Recognition using SIFT and its Variants on Improved Segmented Iris

Apurva Pathak

Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India

Recognition using SIFT and its Variants on Improved Segmented Iris

Thesis submitted in

May 2013

to the department of

Computer Science and Engineering

of

National Institute of Technology Rourkela

in partial fulfillment of the requirements

for the degree of

Bachelor of Technology

in

Computer Science and Engineering

by

Apurva Pathak

[Roll: 109CS0171]

under the guidance of

Prof. Banshidhar Majhi

Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India

Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India. www.nitrkl.ac.in

Dr. Banshidhar Majhi
Professor

May 13, 2013

Certificate

This is to certify that the work in the project entitled *Recognition using SIFT and its Variants on Improved Segmented Iris* by *Apurva Pathak* is a record of his work carried out under my supervision and guidance in partial fulfillment of the requirements for the award of the degree of *Bachelor of Technology* in *Computer Science and Engineering*.

Banshidhar Majhi

Acknowledgment

"Only those who will risk going too far can possibly find out how far one can go"

- T. S. Elliot

All the efforts I have put forward in carrying out this project would have been incomplete, if not for the kind support of many individuals as well as this institute. I would like to express my deep sense of gratitude to all of them.

Foremost, I would like thank my supervisor of this project, Prof. Banshidhar Majhi, Department of Computer Science and Engineering, National Institute of Technology, Rourkela, for his incalculable contribution in the project. He stimulated me to work on the topic and provided valuable information which helped in completing the project through various stages. I would also like to acknowledge his exemplary guidance, monitoring and constant encouragement throughout the course of this thesis.

I am obliged to all the professors of the Department of Computer Science and Engineering, NIT Rourkela for instilling in me the basic knowledge about the field that greatly benefitted me while carrying out the project and achieving the goal.

Also, I am highly indebted to Mrs. Hunny Mehrotra, Department of Computer Science and Engineering, National Institute of Technology, Rourkela who invested her full effort in helping me in finalizing this project within the limited time frame.

Lastly, I am grateful to my friends, for their relentless support in augmenting the value of work; my family, for being considerate and appreciative throughout; and Almighty, for everything.

Apurva Pathak

Abstract

Iris is one of the most reliable biometric traits due to its stability and randomness. Iris is transformed to polar coordinates by the conventional recognition systems. They perform well for the cooperative databases, but the performance deteriorates for the non-cooperative irises. In addition to this, aliasing effect is introduced as a result of transforming iris to polar domain. In this thesis, these issues are addressed by considering annular iris free from noise due to eyelids. This thesis presents several SIFT based methods for extracting distinctive invariant features from iris that can be used to perform reliable matching between different views of an object or scene. After localization of the iris, Scale Invariant Feature Transform (SIFT) is used to extract the local features. The SIFT descriptor is a widely used method for matching image features. But SIFT is found out to be computationally very complex. So we use another keypoint descriptor, Speeded up Robust Features (SURF), which is found to be computationally more efficient and produces better results than the SIFT. Both SIFT and SURF has the problem of false pairing. This has been overcome by using Fourier transform with SIFT (called F-SIFT) to obtain the keypoint descriptor and Phase-Only Correlation for feature matching. F-SIFT was found to have better accuracy than both SIFT and SURF as the problem of false pairing is significantly reduced. We also propose a new method called S-SIFT where we used S Transform with SIFT to obtain the keypoint descriptor for the image and Phase-Only Correlation for the feature matching. In the thesis we provide a comparative analysis of these four methods (SIFT, SURF, F-SIFT, S-SIFT) for feature extraction in iris.

***Keywords*:** SIFT, SURF, F-SIFT, S-SIFT, Iris, Integro-differential operator, Iris Segmentation

Contents

Certificate ii

Acknowledgement iii

Abstract iv

Contents v

List of Figures vii

List of Tables x

1 Introduction 1

1.1 Iris Biometrics . . . 3

1.2 Motivation . . . 4

1.3 Contributions . . . 5

1.4 Various Performance Measures . . . 5

1.4.1 False Acceptance Rate (FAR) . . . 6

1.4.2 False Rejection Rate (FRR) . . . 6

1.4.3 Genuine Acceptance Rate (GAR) . . . 6

1.4.4 Receiver Operating Characteristic (ROC) Curve . . . 6

1.5 Iris Databases used in the Research . . . 7

1.5.1 CASIA version 3 . . . 7

1.5.2 UBIRIS version 1 . . . 7

1.6 Thesis Organization . . . 7

2 Literature Review **9**
2.1 Iris Segmentation 10
2.2 Iris Recognition 12

3 Iris Segmentation **14**
3.1 Removal of Specular Highlights 14
3.2 Pupil Detection 16
3.3 Iris Detection 16
3.4 Annular Iris 17

4 Keypoint Descriptors for Iris **20**
4.1 Scale Invariant Feature Transform (SIFT) 21
4.1.1 Keypoint Detection 21
4.1.2 Keypoint Descriptor 26
4.1.3 Keypoint Pairing 26
4.2 Speeded Up Robust Features (SURF) 28
4.2.1 Detection of Keypoints 28
4.2.2 Keypoint Pairing 34
4.3 F-SIFT 34
4.3.1 Keypoint Detection using SIFT 35
4.3.2 Keypoint Descriptor using Fourier Transform 35
4.3.3 Keypoint Pairing 36
4.4 S-SIFT 37
4.4.1 Keypoint Detection using SIFT 38
4.4.2 Keypoint Descriptor using S Transform 38
4.4.3 Keypoint Pairing 39

5 Experimental Results and Analysis **40**
5.1 Iris Segmentation 40
5.2 Iris Recognition 42

6 Conclusions **45**

Bibliography **46**

List of Figures

1.1 Various forms of authentication. Traditional methods of authentication using token based and knowledge based approaches (left). Use of biometrics to claim identity (right) 2

1.2 Outline of a general biometric system 2

1.3 Anatomy of human eye 4

3.1 Adaptive image thresholding using grid based approach: (a) Iris image with blocks to compute threshold (b) Binary image obtained using adaptive threshold (taken from [1]) 15

3.2 Pupil Detection: (a) Spectrum image (b) Edge detected image with pupil center (c) Pupil localised image 17

3.3 Iris Detection: (a) Contrast enhanced image (b) Concentric circles of different radii (c) Iris localised image 19

3.4 Preprocessing of iris image: (a) Input iris image, (b) Geometrical representation of sectors on iris circles, (c) Noise independent annular iris image after preprocessing (taken from [1]) 19

4.1 Scale space extrema for different octaves. Adjacent Gaussian images are subtracted to produce DOG images on right (taken from [2]) . . . 23

4.2 Detection of scale space extrema 23

4.3 Maxima and minima of DOG images are obtained by comparing a pixel to 26 neighbors in $3 \times 3 \times 3$ regions (taken from [2]) 24

4.4 Interpolation of datapoints to estimate location of extremum (taken from [1]) 25

4.5 Keypoint detection on annular iris image using SIFT (a) Detected keypoints after removing noise and edge responses, (b) Scale and direction of orientation is indicated by arrows 26
4.6 Window is taken relative to direction of dominant orientation. This window is weighted by a Gaussian and histogram is obtained for 4×4 regions (taken from [1]) 27
4.7 Integral images are used to calculate the sum of intensities inside a rectangular region of any size. 29
4.8 Left to right: discrete Gaussian second order derivative in y and xy direction. Approximation for the second order Gaussian partial derivative in $y - (D_{yy})$ and xy-direction (D_{xy}) (taken from [3]). 30
4.9 Use of integral images for upscaling filter masks (taken from [4]) . . . 31
4.10 Filters D_{yy} (top) and D_{xy} (bottom) for two successive filter sizes (9×9 and 15×15) [3]. 32
4.11 Detected interest points on annular iris image 32
4.12 Orientation assignment by taking a sliding window of size $\frac{\pi}{3}$ indicated by shaded region [3] 33
4.13 An oriented window with 4×4 sub-regions is taken in direction of orientation. For each sub-region wavelet responses are obtained [3]. . 33
4.14 Descriptor entries of a sub-region represent the nature of the underlying intensity pattern [3]. 34
4.15 Sample keypoints falsely paired using SIFT due to texture similarity (taken from [5]) 35
4.16 Phase-Only Correlation between similar and dissimilar regions of iris image (taken from [5]) 37
4.17 Block diagram of F-SIFT approach for matching (taken from [5]) . . 38

5.1 Localisation performance using [6] approach and proposed approach on image no S1001L01 of CASIA database 41
5.2 Failure to generate noise independent annular iris due to greater degree of occlusion by upper eyelid 42
5.3 ROC curve for annular iris image 43

5.4 Genuine and Imposter score distribution curve 44

List of Tables

2.1 Performance of some selected localisation approaches (taken from [7]) 12

5.1 Mis-localisation percentage of Sector Based approach of [6] and proposed approach 41

Chapter 1

Introduction

Today personal identification is required in a large range of applications like banking, driver licenses, voter id card, etc. Traditionally several types of authentication methods exists like (i) token based systems: where the protected resources are prevented from the imposters using smart cards, ID cards, etc, (ii) knowledge based systems: where the information like user id and password are used by a person to claim his identity. Some of the systems use the combinations of both the approaches. However, there are many disadvantages of using the traditional methods for authentication. The drawbacks of token based approaches is that the possession could be stolen, lost or misplaced. The problem with the knowledge based systems is to remember the PINs/passwords and that the intruders can guess the easily recallable passwords. Thus, the security requirements could not be satisfied even with the combination of token and knowledge based systems [8]. Biometrics provides a more reliable, capable and trustworthy solution to the problems of traditional authentication approaches.

