 + LL4	78.0	80.0	82.0	81.0
LL2 + LL5	78.0	81.0	81.0	81.0
LL3 + LL4	78.0	81.0	82.0	82.0
LL3 + LL5	79.0	83.0	82.0	82.0
LL4 + LL5	80.0	84.0	83.0	82.0
LL1 + LL2 + LL3	83.0	84.0	84.0	83.0
LL1 + LL2 + LL4	83.0	85.0	86.0	84.0
LL1 + LL2 + LL5	82.0	85.0	86.0	84.0
LL1 + LL3 + LL4	82.0	85.0	87.0	84.0
LL1 + LL3 + LL5	83.0	85.0	87.0	85.0
LL1 + LL4 + LL5	83.0	85.0	87.0	85.0
LL2 + LL3 + LL4	82.0	83.0	86.0	83.0
LL2 + LL3 + LL5	82.0	83.0	86.0	84.0
LL2 + LL4 + LL5	82.0	84.0	87.0	84.0
LL3 + LL4 +LL5	82.0	84.0	86.0	84.0
LL1 + LL2 + LL3 + LL4	84.0	85.0	87.0	86.0
LL1 + LL2 + LL3 + LL5	84.0	86.0	88.0	86.0
LL1 + LL2 + LL4 + LL5	84.0	86.0	88.0	86.0
LL1 + LL3 + LL4 + LL5	85.0	86.0	88.0	87.0
LL2 + LL3 + LL4 + LL5	85.0	85.0	88.0	87.0
LL1 + LL2 + LL3 + LL4 + LL5	85.0	87.0	**89.0**	87.0

Table 5.3: Accuracy using wavelet features for subjects with different ρ

ρ	Number of Subjects	Accuracy (in %)
≤ 0.7	25	88.0
$0.7 < \rho < 0.8$	24	87.5
≥ 0.8	51	90.2

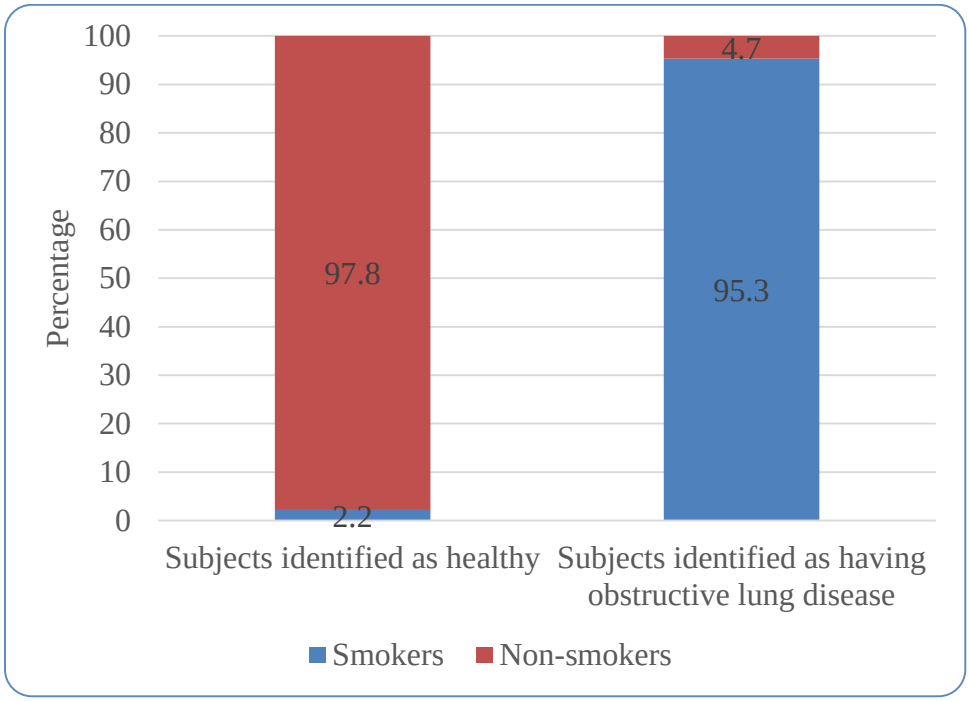


Figure 5.1: Distribution of smokers and non-smokers in subjects identified as healthy and the subjects identified as having obstructive lung disease using wavelet features

5.4.2 Results using Gabor filter

In the next experiment, features have been extracted by convolving the segmented lung region with 2-*D* Gabor filter. The maximum accuracy of the proposed method with these features is achieved when RBF kernel function is employed. This accuracy using 10-fold cross validation is 88.0%.

Table 5.4 illustrates the accuracy of the proposed method for subjects with different spirometry test ratio.

Table 5.4: Accuracy using Gabor filter for subjects with different ρ

ρ	Number of Subjects	Accuracy (in %)
≤ 0.7	25	88.0
$0.7 < \rho < 0.8$	24	83.3
≥ 0.8	51	90.2

Figure 5.2 depicts that 95.2% of the subjects identified as having obstructive lung disease are from the category of smokers and 95.7% of the subjects identified as healthy are from the category of non-smokers.

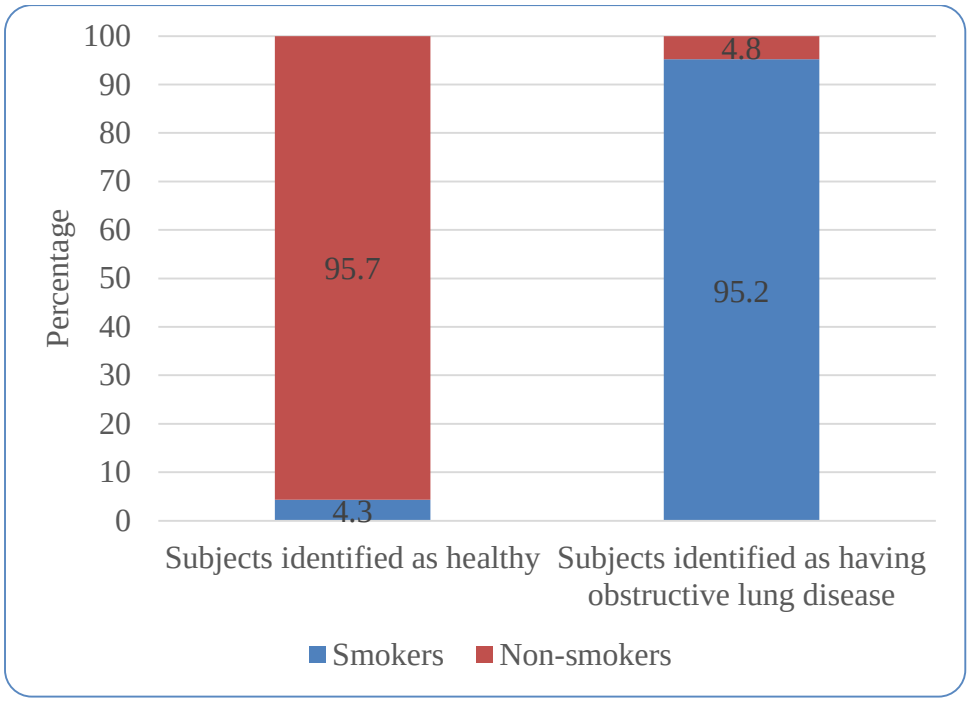


Figure 5.2: Distribution of smokers and non-smokers in subjects identified as healthy and the subjects identified as having obstructive lung disease using Gabor filter

Table 5.5 shows the comparison of proposed models to predict obstructive lung disease with existing irido-diagnosis models. The obtained accuracy with higher sample size demonstrates the effectiveness of proposed system.

Table 5.5: Comparison of accuracy of proposed obstructive lung disease prediction models with existing techniques

S. No.	Disease	Feature Extraction Technique	Classifier	No. of Samples	Accuracy (in %)
1.	Detecting broken tissues in pancreas (Wibawa and Purnomo, 2006)	Minimum filter	Visual inspection	34	94.0
2.	Alimentary canal (Ma and Li, 2007)	Gabor filter	SVM	53	84.9
3.	Nerve system (Ma and Li, 2007)	Gabor filter	SVM	44	86.4
4.	Hearing loss (Stearn and Swanepoel, 2007)	Digital image enhancement	Visual inspection	53	70.0
5.	Pancreas disorder (Lesmana *et al.*, 2011)	GLCM	Neighborhood based modified backpropagation using adaptive learning parameters (ANMBP)	50	83.3
6.	Pulmonary diseases (Sivasankar *et al.*, 2012)	Fuzzy C-means clustering	Gray level analysis	32	84.3
7.	Proposed model	Gabor filter	SVM	100	88.0
8.	Proposed model	Wavelet transform approach 2	SVM	100	89.0

Chapter Summary

In this chapter, an automated system to predict the obstructive lung disease using iris images has been proposed and implemented. Two different models to extract significant features have been proposed, one using wavelets and other using Gabor filter. SVM based classifier has been implemented for classifying subjects into healthy or one with obstructive lung disease. In case of proposed model based on wavelets, it has been found that the level of decomposition and choice of feature vector significantly affects the performance of the system.

An accuracy of 89.0% and 88.0% in case of wavelet based model and Gabor filter based model, respectively with a sample size of 100 has been obtained in this research work.

CHAPTER 6

CONCLUSION AND FUTURE SCOPE OF WORK

6.1 Conclusion

In this work, one of the prominent biometric techniques, *i.e.*, iris recognition has been investigated. Various existing algorithms involved in different stages of iris recognition system have been surveyed.