The term *Biometrics* is related to the field of development of mathematical and statistical methods that is applicable to the data analysis problems existing in the biological sciences. It is the science of establishing an individual's identity based on his physiological and behavioural characteristics. It offers reliable and natural solution to identity management by utilising semi-automated or fully automated schemes to recognise an individual [9]. The primary motivation of using biometrics over knowledge based and token based approaches is that, it cannot be forgotten, misplaced or stolen. Spoofing biometric systems are very difficult because the person is required

Figure 1.1: Various forms of authentication. Traditional methods of authentication using token based and knowledge based approaches (left). Use of biometrics to claim identity (right)

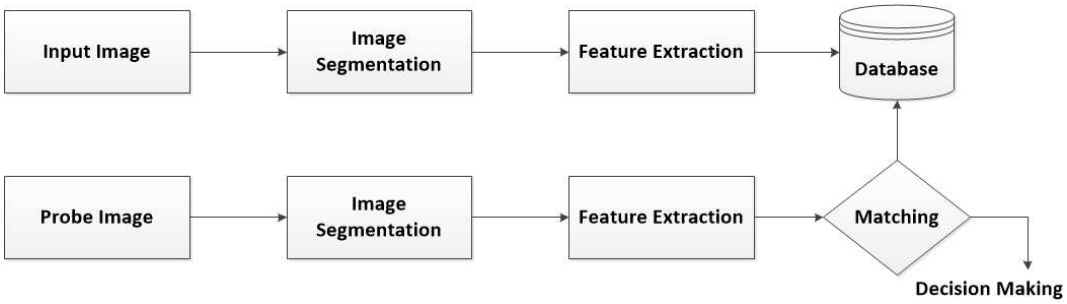


Figure 1.2: Outline of a general biometric system

to be present physically for the authentication. Different forms of authentication are shown in Figure 1.1. The main modules involved in a generic biometric system is given in Figure 1.2. In a general biometric system first the input image is acquired from the user, then it is segmented to obtain the region of interest. Finally features are extracted from the segmented image and stored in the database as a biometric template. To identify a probe image the similar steps are performed to obtain the features, then the biometric template of the probe image is compared one-to-many with the existing templates in the database. If a match is found then the necessary decision/action is taken. The steps are clearly shown in Figure 1.2.

There exists many traits like iris, gait, face, fingerprint, ear etc. Two basic categories can be identified by looking at the nature of the underlying modalities. They are: (i) Physiological (or passive) biometrics like iris, fingerprint, face, hand geometry,

etc, are based on data and measurement derived from direct measurement of a part of the subject's body; and (ii) Behavioral (or active) biometrics like voice recognition, keystroke dynamics, and signature, are based on an action taken by a person and it indirectly measures the characteristics of the human body. A good biometric trait should have highly unique and stable feature that can be easily captured.

1.1 Iris Biometrics

Iris plays a significant role among various available biometric traits to provide a promising solution to authenticate an individual using unique texture patterns [10]. Iris is proved to be the most efficient biometric technique, taking invasiveness and reliability into consideration. From the reliability point of view, each individual has unique spatial patterns. From the invasiveness point of view, iris is a protected internal organ whose random texture is stable throughout the life. It can serve as a kind of password that one always carries along and need not remember.

Iris is the most significant and promising feature in the eye image (shown in Figure 1.3). The iris has the form of circular ring that contains several interlacing minute characteristics such as coronas, freckles, stripes, crypts, furrows and so on. These minute patterns in the iris are unique to each individual and are not invasive to their users. The central dark circle inside the iris is known as pupil. The muscles in the iris cause the pupil to dilate in dim light and constrict in bright light. The amount of light entering the eye is controlled by this pupillary motion. The circumference of iris and pupil is known as iris and pupil boundary respectively. The white, tough and leather-like tissue surrounding the iris is known a sclera. Apart from these features, the eyeball is covered by lower and upper eyelids. The lower eyelid has a smaller degree of motion which is caused by deformation due to eyeball. The upper eyelid on the other hand is a stretchable membrane that has a great freedom of motion, ranging from wide open to close. It can form a cover over the eye [11]. The hairs that grow at the edge of the eyelid and protects the eye from dust are called as eyelashes.

The unique iris patterns can be extracted from the acquired image of an eye by the application of image processing techniques, and generate a biometric template, that can be stored in the database. This biometric template contains the unique

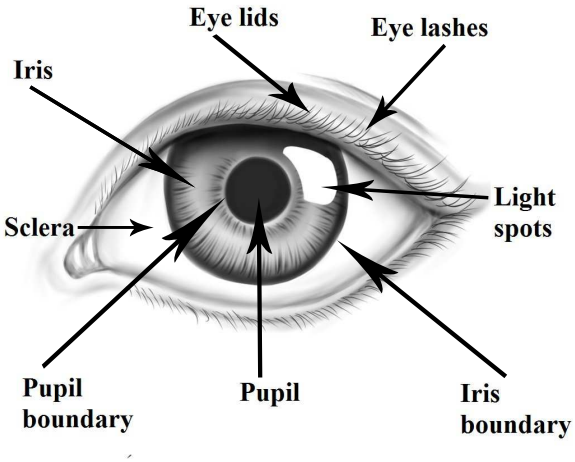


Figure 1.3: Anatomy of human eye

texture information stored in the iris in the form of a mathematical representation. To identify a subject by an iris recognition system, first their eye is photographed, and a template is created for their iris region after that. Then this template is compared with the other templates stored in the database until either a matching template is found and the subject is identified, or no matching template is found and the subject remains unidentified.

1.2 Motivation

The acquired iris image is localized for the detection of annular ring underlying the inner pupil and the outer iris boundary. The traditional method converts this annular ring into doubly dimensionless polar coordinates [10] and hence suffers from the problem of aliasing. The sector based approach in [6] is used to obtain a noise independent annular iris and does not perform well for the non-cooperative databases. A more robust method for the segmentation is required that works better than the existing approaches for the iris segmentation. After localization the main objective is to obtain features from the image that are invariant to scale, position and orientation. Global feature extraction techniques fail due to transformation of features between two image samples [7]. To obtain transformation invariance between the features, local feature extraction algorithms like SIFT [2], SURF [3] and F-SIFT [5] are used.

Each of these algorithms has their own advantages and disadvantages. This thesis gives a comparative analysis of these Scale Invariant Feature Transform (SIFT) based techniques for the feature extraction. It also proposes a S Transform based technique (S-SIFT) for the feature extraction.

1.3 Contributions

The following contributions have been made in this thesis:

- A new and improved algorithm of iris segmentation has been proposed.
- Algorithms for feature extraction viz. SIFT, SURF and F-SIFT have been implemented on two databases.
- A new algorithm for feature extraction has been proposed and implemented on CASIA V3 database.
- A comparative analysis of all the four algorithms for feature extraction has been done.

1.4 Various Performance Measures

The authenticity of a password is checked by finding a perfect match with any one of the alphanumeric strings stored in the database. However, in biometrics the same templates are compared in very rare occasions. There is difference between the two templates due to the change in characteristics with respect to aging, scanning conditions, change in acquisition conditions etc. Thus, the feature sets of the same individual does not look the same. When two different biometric templates of the same individual are different then it is known as **intra-class** variations. However, variations that occurs between templates of two different individuals are known as **inter-class** variations [12].

When the intra-class variations are found by comparing two biometric templates then such scores are known as genuine scores/similarity scores. However, when the inter-class similarity is found by comparing two biometric traits, then the scores are

known as imposter scores. The imposter scores that exceed a predefined threshold (τ), leads to false acceptance. The genuine score that lies below the threshold τ leads to false rejection. The commonly used measures to evaluate the performance of biometrics system are:

1.4.1 False Acceptance Rate (FAR)

FAR is the frequency of fraudulent access to imposters claiming identity [13]. This statistic is used in the verification mode to measure the biometric performance. A false accept occurs when the subject is incorrectly matched to another subject's existing biometric template.

1.4.2 False Rejection Rate (FRR)

FRR is the frequency of rejections of the people who should be correctly verified. This statistic is used in the verification mode to measure the biometric performance. A false reject occurs when the subject is not matched correctly to his existing biometric template in the template.

1.4.3 Genuine Acceptance Rate (GAR)

GAR is the fraction of genuine/similarity scores exceeding the threshold τ. It is defined as

$$GAR = 1 - FRR \tag{1.1}$$

1.4.4 Receiver Operating Characteristic (ROC) Curve

ROC curve is a wide-ranging way to analyze the performance of a biometric system. It shows the dependence of FAR with GAR for a change in the value of threshold τ. The curve is plotted using logarithmic, semi-logarithmic or linear scale. In some cases, ROC is also represented by plotting FAR against FRR for a changing threshold value.

1.5 Iris Databases used in the Research

The databases used in all the experiments in this thesis that are relevant to the research are UBIRIS version 1 [14] and CASIA version 3 [15].

1.5.1 CASIA version 3

CASIA version 3 (CASIAV3) is acquired under ideal comditions in an indoor environment with less noise factors. Majority of the images have been acquired in two sessions with an interval of minimum one month. The database comprises 249 individuals with a total of 2655 images from left and right eyes. CASIAV3 is a superset of CASIAV1 [16].

1.5.2 UBIRIS version 1

UBIRIS version 1 (UBIRIS.v1) database contains 1877 images collected from 241 persons in two different sessions. It is a non-cooperative database as the images for this database are acquired under noisy conditions with less constraints on the user.

1.6 Thesis Organization

The thesis consist of five chapters following this chapter:

Chapter 2: Literature Review

This chapter outlines the existing work on iris segmentation and feature extraction, their performance and limitations.

Chapter 3: Iris Segmentation

This chapter discusses a more robust approach for preprocessing the iris based on image morphology [6].

Chapter 4: Keypoint Descriptors for Iris

This chapter discusses different algorithms (SIFT [2], SURF [3] and F-SIFT [5]) for obtaining keypoint descriptors for iris. It also proposes a S Transform [17] based algorithm S-SIFT for obtaining keypoint descriptor from the iris.

Chapter 5: Experimental Results and Analysis

This chapter shows all the results of the performance measures of the proposed system and the comparative analysis with the existing system.

Chapter 6: Conclusions

This chapter presents analytical remarks to overall achievements.