Implementation of iris recognition system with and without normalization and their comparative study has been carried out. Experiments have been conducted on standard iris databases: *IIT Delhi Iris Database Version 1.0* (*Database A*) and *CASIA-IRIS-V3* (*Database B*). Iris recognition system without normalization is simple, has less algorithmic complexity and requires less memory to store feature vectors of iris images. However, iris recognition systems with normalization have significantly improved EER of 1.5% and 2.7% as compared to 19.3% and 22.0% in case of systems without normalization for *Database A* and *Database B*, respectively. It has also been found that increase in number of virtual circles drawn on iris images improves system performance when we find the features for iris image without normalization. Results are indicative of the fact that normalization is an important step in iris recognition systems.

In this research work, a simple, effective and computationally efficient statistical feature extraction based iris recognition system has been proposed and implemented. Experiments have been performed to extract features in two different directions, namely, radial and angular directions. It is observed that system performance improves with increased number of statistical features. Results also show that increased radial and angular resolution, while normalization in place, improves the accuracy of iris recognition system. Comparison between test image template and templates stored in the database has been carried out using Hamming distance. The minimum error rate achieved in the experiments conducted in this work is 1.3% when 200-bit feature vector along radial direction is considered and is 3.4% when 1000-bit feature vector along

angular direction is considered. Results also show that statistical feature extraction based technique creates compact iris code and takes less time for feature extraction. The results obtained are encouraging and effective.

Iris recognition system is considered as a reliable and accurate biometric authentication system. Some work has been reported in literature to explore clinical applications of iris recognition systems. In this thesis, iris recognition system for clinical applications has been proposed and implemented. Iris recognition system has been combined with iridology to develop a model to predict the human being as healthy or suffering from a disease. Three applications, namely, predicting gender of an imposter, predicting diabetes and predicting obstructive lung disease have been considered.

SVM based gender prediction model has been implemented using two different feature extraction techniques, namely, wavelet transform and Otsu's threshold method. Maximum accuracy of 85.6% obtained in this work is encouraging and it shows an improvement over existing models. It has been noted that SVM as a classifier requires smaller training set and has small computational complexity.

It has been observed that system performance is affected by the sampling procedure, as emphasized by other authors. Human response greatly affects the ground truth while taking samples. One may not give the correct duration of the period they are having disease. They may also not give accurate information about the treatments they are taking.

Features have been extracted from specific location of the iris image using 2-*D* wavelet tree decomposition and also using Gabor filter. It has been found that the combination of wavelet coefficients as well as level of decomposition affects the performance of system significantly. SVM based binary classifier has been implemented to classify the subjects. An overall accuracy of 88.3% in predicting diabetes and 89.0% in predicting obstructive lung disease has been achieved in this work.

This is worth mentioning here that proposed models are of non-invasive and non-contact type. The systems developed using these models are compact and portable. One requires an iris scanner and a computing machine for implementing these systems. We expect that the system proposed in this work shall be used for the benefit of mankind.

The key contributions of this research work are:

i. Use of normalized iris images leads to better recognition results as concluded by the experiments conducted on the two standard databases.
ii. The features used in iris recognition systems should be easy to compute. In this work, six features have been proposed for such a system. The time required to compute the features proposed in this study is least when compared with other existing techniques and using the iris images from the two standard databases.
iii. An SVM based efficient gender prediction model has been proposed in this work. This model achieved a 10-fold recognition accuracy of 85.6% on a database of 400 iris images.
iv. An SVM based diabetes prediction system has been proposed. This system explores the concepts of iridology for predicting the diabetes in a subject. The experiments conducted for this model achieved a 4-fold recognition accuracy of 88.3% on a database of 80 iris images.
v. An SVM based obstructive lung disease prediction model has also been proposed in this work. This system explores the concepts of iridology to predict obstructive lung disease. The experiments conducted for this model achieved a 10-fold recognition accuracy of 89.0% on a database of 100 iris images.

6.2 Future Scope of Work

The work carried out in this thesis has a great potential for its extension. Following key points can be considered for extending this work.

- Performance of proposed statistical feature extraction based iris recognition system is effective. However, the accuracy can further be improved by considering more efficient features and classifiers.
- Segmentation of iris region from eye images can be improved by considering other segmentation techniques. Efficient techniques to remove noise due to occlusion and reflection can further improve system performance.
- Proposed methods predict the human being as healthy or suffering from a disease (diabetes/obstructive lungs disease). However, system capabilities can further be enhanced to predict the level/stage of the disease.

- SVM as a classifier has been used in the proposed models. One can explore the use of other classifiers on same dataset to achieve even better an accuracy.
- Age of subject and duration of smoking/diabetes can be considered as one of the features along with other extracted features in order to predict the level of disease with suitable modifications in training patterns.
- Here, dataset used for clinical applications is taken from Indian subjects. This work can be extended at international level.
- Iris dataset for clinical applications has been created using I-SCAN-2 dual iris scanner. In order to have an inclusive usage of the proposed models, the experiments need to be conducted on a low-price scanner.
- A web based application for predicting diabetes and predicting obstructive lung disease can be developed.
- The scope of proposed models can be extended for forensic applications to determine gender, age or ethnicity of unknown person/body.

LIST OF PUBLICATIONS

Publications in International Journals

Bansal A., Agarwal R. and Sharma R.K. (2015) 'Determining diabetes using iris recognition system' *Springer's International Journal of Diabetes in Developing Countries***.** Online available at http://link.springer.com/article/10.1007/s13410-015-0296-1, **DOI: 10.1007/s13410-015-0296-1.**

Bansal A., Agarwal R. and Sharma R.K. (2014) 'Predicting gender using iris images' *Research Journal of Recent Sciences***,** 3(4), 20-26.

Bansal A., Agarwal R. and Sharma R.K. (2014) 'Statistical feature extraction based iris recognition system' *Image Analysis and Stereology*, **Communicated.**

Bansal A., Agarwal R. and Sharma R.K. (2015) 'Pre-diagnostic tool to predict obstructive lung disease using iris recognition system' *Elsevier's Computerized Medical Imaging and Graphics*, **Communicated.**

Publications in Conferences

Bansal A., Agarwal R. and Sharma R.K. (2012) 'SVM based gender classification using iris images' in *Proc. of 4th IEEE International Conference on Computational Intelligence and Computer Networks, CICN 12*, 3-5 Nov., 2012, 425-429.

Bansal A., Agarwal R. and Sharma R.K. (2012) 'FAR and FRR based analysis of iris recognition system' in *Proc. of IEEE International Conference on Signal Processing, Computing and Control, ISPCC 12*, 15-17 March, 2012, 1-6.

Bansal A., Agarwal R. and Sharma R.K. (2010) 'Trends in iris recognition algorithms' in *Proc. of 4th IEEE Asia International Conference on Mathematical Modeling & Simulation, AMS 10*, 26-28 May, 2010, 337-340.

Bansal A., Agarwal R. and Sharma, R. K. (2014) 'An Iris recognition based health examination system' in *Proc. of 3rd National Conference on Advances in Metrology, AdMet-2014*, 19-21 Feb., 2014, pp. 67-68.

(Best Paper Award)

REFERENCES

Almaadeed N., Aggoun A. and Amira A. (2015) 'Speaker identification using multimodal neural networks and wavelet analysis', *IET Biometrics,* vol.4, no.1, pp.18-28

Angadi S.A. and Gour S. (2014) 'Euclidean distance based offline signature recognition system using global and local wavelet features', in *Proc. of 2014 Fifth International Conference on Signal and Image Processing (ICSIP),* vol. 1, 8-10 Jan. 2014, pp. 87-91

Bachoo A.K. and Tapamo J.R. (2005) 'Texture detection for segmentation of iris images', in *Proc. of ACM International Conference South African Institute of Computer Scientists and Information Technologists*, White River, South Africa, 20-22 Sep. 2005, pp. 236–243

Bodade R.M. and Talbar S.N. (2009) 'Shift invariant iris feature extraction using rotated complex wavelet and complex wavelet for iris recognition system', in *Proc. of Seventh International Conference on Advances in Pattern Recognition, ICAPR'09,* 4-6 Feb. 2009, pp. 449-452

Boles W.W. and Boashash B. (1998) 'A human identification technique using images of the iris and wavelet transform', *IEEE Transaction on Signal Processing*, vol. 46, no. 4, pp. 1185-1188

Burges C.J.C. (1998) 'A tutorial on support vector machines for pattern recognition', *Boston: Kluwer Academic Publishers*

Burke M. J. and Nasor M. (2004) 'Wavelet based analysis and characterization of the ECG Signal'', *Journal of Medical Engineering and Technology,* vol. 28, no. 2, pp. 47-55

Cao K., Eryun L., and Jain A.K. (2014) 'Segmentation and enhancement of latent fingerprints: A Coarse to Fine Ridge Structure Dictionary', *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 36, no. 9, pp. 1847-1859

Chaskar U.M. and Sutaone M.S. (2012) 'On a methodology for detecting diabetic presence from iris image analysis', in *Proc. of International Conference on Power, Signals, Controls and Computation (EPSCICON),* 3-6 Jan., pp. 1-6

Chen Y., Dass S., and Jain A. (2005) 'Localized iris image quality using 2-*D* wavelets', *Advances in Biometrics, Springer Berlin,* vol. 3832/2005, pp.373-388

Chou Chia-Te, Shih Sheng-Wen, Chen Wen-Shiung, Cheng V.W. and Chen Duan-Yu (2010) 'Non-Orthogonal vies iris recognition system', *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no.3, pp. 417-430

Chu W.S., Rong C. and Song Chen C. (2010) 'Identifying gender from unaligned facial images by set classification', in *Proc. of* 20th *IEEE International Conference on Pattern Recognition,* 23-26 Aug. 2010, pp.2636 -2739

Cortes C. and Vapnik V. (1995) 'Support vector networks', *Machine Learning,* 20:273–97

Costa R. M.D. and Gonzaga A. (2012) 'Dynamic features for iris recognition', *IEEE Transactions on Systems, Man, and Cybernetics part-B: Cybernetics,* vol. 42, No.4, pp. 1072-1082