Chapter 2

Literature Review

The first operational biometric system for iris has been developed by Daugman at University of Cambridge [18]. To control the illumination the digital images of eye has been captured using near-infrared light source. The algorithm of the image acquisition system adjusts focus of the system to maximize the spectral power, thus is highly robust. In the next step iris in the image is found that uses deformable templates. Some parameters and shape of the eye are used to train a deformable template to guide the detection process [19]. Daugman used a model of iris where the iris and pupil boundaries are circular thus the boundary of circle can be described using three parameters: center of the circle x_0, y_0 and radius r[20]. The operator is defined as

$$max(r, x_0, y_0)|G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r,x_0,y_0} \frac{I(x,y)}{2\pi r} ds| \tag{2.1}$$

where $G_\sigma(r)$ is a blurring function and $I(x, y)$ is the image of the eye. The entire image domain (x, y) is searched by the operator for the maximum in the blurred partial derivative with respect to the increasing radius (r) of the normalised contour integral of the image $I(x, y)$ along a circular arc ds with centre coordinates (x_0, y_0) and of radius r. After the segmentation of the iris, features of the iris is obtained for comparison. The biggest difficulty in iris comparison is that, all iris images are of different size. The iris features should be invariant to change in scale, size, orientation, etc. The distance between the eye and the camera affects the size of iris in an image. The problem of linear deformation of the iris pattern due to change in orientation

of iris due to head tilt, camera position, movement of eyeball, etc, and change in illumination that causes pupil to dilate or contract has been addressed by Daugman by mapping the iris into a dimensionless polar coordinate system [20]. The matching score is generated by the similarity between two iris representations.

A different approach was followed in the iris biometric system that was developed at Sarnoff labs [21]. The authors have used a diffused source of light with low level light camera for the image acquisition. Segmentation of pupil and iris was done using Hough transform. Laplacian of Gaussian filter at multiple scales was used for matching two iris images to produce template and compute the normalised correlation as a similarity measure [21]. This chapter discusses about work done in two most significant areas of iris recognition which are preprocessing and feature extraction.

2.1 Iris Segmentation

Iris preprocessing involves finding the pupil and iris boundaries from the image of the eye. The pupil and the iris are presumed to be circular. To further improve the localisation performance few authors have also worked on detecting eyelids/eyelashes [7]. As mentioned already, Daugman used an integro-differential operator for iris localisation but since the location of iris varies for different images; so the global search reduces speed. To address this Huang et. al [22] proposed coarse to fine strategy. The technique first finds the outer iris boundary in the rescaled image, then uses this information to find iris circles using intergro-differential operator. Further the method for detection of eyelids and eyelashes have been proposed by the authors. Eyelids can detected by searching two curves that satisfies polynomial equation $x(t) = at^2 + bt + c$, $t \in [0, 1]$. Eyelashes can also be detected by checking the variance for each block.

In [23], the authors have improved the localisation speed by using canny edge detector with Hough transform. With this approach, the center and the inner edge was found by using normal line algorithm, while the outer edge was found using homocentric circle algorithm. The authors in [24] uses the bisection method to find the inner boundary. The eyelid position is used to find outer boundary as it is difficult to locate the outer boundary when the iris image is blurred.

Some authors have also used thresholding based approaches for the localisation of

the pupillary region. The authors in [25] takes the pixels below a threshold as pupil and then find the circles using Hough transform and edge detection in the limited area. Further, an iris segmentation based on adaptive threshold is proposed in [26]. In this approach, the iris image is first divided into rectangular regions and intensity means is obtained for each of the region. Then the minimum value of the mean is used as threshold for converting the image into binary. Further, split and merge algorithm is used in [27] to detect connected regions in the image. Authors in [28] have used a concept similar to Daugman for iris segmentation. Authors in [28] have used a concept for iris segmentation which was similar to Daugman. First the irregularities are removed using bilinear interpolation, then the candidate locations are generated to provide the initial conditions for pupil and iris boundary. After that, pupil and iris parameters are recovered for each seed (x, y) . Authors in [29] find the pupil using least significant bit planes. The authors in [6] proposed a novel sector based method for the iris segmentation and achieved an accuracy of 99.07% on BATH database and 95.76% with CASIA database. The iris image was converted into binary using an adaptive thresholding and the inner iris boundary was detected. The outer iris boundary was obtained by first taking the sum of concentric circles of incrementing radii starting from the pupil radius. Then the difference of adjacent circles were obtained. The circle having the maximum difference were taken as the outer boundary of the iris. Further the noise due to eyelids and eyelashes were removed by using a sector based approach.

Some work has also been proposed in the direction of localisation of non-cooperative iris. The authors in [30] have implemented the segmentation methodology that was proposed by Tuceryan [31]. It used the moments of the image in small windows as texture features and then used a clustering algorithm to segment the image. Further in [32] a more robust segmentation approach for the non-ideal images has been developed using graph cuts. Performance of some selected localisation approaches is given in Table 2.1.

Table 2.1: Performance of some selected localisation approaches (taken from [7])

First Author	Approach	Database	Results
Camus [28]	Multiresolution coarse to fine strategy	670 images without glasses and 30 with glasses	99.5% without glasses and 66.6% wearing glasses
Sung [24]	3176 images	Bisection method, canny edge detector and histogram equalisation	100% inner boundary and 94.5% for collarette boundary
Bonney [29]	108 CASIA v1 and 104 UNSA	Least significant bit planes	Pupil detection 99.1% and limbic detection 66.5%
Liu [33]	Modification to Hough transform	4249 images	97.08% Rank 1 recognition
Proenca [30]	Moments	1214 good quality images, 663 noisy images	98.02% good dataset and 97.88% noisy dataset
Pundlik [32]	Graph Cuts	WVU Non-ideal database	Pixel label error rate 5.9%

2.2 Iris Recognition

Significant researches for feature extraction and representation exist in the literature. Daugman has used Gabor filter to obtain binary representation of iris [10]. He used a 2D wavelet demodulation approach on 4258 different iris images and obtained a correct match rate of 100%. In [34] the texture are represented using Gaussian filter. A local orientation at each pixel is obtained by convolving the gradient vector field of an iris image with a Gaussian filter from normalised iris image. The angle is quantized into six bins. This method has been tested on 2255 images of CASIA database and was found to have a correct recognition rate of 100%. In [35], dyadic wavelet transform has been applied on a sequence of 1-D intensity signals around the inner part of the iris to create a binary iris code. The system achieved 100% correct recognition rate with an EER of 0.07%. Modified Log-Gabor filters are used in [36] because unlike Log-Gabor filters Gabor filters are not bandpass filters. In [37] Discrete Cosine Transform (DCT) is used for feature extraction. DCT is applied to the rectangular patches rotated at 45 degrees from radial axis. By keeping the three

most discriminating binarized discrete cosine transform coefficients the dimensionality of feature set is reduced. This approach has been tested on 2156 CASIA images and is reported to have 100% accuracy. In [38] the authors have done texture analysis by computing the analytic image. The analytic image can be obtained by taking the sum of the original image signal with Hilbert transform of the original signal.

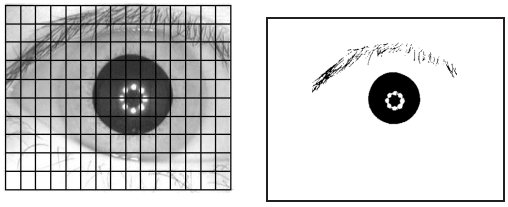
Chapter 3

Iris Segmentation

The image captured by the image acquisition system contains a larger portion of image that includes data from immediately surrounding eye region [21]. Thus before the extraction of features it is necessary to localize only the effective portion of the iris. As per [6], the steps followed for the segmentation of noise free iris from the acquired image are (i) to remove the effect of specular highlights from pupil, (ii) to localise the inner and outer iris circles, (iii) to remove the noise due to the upper and lower eyelids. However, it has been observed in our research that the presence of noise due to eyelids causes problem for localization of the outer iris circle with the existing approach. Hence, in the proposed research the method to localize the outer iris circle in [6] has been modified. The detailed description of all the steps involved in the preprocessing are given as follows:

3.1 Removal of Specular Highlights

Pupil is a dark circular disk-like shape in the eye with significantly low occlusion. However, it contains some specular highlights which are non-singular features in the pupil. The position of the specular highlights is determined by the position of the light source. This light spot forms a hole in the pupil that has to be detected and filled. To begin the hole filling, the input iris image is converted into binary by using an adaptive threshold. The adaptive value of threshold allows the binarization to work under varying illumination conditions. To calculate the suitable value of threshold,



(a) (b)

Figure 3.1: Adaptive image thresholding using grid based approach: (a) Iris image with blocks to compute threshold (b) Binary image obtained using adaptive threshold (taken from [1])

the input image is divided into blocks of size $w \times w$ (shown in Figure 3.1(a)). The mean of the intensities is obtained for each block. The threshold (T) is computed by taking the minimum value of the mean [6]. As the pupil is the darkest portion in an eye image, the block with minimum value of mean will lie on pupillary area. The input image is compared against T to obtain the binary image as shown in Figure 3.1(b).

The image obtained in Figure 3.1(b) contains light spots known as specular highlights. These spots are detected and filled to efficiently localize the pupil. These holes are filled by using morphological region filling approach [6] [39]. Hole filling operation begins by taking the complement of the binary image. The boundary pixels are labelled as 1. If p is a point inside the hole (boundary) the 1 is assigned to it and the following transformation fills the region with ones

$$X_k = (X_{k-1} \oplus B) \cap A^c \tag{3.1}$$

where $X_0 = p$, $k = 1, 2, 3....$ $\oplus$ is used for dilation of X_{k-1} by B which is defined as

$$X_{k-1} \oplus B = \{z|(\hat{B})_z \cap X_{k-1} \neq \phi\} \tag{3.2}$$

B is the symmetric structuring element defined as

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Algorithm 1 Hole_Filling

Require: A: Binary Image, B: Structuring element, p: Point inside the boundary, r: Rows, c: Columns
Ensure: H: Hole filled Image
$C \Leftarrow A^c$ {Complement of an image}
$X_0 = zeros(r, c)$
$X_0(p) = 1$
$k \Leftarrow 0$
repeat
$k \Leftarrow k + 1$
$X_k = (X_{k-1} \oplus B) \cap C$
until $X_k \neq X_{k-1}$
$H = X_k \cup A$

This algorithm terminates at k^{th} iteration if $X_k = X_{k-1}$. The image generated from last iteration X_k is combined with A using bitwise OR that contains the boundary filled image. Algorithm 1 [1] describes the steps involved in hole filling.