Cristianini N. and Shawe T.D. (2000) 'An introduction to support vector machines and other kernel-based learning methods', *Cambridge: Cambridge University Press*

Cui J., Wang Y., Tan T., Ma L. and Sun Z. (2004) 'A fast and robust iris localization method based on texture segmentation', *SPIE Defense and Security Symposium*, vol. 5404, pp. 401-408

Daugman J. (1993) 'High confidence visual recognition of persons by a test of statistical independence', *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 15, pp 1148-1161

Daugman J. (2004) 'How iris recognition works?', *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 14, no.1, pp.21-30

Dubey S.K., Anna T., Shakher C. and Mehta D. S. (2007) 'Fingerprint detection using full-field swept-source optical coherence tomography', *Applied Physics Letters* 91, 181106

Field D. (1987) 'Relations between the statistics of natural images and the response properties of cortical cells', *Journal of the Optical Society of America*, vol.4, Issue-12, pp. 2379-2394

Gaidhane V.H, Hote Y.V and Singh V. (2012) 'An efficient similarity measure technique for medical image registration', *Sadhna,* vol. 37, 6, pp.709-721

Giot R., Hemery B. and Rosenberger C. (2010) 'Low cost and usable multimodal biometric system based on keystroke dynamics and 2-*D* face recognition', in *Proc. of 20th IEEE International Conference on Pattern Recognition,* 23-26 Aug. 2010, pp.1128-1131

Giovanni N. (2012) 'Relevance of Iridology' *European Journal of Integrative Medicine*, vol. 4, no. 1, pp. 37

Grabowski K., Sankowski W., Zubert M. and Napieralska M. (2006) 'Reliable iris localization method with application to iris recognition in near infrared light', in *Proc. of International Conference on Mixed Design of Integrated Circuits and System,MIXDES'06,* Poland, 22-24 June 2006, pp. 684-687

Han X., Ugail H. and Pahnar I. (2009) 'Gender classification based on 3-*D* face geometry features using SVM', in *Proc. of IEEE International Conference on Cyber World,* 7-11 Sept., pp. 114 -118

Haykin S. (2004) 'Neural networks—a comprehensive foundation', *Pearson Education, second edition*

He X. and Shi P. (2008) 'Extraction of complex wavelet features for iris recognition', in *Proc. of 19th International Conference on Pattern Recognition, ICPR'08,*8-11 Dec. 2008, pp. 1–4

He Y., Feng G., Hou Y. and Li L., Tzanakou E. M. (2011) 'Iris feature extraction method based on LBP and chunked encoding', in *Proc. of IEEE 7th International Conference on Natural Computation,* 26-28 July 2011, vol.3, pp. 1663-1667

Huang J., Wang Y., Cui J., and Tan T. (2004) 'Noise removal and implanting model for iris image', in *Proc. of IEEE International Conference on Image Processing*, 24-27 Oct., vol. 2, pp. 869–872

Huang J., Wang Y., Tan T., and Cui J. (2004) 'A new iris segmentation method for recognition', in *Proc. of 17th International Conference on Pattern Recognition,* 23-26 Aug. 2004, vol. 3, pp. 554-557

Hussein S.E., Hassan O.A. and Granat M.H. (2013) 'Assessment of the potential iridology for diagnosing kidney disease using wavelet analysis and neural networks', *Biomedical Signal Processing and Control*, vol. 8, issue 6, pp. 534-541

Jayasekara B., Jayasiri A. and Udawatta L. (2006) 'An Evolving Signature Recognition System', in *Proc. of First International Conference on Industrial and Information Systems,* 8 - 11 August 2006, pp.529-534

Jensen B. (1985) 'The science and practice of iridology', *California Bernard Jensen Co.*, 1

Kang B.J. and Park K.R. (2007) 'A robust eyelash detection based on iris focus assessment', *Pattern Recognition,* vol. 28, pp.1630–1639

Khan S.A., Nazir M., Akram S. and Riaz N. (2011) 'Gender classification using image processing techniques: Survey', in *Proc. of IEEE 14th International Conference on Multitopic (INMIC), 22*-24 Dec. 2011, pp. 25-30

Ko Jong Gook, Gil Yeon Hee and Yoo Jang Hee (2006) 'Iris recognition using cumulative sum based change analysis", in *Proc. of International symposium on intelligent signal processing and communication system,* 12-15 Dec. 2006, pp. 275-278

Kong W. and Zhang D. (2001) 'Accurate iris segmentation based on novel reflection and eyelash detection model', in *Proc. of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing*, Hong Kong, 2-4 May. 2001, pp. 263-266

Kumar A. and Anand S. (2006) 'EEG signal processing for monitoring depth of anesthesia', *IETE Technical Review*, vol. 23, no. 3, pp 179-186

Kumar A., Garg S. and Hanmandlu M. (2014) 'Biometric authentication using finger nail plates', *Expert Systems with Applications,* vol. 41, no. 2: 373-386