3.2 Pupil Detection

After removing the specular highlights from the pupillary region the next step is to obtain the center and the radius of the pupil. To obtain the center of the pupil, the distance of every pixel from the nearest non-zero pixel is obtained in the complemented hole filled image. The point having the maximum distance from its nearest non-zero pixel is the center of the pupil and the distance from the pupil radius to the nearest non-zero pixel is the pupil radius [6]. The results obtained at different steps of pupil detection are shown in Figure 3.2. The algorithm for detecting pupil center and radius is given in Algorithm 2 [6].

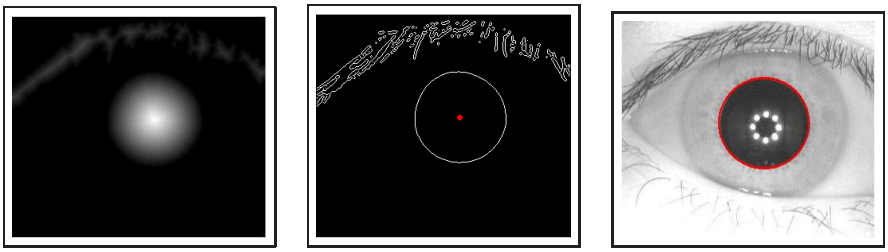
3.3 Iris Detection

Iris detection involves the detection of iris radius from the image. First the image is blurred using median filter to remove external noise, then the contrast of the image is enhanced by using histogram equalisation to have sharp variation at image boundaries [39] as shown in Figure 3.3 (a). This contrast enhanced image is then used for finding the boundary of outer iris drawing concentric sectors (θ = 35 to 145 and θ = 215

to 325) of different radii from the center of the pupil and the intensities lying over the perimeter of the sectors are summed up (Figure 3.3 (b) shows an example). Note that, the sector is taken instead of complete circle to remove the effect of noise due to eyelids. The sector having the maximum change in intensity with respect to the previous drawn sector represents the the iris outer boundary as shown in Figure 3.3 (c). The algorithm for detection of iris radius (r_i) is given in Algorithm 3 [1].

3.4 Annular Iris

After localizing the pupil and iris from the image we obtain an annular region that contains the iris. Further, the noise due to eyelashes and eyelids should be detected and removed from the annular region. In a normal gaze, approximately half of the upper iris circle and one-fourth of the lower iris circle are covered by the edge of the upper and lower eyelids respectively. However, such occlusions are not there in the left and the right regions. It has been observed that the ranges of angular values (θ) for the regions that are not occluded due to eyelids are given by $[35°, 145°]$ and $[215°, 325°]$ and for the upper and lower region, only partial values are taken from a sector [6]. Given the pupil center (x_c, y_c), pupil radius (r_p) and iris radius (r_i) the



(a) (b) (c)

Figure 3.2: Pupil Detection: (a) Spectrum image (b) Edge detected image with pupil center (c) Pupil localised image

Algorithm 2 Pupil_Detect

Require: H: Hole Filled Image

Ensure: x_c: xcenter of pupil, y_c: ycenter of pupil, r_p: Radius of pupil {Estimation of pupil center}

$C \Leftarrow H^c$ {Complement of hole filled image}

$[x\ y] = find(C == 1)$ {Find location of ones in an image}

$l \Leftarrow length(x)$ {To find the number of elements in an array}

for $i = 1$ to r **do**

  for $j = 1$ to c **do**

    for $k = 1$ to l **do**

      $D_k \Leftarrow \sqrt{(x_k - i)^2 + (y_k - j)^2}$

    end for

    $DN = sort(D)$ {Sort the values in D in increasing order}

    $S_{i,j} = DN_1$ {Take the smallest value of DN}

  end for

end for

$[x_c\ y_c] \Leftarrow max(S)$

$E = edge(C)$ {Edge detection using Canny}

$j \Leftarrow y_c$ {Estimation of pupil radius}

$r_p \Leftarrow 0$

while $E_{x_c,j} \neq 1$ **do**

  $r_p = r_p + 1$

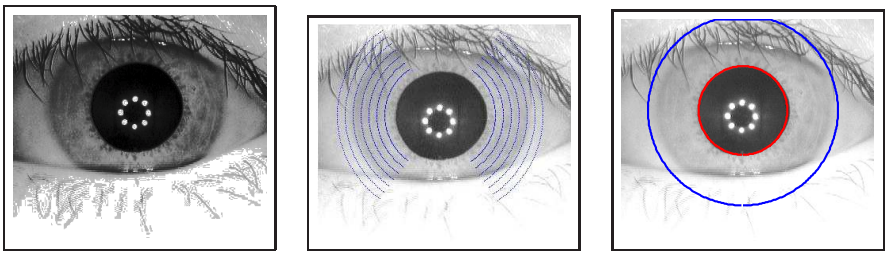
  $j = j + 1$

end while

value of r_i changes depending upon the range of θ as defined by

$$r_i = \begin{cases} \frac{3}{4}r_i & \text{if } 0^\circ < \theta < 35^\circ \\ r_i & \text{if } 35^\circ \leq \theta \leq 145^\circ \\ \frac{1}{2}r_i & \text{if } 145^\circ < \theta \leq 215^\circ \\ r_i & \text{if } 215^\circ < \theta \leq 325^\circ \\ \frac{3}{4}r_i & \text{otherwise} \end{cases} \tag{3.3}$$

The quantisation scheme of (3.3) is used to obtain sector based annular iris image. Figure 3.4 (b) shows the geometrical representation of regions on annular iris circle that are taken into consideration. The annular iris image thus obtained is free from the problem of aliasing. The final preprocessed image is shown in Figure 3.4 (c).



(a) (b) (c)

Figure 3.3: Iris Detection: (a) Contrast enhanced image (b) Concentric circles of different radii (c) Iris localised image

Algorithm 3 Iris_Detect

Require: I: Input image, r_p: Radius of pupil, x_c: xcenter of pupil, y_c: ycenter of pupil

Ensure: r_i: Radius of iris

$F \Leftarrow$ medianFilt(I){Median Filtering on input image}
$H \Leftarrow$ Histeq(F) {Histogram equalisation}
$[r\ c] \Leftarrow$ size(I) {Finding image dimensions}
{Finding the intensity over perimeter of the sector}
for $r_i = r_p \times 1.5$ to $\frac{r}{2}$ **do**
  $sum_{r_i} \Leftarrow 0$
  for $\theta = 35$ to 145 and $\theta = 215$ to 325 **do**
    $x = x_c + r_i \times \cos(\theta)$
    $y = y_c + r_i \times \sin(\theta)$
    $sum_{r_i} = sum_{r_i} + H_{x,y}$
  end for
  $r_i = r_i + 2$
end for
{Change in intensity over circumference}
for $i = 1$ to r_i **do**
  $D_i = |sum_i - sum_{i+1}|$
end for
$[d\ r_i] = \max(D)$ {Maximum change in intensity}

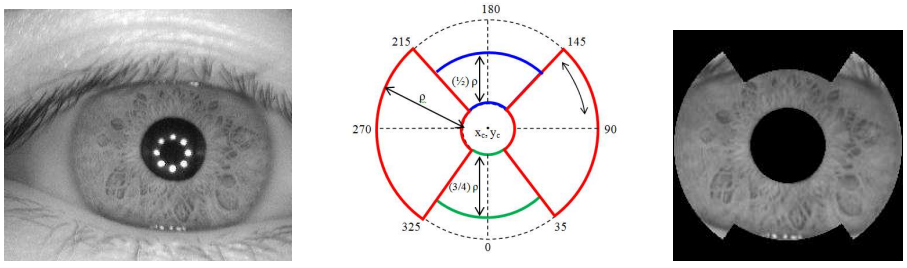


Figure 3.4: Preprocessing of iris image: (a) Input iris image, (b) Geometrical representation of sectors on iris circles, (c) Noise independent annular iris image after preprocessing (taken from [1])

Chapter 4

Keypoint Descriptors for Iris

Feature extraction refers to describing an input image with a simplified amount of information. There already exists various global feature extraction techniques for iris [40, 41]. The main drawback of these global techniques is that they vary with change in illumination, pose and viewpoint. Local features on the other hand are invariant to image scaling, translation and rotation, and partially invariant to change in viewpoint and illumination. These local features perform well under partial occlusions as well. To extract local features from iris, some special points known as *keypoints* are detected where there can be line endings, an isolated point of local intensity maximum or minimum, a corner, or a point on a curve where the curvature is locally maximal. Around the neighborhood of every keypoint a descriptor that is robust to displacements, noise and geometric and photometric deformations is taken that represents the feature vector [3].

In the proposed work an effort has been made to extract the local features from the annular iris image. As discussed earlier the reason for considering annular iris over the polar transformation is to overcome the aliasing errors. To begin with Scale Invariant Feature Transform (SIFT) has been applied to the iris to obtain the keypoint descriptor [2]. SIFT descriptors are invariant to scale, translations and rotation but it suffers from the problem of high computational cost for matching due to higher dimensional descriptor. Speeded Up Robust Features (SURF) [4] uses a faster keypoint detection scheme with reduced dimensional descriptor. Although the descriptors of SURF are more robust than SIFT, yet both SIFT and SURF suffers from the problem

of false pairing. To overcome the problem of false pairing F-SIFT has been used in which Fourier Transform has been applied after detecting SIFT keypoints [**?**]. With the use of Fourier Transform with SIFT the problem of false pairing reduced significantly. Due to inherent advantages of application of Fourier transform with SIFT, S Transform has been used along with SIFT in the proposed work. The resulting algorithm is termed as S-SIFT.

4.1 Scale Invariant Feature Transform (SIFT)

Scale Invariant Feature Transform (SIFT) is a widely used algorithm to extract features from the images [2]. The descriptors obtained by SIFT are invariant to scale, translations and rotation. However, they are less sensitive to the local image distortions. For computing the local features of an image cascade filtering approach can be used to minimize the feature extraction cost as it applies more expensive operations at those locations which pass an initial test. The difference of Gaussian images (DOG) are used to detect keypoints. Local image gradients are measured during feature extraction, at selected scale in the region around each keypoint so as to form a descriptor vector. Following subsections contain a more detailed description of the steps which have been outlined above.

4.1.1 Keypoint Detection

In the first step the potential keypoints which do not vary according to scale and orientation is found. A detailed model is fit to determine the location and scale for each of the detected keypoints. Then based on image gradients, an orientation is assigned to each location. The steps for keypoint detection are explained below.

Detection of Scale Space Extrema

The first step of keypoint detection involves identification of locations that can be assigned with a change in view and scale. Such locations, that are invariant to scale changes, are found by searching stable features across all the possible scales using scale space that is a continuous function of scale [2]. Gaussian function is the only

possible scale space function. Therefore the scale space of image is defined as,

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{4.1}$$

where $I(x, y)$ is the input image and $*$ is the convolution operation in x and y. $G(x, y, \sigma)$ is the variable scale Gaussian defined as

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{(-x^2+y^2)/2\sigma^2} \tag{4.2}$$

DOG (Difference of Gaussian) function is convolved with the image to detect stable keypoint locations. For two nearby scales of an iris image I, the Difference of Gaussian (DOG) is computed as

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \tag{4.3}$$

where k is a constant multiplicative factor that is used for changing the scale and x, y are the coordinates of a pixel in image I. Figure 4.1 illustrates the scale space for two different scales. This technique is scale invariant, hence is appropriate for annular iris images as the dimension of iris varies due to dilation and contraction of the pupil. Figure 4.2 shows the iris images, which have been blurred using the Gaussian filter and computation of DOG for change in octave, scale and σ. These images are generated using SIFT code in [42].

Keypoint Localisation

To detect the interest points, DOG images are used and local maxima and local minima are computed across different scales. Each pixel of a DOG image is compared to 8 neighbors in the same scale and 9 neighbors in the neighboring scales. If a particular pixel is either local maxima or minima in its 3×3×3 neighbouring region as shown in 4.3, then it is selected as a candidate keypoint.

After keypoint detection, the next step is performing the detailed fit to the adjoining data for location, scale and the ratio of principal curvature. The basic idea behind this is to reject all those keypoints which are low in contrast. These low con-

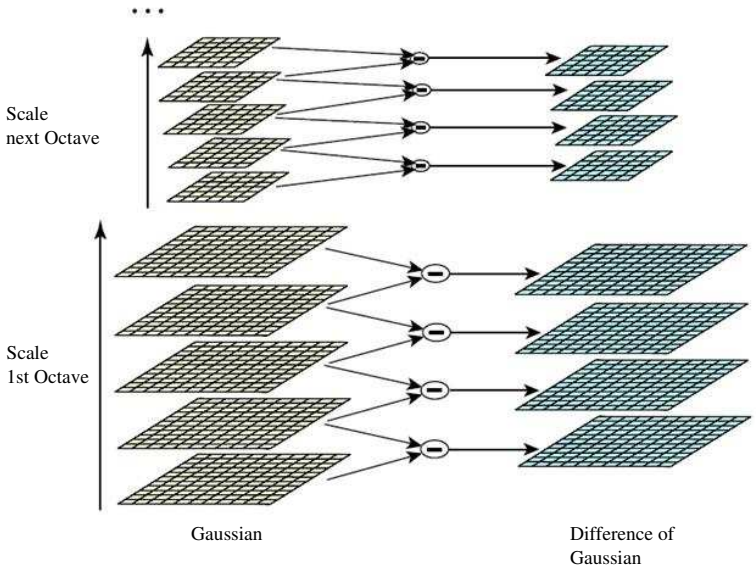


Figure 4.1: Scale space extrema for different octaves. Adjacent Gaussian images are subtracted to produce DOG images on right (taken from [2])

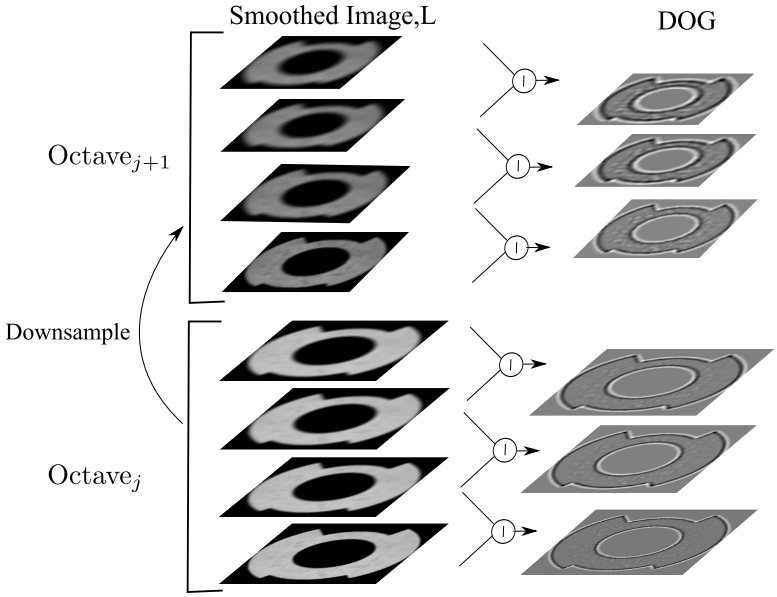


Figure 4.2: Detection of scale space extrema

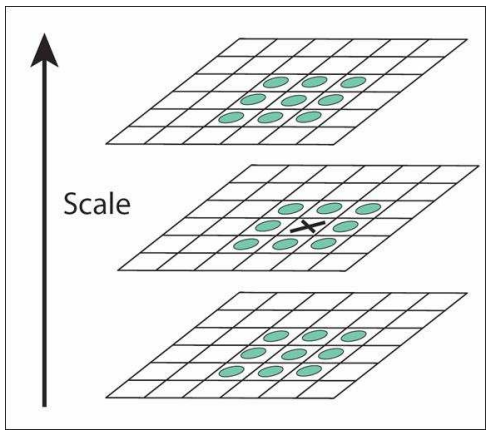


Figure 4.3: Maxima and minima of DOG images are obtained by comparing a pixel to 26 neighbors in $3 \times 3 \times 3$ regions (taken from [2])

trast keypoints are not considered because as stated in [2], such keypoint are sensitive to noise or poorly localized. 3D quadratic function is fitted to local keypoint [43] in order to determine the interpolated location of maximum. The authors have used Taylor expansion of the scale space function, $D(x, y, \sigma)$ shifted so that the origin lies at the sample point

$$D(\mathrm{x}) = D + \frac{\partial D^T}{\partial \mathrm{x}}\mathrm{x} + \frac{1}{2}\mathrm{x}^T\frac{\partial^2 D}{\partial \mathrm{x}^2}\mathrm{x} \tag{4.4}$$

where D and its derivatives are calculated at the sample point and $\mathrm{x}=(x, y, \sigma)^T$ is an offset from this point. The location of extremum ($\hat{\mathrm{x}}$) is obtained by taking the derivative of this function with respect to x and setting it to zero, thus giving

$$\hat{\mathrm{x}} = -\frac{\partial^2 D^{-1}}{\partial \mathrm{x}^2}\frac{\partial D}{\partial \mathrm{x}} \tag{4.5}$$

The offset is compared to a predefined threshold and if it is larger, then it implies that $\hat{\mathrm{x}}$ is close to some different sample point. In this case sample point is changed and interpolation is performed about that point. The final offset is then added to the sample point to get the interpolated location of extremum. Figure 4.4 illustrates the interpolation of datapoints to get an estimate of the extremum. Figure 4.5(a) shows a sample iris image after the detection of keypoints.

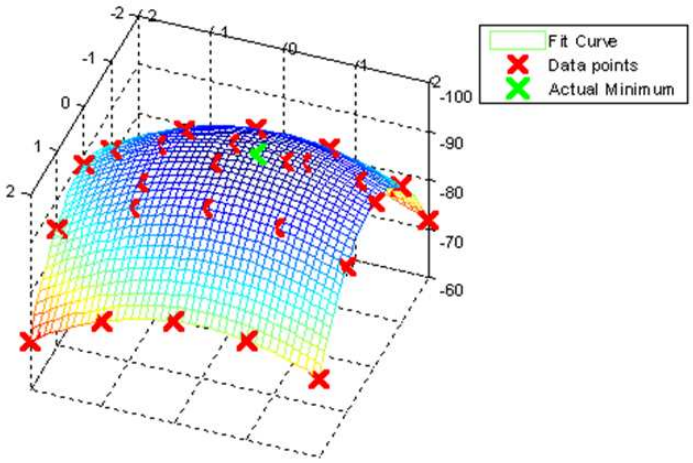


Figure 4.4: Interpolation of datapoints to estimate location of extremum (taken from [1])

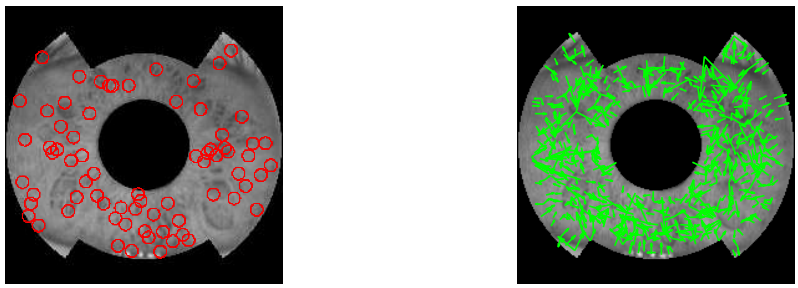
Orientation Assignment

To attain invariance to image rotations, an orientation is assigned to each of the keypoint locations. The descriptor can be represented relative to this orientation. For determination of the keypoint orientation, a gradient orientation histogram is worked out in the neighborhood of the keypoint. A Gaussian smoothed image L is selected using the scale of keypoint. For a Gaussian smoothed image $L(x, y)$, magnitude $(m(x, y))$ and orientation $(\theta(x, y))$ are calculated as

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \tag{4.6}$$

$$\theta(x,y) = \tan^{-1}\left(\frac{(L(x,y+1) - L(x,y-1))}{(L(x+1,y) - L(x-1,y))}\right) \tag{4.7}$$

This is followed by formation of the orientation histogram for gradient orientation around each of the keypoints. The histogram contains 36 bins for 360 orientations and before adding it to the histogram, every sample is weighted by gradient magnitude and Gaussian weighted circular window, with σ of 1.5 times the scale of keypoint. Peaks in histogram correspond to the orientations. Keypoint with the computed orientation is created by using any other local peak within 80% of largest. By doing



(a) (b)

Figure 4.5: Keypoint detection on annular iris image using SIFT (a) Detected keypoints after removing noise and edge responses, (b) Scale and direction of orientation is indicated by arrows

this, stability during matching increases [2]. In figure 4.5(b), the arrows indicates the scale and direction of orientation.