Kumar A. and Passi A. (2008) 'Comparison and combination of iris matchers for reliable personal identification', in *Proc. of Computer Vision and Pattern Recognition 2008*, Anchorage, Alaska, 23-28 June, pp. 21-27

Kyaw K.S.S. (2009) 'Iris recognition system using statistical features for biometric identification', in *Proc. of International Conference on Electronic Technology, IEEE Computer Society,* 20-22 Feb., pp.554-556

Lagree S. and Bowyer K.W. (2011) 'Predicting ethnicity and gender from iris texture', in *Proc. of IEEE International Conference on Technologies for Homeland Securities (HST),* 15-17 Nov. 2011, pp. 440-445

Lai C, Chiu C. (2010) 'Health examination based on iris images', in *Proc. of 9th International Conference on Machine Learning and Cybernetics, Qingdo,* 11-14 July 2010, vol. 5, pp. 2616-2621

Liau H.F. and Isa D. (2011) 'Feature selection for support vector machine-based face-iris multimodal biometric system', *Expert Systems with Applications*, vol. 38, pp. 11105–11111

Lesmana I.P.D., Purnama I.K.E. and Purnomo M.H. (2011) 'Abnormal condition detection of pancreatic Beta-Cells as the cause of diabetes mellitus based on iris image', in *Proc. of International Conference on Instrumentation, Communication, Information Technology and Biomedical Engineering*, 8-9 Nov. 2011, pp. 150-155

Lundback B. *et al.* (2003) 'Not 15 but 50% of smokers develop COPD?-Report from obstructive lung disease Northern Sweeden studies", *Respiratory Medicine,* vol. 97, pp. 115-122

Ma L. and Li N. (2007) 'Texture feature extraction and classification for iris diagnosis', *International conference on medical biometrics, Lecture notes in computer science, Springer-Verlag,* vol. 4901, pp. 168-175

Ma L., Wang Y. and Tan T. (2002) 'Iris recognition using circular symmetric filters', *National Laboratory of Pattern Recognition*, Institute of Automation, Chinese Academy of Sciences, pp. 414-417

Mamta and Hanmandlu M. (2015) 'Multimodal biometric system built on the new entropy function for feature extraction and the Refined Scores as a classifier', *Expert Systems with Applications,* vo. 42, no. 7, pp. 3702-3723

Masek, L. (2003) 'Recognition of human iris patterns for biometric identification', *Technical report,* School of Computer Science and Software Engineering, University of Western Australia

Min J. Y., Min K. B., Cho S. and Paek D. (2008) 'Combined effect of cigarette smoking and sulphur dioxide on lung functions in Koreans' *Journal of Toxicology and Environmental Health, Part-A,* vol. 71, pp. 301-303

Monro D. M., Rakshit S., and Zhang D. (2007) 'DCT based iris recognition', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no.4, pp. 586-595

Nazir M., Ishtaiq M., Batool A., Jaffar A. and Mirza A.M. (2010) 'Feature selection for efficient gender classification', In *Proc.* of *WSEAS International conference*, 13 June. 2010, pp. 70-75

Nova R., Escalante-Ramirez B., Cristobal G. and Estepar P.S.J. (2014) 'Extended Gabor approach applied to classification of emphysematous patterns in computed tomography', *Medical Biological Engineering and Computing*, vol. 52, pp. 393-403

Othman Z, Prabuwono A S. (2010) 'Preliminary study on iris recognition system: tissues of body organs in iridology', in *Proc. of IEEE, EMBS, Conference on Biomedical Engineering & Sciences (IECBES2010), Kuala Lumpur, Malaysia*, Nov. 30 - Dec. 2, pp. 115-119

Otsu N. (1979) 'A threshold selection method from gray-level histograms', *IEEE Transaction on System, Man and Cybernetics*, vol. 9, no. 1, pp 62–66

Peng-Fei Zhang, De-Sheng Li and Qi Wang (2004) 'A novel iris recognition method based on feature fusion', in *Proc. of 2004 International Conference on Machine Learning and Cybernetics,* vol. 6, pp. 3661-3665

Pierce R. (2005) 'Spirometry: an essential clinical measurement' *Australian family physician*, vol. 37, no. 7, pp. 535-539

Qiu X.C., Sun Z.A., and Tan T.N. (2006) 'Global texture analysis of iris images for ethnic classification', *Springer Lecture note in Computer Science, Advances in Biometrics*, vol. 3832, pp.411-418

Qiu X.C., Sun Z.A., and Tan T.N. (2007) 'Learning appearance primitives of iris images for ethnic classification', in *Proc. of IEEE International Conference on Image Processing (ICIP)*, Sept. 16 - Oct. 19, pp. II-405–II-408

Rahul P., Katariya A., Tomar G.S. and Prajapat K.K. (2011) 'DWT based multi-level digital watermarking technique for images', *Journal of Signal and Image Processing*, vol.2, pp. 36-41

Rahulkar A. D. and Holambe R. S. (2012) 'Half-Iris feature extraction and recognition using a new class of bi-orthogonal triplet half-band filter bank and flexible k-out-of-n: A post classifier', *IEEE Transaction on Information Forensics and Security*, vol. 7, no. 1, pp. 230-240

Rahulkar A. D., Waghmare L. M. and Holambe R. S. (2013) 'A new approach to the design of hybrid finer directional wavelet filter bank for iris feature extraction and

classification using k-out-of-n: A post classifier', *Pattern Analysis and Applications*, 2013 doi: 10.1007/s 10044-013-0334-x

Rai P. And Khanna P. (2010) 'Gender classification using Randon and wavelet transform', in *Proc. of IEEE International Conference on Industrial and Information Systems*, 29 July - 1 Aug., pp. 448 – 451

Rajan R., Indu M.G. (2014) 'A novel finger vein feature extraction technique for authentication', in *Proc. of 2014 Annual International Conference on Emerging Research Areas: Magnetics, Machines and Drives (AICERA/iCMMD)*, 24-26 July, pp.1-5

Ramlee R.A. and Ranjit S. (2009) 'Using iris recognition algorithms, detecting cholesterol presence', in *Proc. of International conference on Information Management and Engineering*, 3-5 April, pp. 714-717

Ritter N. (1999) 'Location of the pupil-iris border in slit-lamp images of the cornea', in *Proc. of the International Conference on Image Analysis and Processing*, 27-29 Sep., pp. 740-745

Roy K. and Bhattacharya P. (2007) 'Collarette area localization and asymmetrical support vector machines for efficient iris recognition', in *Proc. of 14th International Conference on Image Analysis and Processing,* 10-14 Sept. 2007, pp.3-8

Roy K. and Bhattacharya P. (2008) 'Optimal features subset selection and classification for iris recognition, *EURASIP Journal on Image and Video Processing*, doi:10.1155/2008/743103

Rydgren E., Thomas E.A., Amiel F., Rossant F. and Amara A. (2004) 'Iris features extraction using wavelet packets", *ICIP'04 International Conference on image processing*, 24-27 Oct. 2004, vol. 2, pp: 861-864,

Senoussaoui M., Kenny P., Stafylakis T., and Dumouchel P. (2014) 'A Study of the cosine distance-based mean shift for telephone speech diarization', *IEEE Transaction on Audio, Speech and Language Processing*, vol.22, no. 1, pp. 217-227, 2014

Sharma P., Khan Y.U., Farooq O., Tripathi M. and Adeli H. (2014) 'A wavelet-statistical features approach for non-convulsive seizure detection', *Clinical EEG and Neuroscience,* pii: 1550059414535465

Sivasankar K., Sujaritha M., Pasupathi P. and Muthukumar, S. (2012) 'FCM based iris image analysis for tissue imbalance stage identification', in *Proc. of International Conference on Emerging Trends in Science, Engineering and Technology (INCOSET)*, 13-14 Dec., pp. 210-215

Sizov A., Khoury E., Kinnunen T., Wu Z. and Marcel S. (2015) 'Joint speaker verification and anti-spoofing in the i -vector space', *IEEE Transactions on Information Forensics and Security,* vol.10, no.4, pp.821-832

Smith E., Stein P., Furst J. and Raicu D.S. (2013) 'Weak segmentations and ensemble learning to predict semantic ratings of lung nodules', in *Proc. of 12th International Conference on Machine Learning and Applications (ICMLA),* 4-7 Dec., vol. 2, pp. 519-524

Stearn N. and Swanepoel D.W. (2007) 'Identifying hearing loss by means of iridology', *African Journal of Traditional, Complimentary Alternative Medicine*, vol. 4, pp. 205-210

Sundaram R.M. and Dhara B.C. (2011) 'Neural network based iris recognition system using Haralick features', in *Proc. of IEEE 3rd International Conference on Electronics Computer Technology,* 8-10 April, vol. 3, pp. 19-23

Sung H., Lim J., Park J. and lee Y. (2004) 'Iris recognition using collarette boundary localization', in *Proc. of 17th International Conference on Pattern Recognition*, 23-26 Aug. 2004, vol.4, pp. 857-860

Teo C. and Ewe H.T. (2005) 'An efficient one-dimensional fractal analysis for iris recognition', in *Proc. of the 13th WSCG International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision,* 31 Jan.- 4 Feb., pp. 157-160

Thomas V., Chawla N., Bowyer K.W. and Flynn P.J. (2007) 'Learning to predict gender from iris images', in *Proc*. of *IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS),* 27-29 Sept., pp. 1-5

Tisse C., Martin L., Torres L. and Robert M. (2002) 'Person identification technique using human iris recognition', in *Proc. of International Conference on Vision Interface, Canada*, 2002, pp. 294-299

Vestbo J. (2013) 'Definition and Overview', *Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease. Global Initiative for Chronic Obstructive Lung Disease,* pp. 1–7

Wang J., Valle M.D., Goryawala M., Franquiz J.M. and Mcgoron A.J. (2010) 'Computer-assisted quantification of lung tumors in respiratory gated PET/CT images: phantom study', *Medical Biological Engineering and Computing*, vol. 48, pp. 49-58