4.1.2 Keypoint Descriptor

After orientation has been selected, the feature descriptor is calculated as a set of the orientation histograms on 4×4 pixel neighborhoods. The orientation histograms and keypoint orientation are relative to each other. This can be seen in Figure 4.6. Histogram has 8 bins each and every descriptor contains an array of 16 histograms around the keypoint. Thus a SIFT feature descriptor is generated, containing $4 \times 4 \times 8 = 128$ elements. The descriptor vector obtained is invariant to scaling, rotation and illumination.

4.1.3 Keypoint Pairing

Let $p = \{p_1, p_2, p_3 ... p_n\}$ and $q = \{q_1, q_2, q_3 ... q_n\}$ be n dimensional feature descriptor for each keypoint from database and query images respectively. The Euclidean distance between p and q is described by

$$D(p, q) = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2} \tag{4.8}$$

where n is a 128 dimensional feature descriptor. A simple approach to nearest neighbor

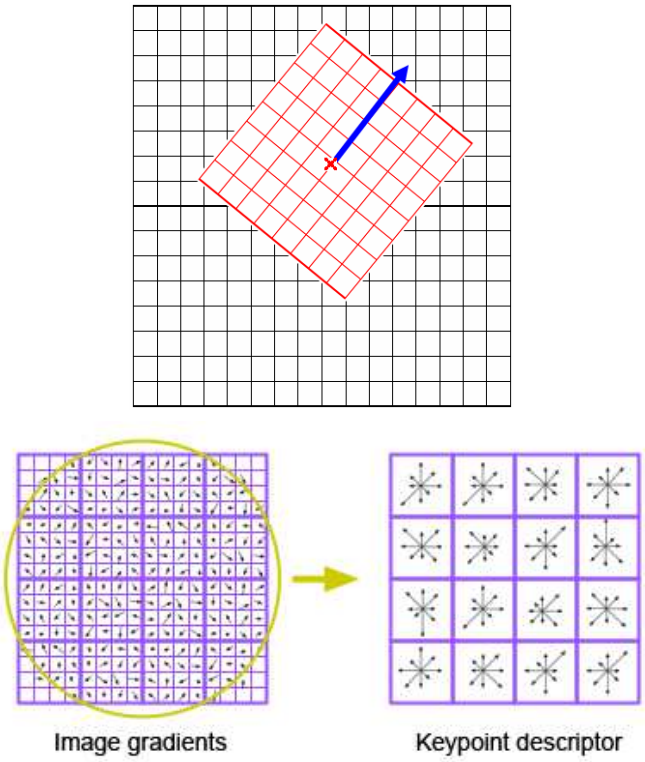


Figure 4.6: Window is taken relative to direction of dominant orientation. This window is weighted by a Gaussian and histogram is obtained for 4×4 regions (taken from [1])

matching is to iterate through all the points in the database so as to determine the nearest neighbor..

4.2 Speeded Up Robust Features (SURF)

In comparison SIFT [2], Speeded Up Robust Features (SURF) detector and descriptor are not just faster, but also far more repeatable and distinctive [4]. The requirement is to apply a feature descriptor which is faster in computation as compared to the existing keypoint approaches, without sacrificing performance. This can be attained by simplifying the detection scheme without compromising on accuracy, and by reducing the descriptors size and at the same time keeping it sufficiently distinctive [3]. The descriptors found using SURF are more stable, repeatable and also performs faster due to low descriptor dimensionality.

SURF has two major advantages over SIFT. First of all, SURF uses the sign of Laplacian to get sharp distinction between the background features and foreground features. Furthermore, SURF uses 64 dimensional vectors whereas SIFT uses 128 dimensional vector. This in turn reduces feature calculation time and also allows quick matching with increased robustness at the same time [44]. Using a Hessian matrix, the operator extracts keypoints and defines the distribution of Haar Wavelet responses from the descriptors, which is actually a window around the interest points. The determination of local descriptor vector is carried out in two steps which are (i) Detection of keypoints (ii) Keypoint descriptor. The above stated steps are explained as follows:

4.2.1 Detection of Keypoints

SURF uses Hessian Matrix approximation for interest point detection. As proposed in [45], integral images are used for faster computation of interest points. The concepts of boxlets proposed by Simard et al. [46] is used in integral images.

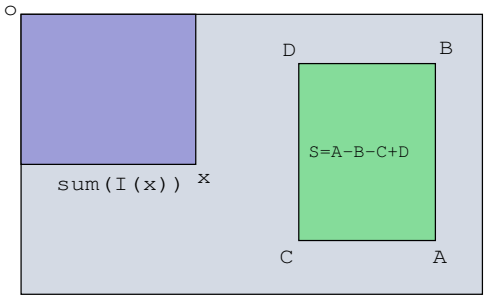


Figure 4.7: Integral images are used to calculate the sum of intensities inside a rectangular region of any size.

Integral Images

The computation time is considerably reduced by integral images which allow faster computation of box-type convolution filters. Given the location x = $(x, y)^T$ and an integral image $I_{\sum}$(x) is entered at that location, it denotes the sum of all the pixels in the input image (I) within a rectangular region which is formed by the origin and x

$$I_{\Sigma}(\mathrm{x}) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(x, y) \tag{4.9}$$

After finding the integral images it takes just three additions to compute sum of intensities over the integral area as illustrated in Figure 4.7. The computation time does not depend on the filter size.

Interest Points by using Hessian matrix

The determinant of Hessian matrix is used for choosing location and scale so as to determine the keypoints. Given a point $P = (x, y)$ within an image I, then the Hessian matrix $H(P, \sigma)$ in P at a scale σ is defined as

$$H(P, \sigma) = \begin{bmatrix} L_{xx}(P, \sigma) & L_{xy}(P, \sigma) \\ L_{xy}(P, \sigma) & L_{yy}(P, \sigma) \end{bmatrix} \tag{4.10}$$

where $L_{xx}(P, \sigma)$ is the convolution of the Gaussian second order derivative $(\frac{\sigma^2}{\sigma x^2} g(\sigma))$ with the image I at the point P. Similarly $L_{xy}(P, \sigma)$ and $L_{yy}(P, \sigma)$ are found. Gaussian is discretized and cropped as illustrated in Figure 4.8. These approximated values

of Gaussian second order derivatives can be calculated at a very low computational cost by using integral images. The computation time is therefore independent of the size of filter. The 9×9 box filters as illustrated in Figure 4.8 are the approximations of a Gaussian at $\sigma = 1.2$ [4]. These are symbolized by D_{xx}, D_{xy} and D_{yy} [47]. The approximated values for the Hessians determinant are computed by choosing the weights for the box filters adequately and using following equation

$$Det(H_{approx}) = D_{xx}D_{yy} - (0.9D_{xx})^2 \tag{4.11}$$

Scale Space Representation

Since box filters and integral image are being used, the iterative application of the same filter to output of the previously filtered image is not required. The computational efficiency can be improved by applying any size box-filter on the original image as described in Figure 4.9. Therefore, the scale space is analyzed by up scaling the filter size rather than by reducing the image size. The output of the 9×9 filter, which was introduced in previous section, is taken to be the initial scale layer. Consequent layers are obtained by filtering the image with larger masks to localize keypoints which do not vary according to the scale. The benefit of such a scale space creation is that it is computationally more efficient, because the image is not downsampled so effect of aliasing is eliminated.

The entire scale space can be divided into octaves, each of which is represented by a series of filter responses which are obtained by convolving the input image with a filter whose size keeps increasing. Every octave is further divided into a fixed number of scale levels. The length (l_0) of the positive or the negative lobe of the partial second

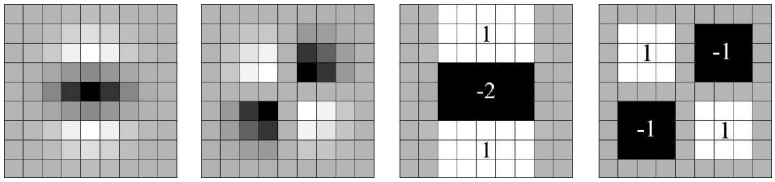


Figure 4.8: Left to right: discrete Gaussian second order derivative in y and xy direction. Approximation for the second order Gaussian partial derivative in $y-(D_{yy})$ and xy-direction (D_{xy}) (taken from [3]).

order derivative in the direction of the derivation (x or y) is set to the third of filter size length. For the 9×9 filter, this length l_0 is 3. For two consecutive levels, this size is increased by minimum 2 pixels (1 pixel on each side) so as to keep the size uneven and therefore ensure the presence of a central pixel. This leads to an increase in the mask size by 6 pixels as depicted in Figure 4.10.