Wang J. and Xie M. (2006) 'Iris feature extraction based on wavelet packet analysis', in *Proc*. of *IEEE International Conference on Communications, Circuits and Systems, June 2006,* vol.1, pp.31-34

Wang K., An N., Li B.N., Zhang Y. and Li L. (2015) 'Speech emotion recognition using Fourier parameters," *IEEE Transactions on Affective Computing,* vol.6, no.1, pp.69-75

Wang Z., Mat Y. and Guangzhu X. (2006) 'Novel method of iris feature extraction based on the PCM', in *Proc. of the 2006 IEEE International Conference on Information Acquisition, Weihai, Shandong, China*, 20-23 Aug. 2006, pp.814-818

Wang Z., Zhang H., Bi G. and Li X. (2014) 'Automatic multi-speaker speech recognition system based on time-frequency blind source separation under ubiquitous environment', in *Proc. of IEEE 9th Conference on Industrial Electronics and Applications (ICIEA), 2014*, 9-11 June 2014,pp. 101-106

Wen-Shiung Chen, Kun-HueiChih, Sheng-Wen Shih, Chih-Ming Hsieh (2005) 'Personal identification technique based on human iris recognition with wavelet transform', in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '05*, 18-23 March, vol. 2, pp. ii/949-ii/952

Wibawa A.D. and Purnomo M.H. (2006) 'Early detection on the condition of pancreas organ as the cause of diabetes mellitus by real time iris image processing', in *Proc. of IEEE Asia Pacific Conference on Circuits and Systems, APCCS-2006*, 4-7 Dec., pp. 1008-1010

Wildes R., Asmuth J., Green G., Hsu S., Kolczynski R., Matey J. and McBride S. (1994) 'A system for automated iris recognition', in *Proc. of IEEE Workshop on Applications of Computer Vision, Sarasota, FL*, 5-7 Dec., pp. 121-128

Wildes, R. P. (1997) 'Iris recognition: An emerging biometric technology', *Proceedings of the IEEE*. vol. 85, no. 9, pp.1348-1363

Xiang D., Yan H., Chen X. and Cheng Y. (2010) 'Offline Arabic handwriting recognition system based on HMM," in *Proc. of 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT)*, 9-11 July 2010, vol.1, pp. 526-529

Xiaofu He and PengfeiShi (2008) 'Extraction of complex wavelet features for iris recognition', in *Proc. of 19th International Conference on Pattern Recognition, ICPR*, 8-11 Dec., pp.1-4

Xu G., Zhang Z., and Ma Y. (2006) 'Improving the performance of iris recognition system using eyelids and eyelashes detection and iris image enhancement," in *Proc. of IEEE International Conference on Cognitive Informatics*, 17-19 July, vol. 2, pp. 871–876

Yang H., Xu Y., Huang H., Zhou R. and Yan Y. (2014) 'Voice biometrics using linear Gaussian model', *IET Biometrics,* vol. 3, no. 1, pp. 9-15

Young-Woo L., Young-Bae P. and Young-Jae P. (2014) 'Experimental study of reliable iris parameters and their relationships with temperament, character, and heart rate variability', *European Journal of Integrative Medicine*, vol. 6, issue 5, pp. 524-531

Yuh-Chin and Huang T. (2012) 'A clinical guide to occupational and environmental lung diseases*', Humana Press,* ISBN 978-1-62703-149-3. p. 266.

Zahedi M. and Zahra M. (2013) 'A fully automatic and Haar like feature extraction based method for lip contour detection", *Research Journal of Recent Sciences*, vol. 2, no. 1, pp. 17-20

Zahra M. (2012) ‘Image duplication forgery detection using two robust features’, *Research Journal of Recent Sciences*, vol. 1, no. 12, pp. 1-6

Zhang D., Monro D.M., and Rakshit S. (2006) ‘Eyelash removal method for human iris recognition’ in *Proc. of IEEE International Conference on Image Processing*, 8-11 Oct. 2006, pp. 285–288

Zhu Yong, Tan Tieniu and Wang Yunhong (2000) ‘Biometric personal identification based on iris patterns’, in *Proc. of the IEEE International Conference on Pattern Recognition*, 3-7 Sept., pp. 2801-2804

‘Iridology charts’, http://www.bernardjensen.com/Charts_c_17.html. [Accessed 5 June, 2011]

‘IIT Delhi Iris Database Version 1.0’, http://web.iitd.ac.in/~biometrics/Database_Iris.htm. [Accessed: June, 2010]

‘CASIA Iris Image Database’, http://www.sinobiometrics.com/Databases.htm. [Accessed: June, 2010]

‘Diabetes mellitus’, http://en.wikipedia.org/wiki/Diabetes_mellitus. [Accessed: 15 Sep. 2011]

A.D.A.M. Medical encyclopedia ‘Diabetes’, US National library of Medicine, http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0002194/. [Accessed: 15 Sep. 2011]