Scale space construction begins with the initial 9×9 filter for which the scale s=1.2. Following this, filters with sizes 15×15, 21×21, and 27×27 are applied. By these filters even more than a scale change of two has been realized. The increase in size of filter is doubled for each new octave (from 6-12 to 24-48). The size of filter is increased for every corresponding octave until the condition is achieved when size of image is larger than the size of filter.

Interest Point Localisation

Keypoints are localized in terms of scale and image space by relating a non-maximum suppression in a 3×3×3 neighborhood. As proposed in [43], the local maxima that is found on the determinant of Hessian matrix is interpolated to the image space. Figure 4.11 illustrates the detected interest points on an annular iris image.

Orientation Assignment

The orientation is recognized for each keypoint so as to achieve invariance to image rotation. To achieve this, Haar wavelet responses are computed in the x and y directions within a circular neighborhood of radius of $6s$ around the interest point, where s is the scale at which the interest point was identified. The wavelet sizes depend on scale and are set to a side length of $4s$. After the wavelet responses are evaluated and

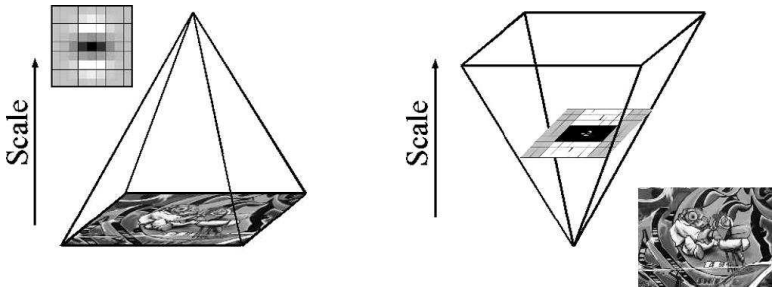


Figure 4.9: Use of integral images for upscaling filter masks (taken from [4])

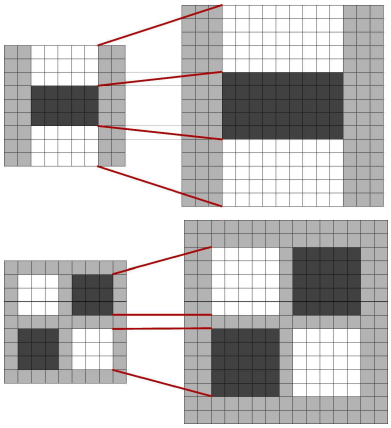


Figure 4.10: Filters D_{yy} (top) and D_{xy} (bottom) for two successive filter sizes (9×9 and 15×15) [3].

weighted with a Gaussian ($\sigma = 2s$), the dominant orientation is attained by adding all the responses within a sliding orientation window of size $\frac{\pi}{3}$ (refer Figure 4.12). The horizontal as well as the vertical responses within the window are added. The longest of all such vectors over all of the windows determines the orientation.

Keypoint Descriptor

The descriptor vector is acquired around each detected keypoint by considering a square window of size of 20s centered on the interest point. The window is aligned relative to the direction of orientation. In order to preserve spatial information, the region is split into smaller 4×4 sub-regions as illustrated in Figure 4.13. For each of the sub-regions, Haar Wavelet responses are attained in the horizontal (d_x) as well as the vertical direction (d_y). The responses (d_x and d_y) are first weighted with a

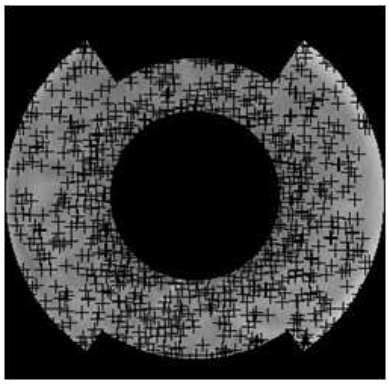


Figure 4.11: Detected interest points on annular iris image

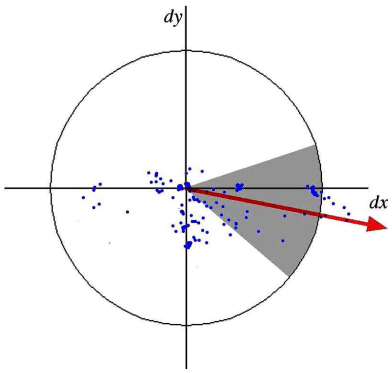


Figure 4.12: Orientation assignment by taking a sliding window of size $\frac{\pi}{3}$ indicated by shaded region [3]

Gaussian ($\sigma = 3.3s$) centered at the interest point, so as to increase the robustness towards geometric deformations and localization errors.

Lastly, the descriptor vector is added up for each sub-region in order to form elements of feature vector. Then the sum of the absolute values of the responses are obtained ($|d_x|$ and $|d_y|$) to bring in information about the polarity of the intensity changes. Therefore, each sub-region is a 4D feature vector comprising of

$$v = \left\{ \sum d_x, \sum d_y, \sum |d_x|, \sum |d_y| \right\} \tag{4.12}$$

Concatenating this for all 4×4 sub-regions leads to a descriptor vector of length 64. The property of a descriptor for three different image-intensity patterns within a sub-region is clearly illustrated in Figure 4.14. The values are comparatively low for

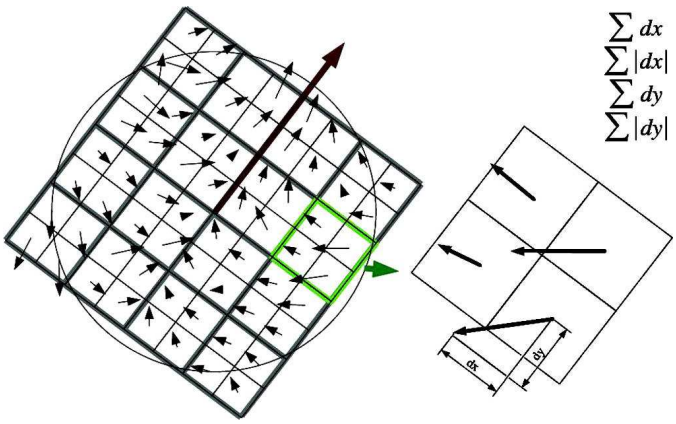


Figure 4.13: An oriented window with 4×4 sub-regions is taken in direction of orientation. For each sub-region wavelet responses are obtained [3].

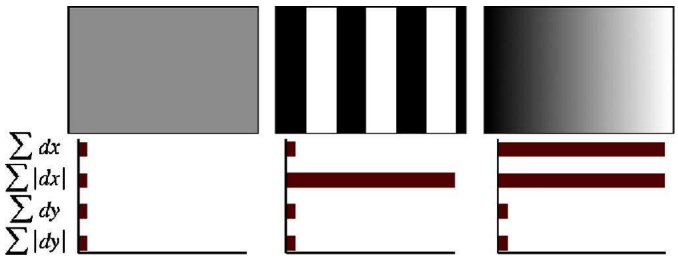


Figure 4.14: Descriptor entries of a sub-region represent the nature of the underlying intensity pattern [3].

a homogeneous sub-region i.e. left of Figure 4.14. The value of $\sum |d_x|$ is high but all others are low (middle), for the presence of frequencies in x direction. But the values of both $\sum d_x$ and $\sum |d_x|$ are high if intensity is such that it slowly decreasing in x direction.

4.2.2 Keypoint Pairing

Once the keypoints in the database image and query image are determined, interest point pairing approach is used to carry out matching. The pairing of the SURF descriptors are done in the same way as discussed in Section 4.1.3.

4.3 F-SIFT

SIFT extract uses gradient information to extract features may not be suitable particularly for iris images. In addition to this, SIFT also suffers from the problem of false pairing as shown in Figure 4.15. This can be minimized by extracting features that are suitable for iris texture pattern. Fourier transform is one such method because of its property of describing periodic function like texture that contains repetitive patterns [5]. Further, SIFT is robust to various transformations, occlusion and illumination [2]. F-SIFT uses the good properties of both Fourier and SIFT. The keypoints are obtained using SIFT and then the phase of the Fourier transform is used to obtain the keypoint descriptor. The details of the steps involved in obtaining the descriptor is discussed below.

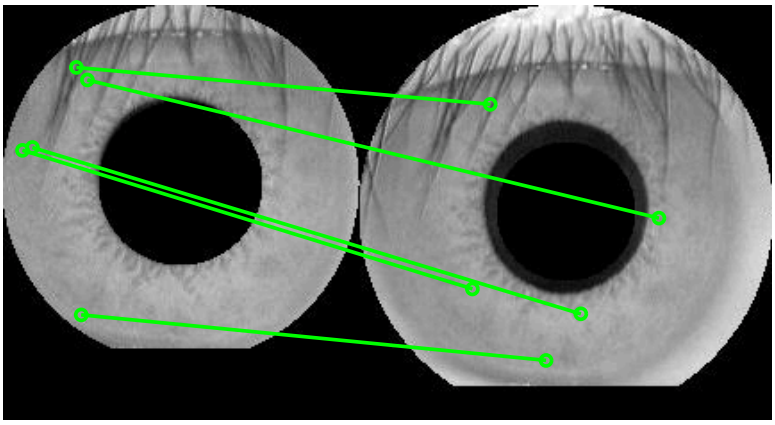


Figure 4.15: Sample keypoints falsely paired using SIFT due to texture similarity (taken from [5])

4.3.1 Keypoint Detection using SIFT

The first step is to obtain all the potential keypoints in the image that are invariant to scale and orientation. SIFT is used to obtain such keypoints as discussed in Section 4.1.1. Each detected keypoint is associated with its location, scale and orientation information as in SIFT [2]. The steps to obtain the keypoints are same as discussed in Section 4.1.1.

4.3.2 Keypoint Descriptor using Fourier Transform

The descriptor vector is obtained by taking a $w \times w$ window around each detected keypoint with center (x, y) relative to the direction of orientation (ϕ). Fourier transform is used to define the local descriptor for each keyblock. The Fourier transform describes the global frequency content of each keyblock efficiently. The phase information of the Fourier Transform is used to robustly define the texture pattern [5]. Each descriptor keyblock is obtained using

$$\begin{aligned} K_i(u,v) &= \frac{1}{W^2} \sum_{n_1=(x-\frac{W}{2})}^{(x+\frac{W}{2})} \sum_{n_2=(y-\frac{W}{2})}^{(y+\frac{W}{2})} I(n_1,n_2) e^{-i2\pi(\frac{n_1 u}{W}+\frac{n_2 v}{W})} \\ &= A(u,v) e^{i\theta(u,v)} \end{aligned} \tag{4.13}$$

where $u < 0$ and $(x - \frac{W}{2}) \leq u \leq (x + \frac{W}{2})$, $v > 0$ and $(y - \frac{W}{2}) \leq v \leq (y + \frac{W}{2})$. $A(u, v)$ are the amplitude component of each keyblock and $\theta(u, v)$ are the phase component

of each keyblock.

4.3.3 Keypoint Pairing

Phase based image matching is used to pair the keypoints between gallery and probe images. The phase information between the i^{th} keyblock in the probe image is paired to the j^{th} keyblock in the gallary image using the Phase-Only Correlation function [5]. Let, $A_i(u,v)$ and $\theta_i(u,v)$ be the amplitude and phase components of Fourier Transform from i^{th} keyblock. Similarly, $A_j(u,v)$ and $\theta_j(u,v)$ be the amplitude and phase components of Fourier Transform from j^{th} keyblock [5]. Thus, the cross phase spectrun between K_i and K_j is calculated using

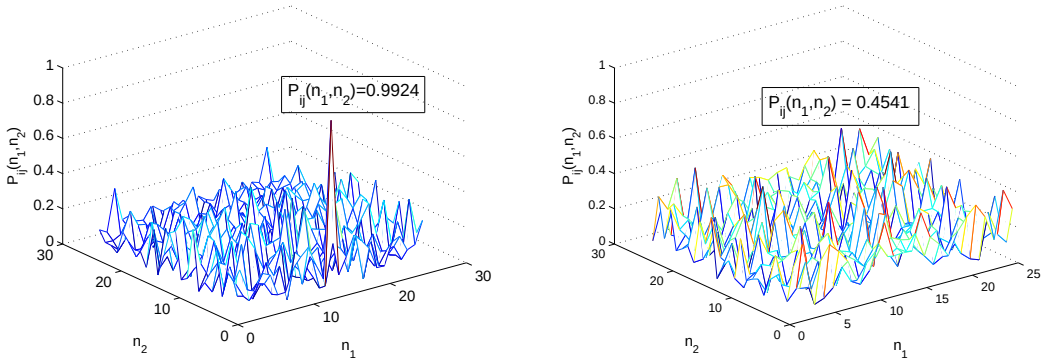
$$R_{ij}(u,v) = e^{i\{\theta_i(u,v)-\theta_j(u,v)\}} \tag{4.14}$$

The POC function P_{ij} is the inverse Fourier transform of $R_{ij}(u,v)$ which is defined as [5]

$$P_{ij}(n_1,n_2) = \frac{1}{W^2}\sum_{u=(x-\frac{W}{2})}^{(x+\frac{W}{2})}\sum_{v=(y-\frac{W}{2})}^{(y+\frac{W}{2})} R_{ij}(u,v)e^{i2\pi(\frac{n_1 u}{W}+\frac{n_2 v}{W})} \tag{4.15}$$

This POC function is used to define the similarity between the two keyblocks. If the two keyblocks are from similar texture regions of iris, then their POC function gives a distinct sharp peak as shown in Figure 4.16(a). However, if keyblocks are from two dissimilar regions of iris the peak drops significantly (refer Figure 4.16(b)). Thus, the height of the peak is taken as a similarity measure for pairing the keyblocks. If peak in P_{ij} is greater a predefined threshold (T) the corresponding keyblocks are paired and removed from the list of unpaired keyblocks. The process is iterated for all the remaining keypoints in the probe set.

The diagrammatic representation of F-SIFT approach is given in Figure 4.17.



(a) P_{ij}: Similar region (b) P_{ij}: Dissimilar region

Figure 4.16: Phase-Only Correlation between similar and dissimilar regions of iris image (taken from [5])

4.4 S-SIFT

The success of using Fourier Transform with the SIFT in F-SIFT [5] gives the motivation to try out other image transforms with the SIFT. In the proposed work S Transform is used with the SIFT to obtain the keypoint descriptors. Some of the properties of the S Transform are given below [17]:

- It gives both time and frequency image unlike Fourier Transform which gives only frequency information.
- It can be derived from Short Time Fourier Transform(STFT) or Wavelet Transform.
- It provides adaptive window superior to fixed window in STFT.
- It retains absolute phase which is not provided by Wavelet Transform.
- It is directly related to Fourier transform.

The keypoints are obtained using SIFT and then the phase of the S Transform is used to obtain the keypoint descriptor. The steps followed to obtain the descriptor in S-SIFT is same as F-SIFT, only the Fourier transform is replaced with S transform. The details of the steps involved in obtaining the descriptor is discussed below.

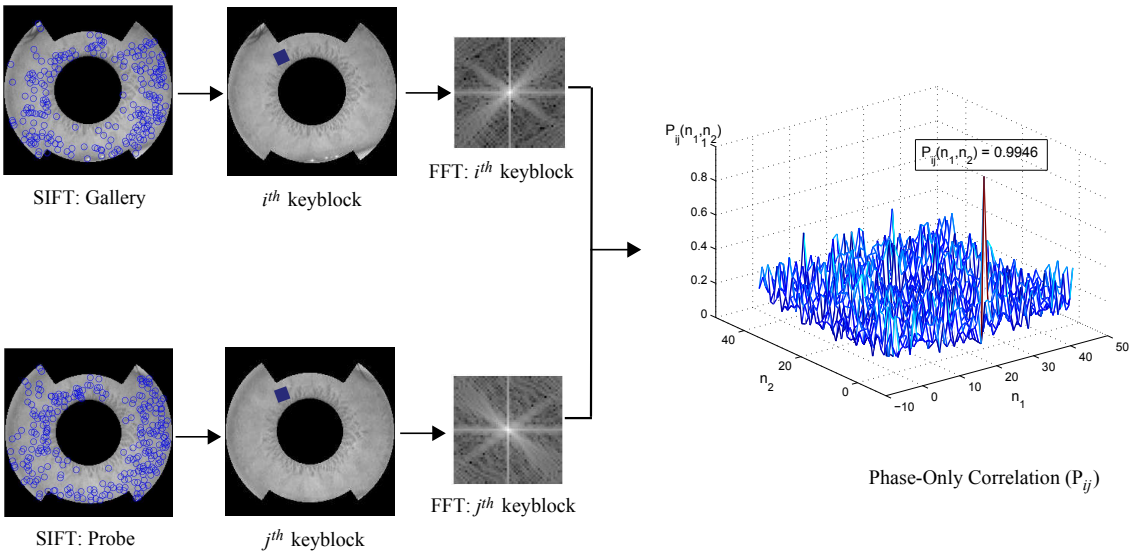


Figure 4.17: Block diagram of F-SIFT approach for matching (taken from [5])

4.4.1 Keypoint Detection using SIFT

The keypoints are detected using the method discussed in Section sect:keypointdetectfsift. Each detected keypoint is associated with its location, scale and orientation information as in SIFT [2].

4.4.2 Keypoint Descriptor using S Transform

The descriptor vector is obtained by taking a $w \times w$ window around each detected keypoint with center (x, y) relative to the direction of orientation (ϕ). Due to the computational complexity of S transform $(O(N^4(logn))$ a small value (4 in this thesis) w is chosen. S transform is used to define the local descriptor for each keyblock. The phase information of the S Transform is used to define the texture pattern. Each descriptor keyblock is obtained using

$$\begin{aligned} K_i(x,y,u,v) &= \sum_{n_1=0}^{W-1}\sum_{n_2=0}^{W-1} I(x+n_1, y+n_2) e^{-\frac{i2\pi^2 n_1^2}{u^2}} e^{-\frac{i2\pi n_1 x}{W}} e^{-\frac{i2\pi^2 n_2^2}{v^2}} e^{-\frac{i2\pi n_2 y}{W}} \\ &= A(x,y,u,v) e^{i\theta(x,y,u,v)} \end{aligned} \tag{4.16}$$

4.4.3 Keypoint Pairing

Similar to F-SIFT Phase-Only Correlation is used to pair the keypoints between gallery and probe images. Let, $A_i(x, y, u, v)$ and $\theta_i(x, y, u, v)$ be the amplitude and phase components of Fourier Transform from i^{th} keyblock. Similarly, $A_j(x, y, u, v)$ and $\theta_j(x, y, u, v)$ be the amplitude and phase components of Fourier Transform from j^{th} keyblock [5]. The cross phase spectrun between K_i and K_j is calculated using

$$R_{ij}(x, y, u, v) = e^{i\{\theta_i(x,y,u,v) - \theta_j(x,y,u,v)\}} \tag{4.17}$$

The POC function P_{ij} is the inverse S transform of $R_{ij}(x, y, u, v)$ which is can be calculated by first computing the sum of all the values for each of the frequency as in 4.18 and then obtaining the inverse Fourier transform of the resultant matrix as in 4.19.

$$R'_{ij}(u, v) = \sum_{x=0}^{W-1} \sum_{y=0}^{W-1} R_{ij}(x, y, u, v) \tag{4.18}$$

$$P_{ij}(n_1, n_2) = \frac{1}{W^2} \sum_{u=(x-\frac{W}{2})}^{(x+\frac{W}{2})} \sum_{v=(y-\frac{W}{2})}^{(y+\frac{W}{2})} R'_{ij}(u, v) e^{i2\pi(\frac{n_1 u}{W} + \frac{n_2 v}{W})} \tag{4.19}$$

As in F-SIFT discussed in Section 4.3, the POC function gives a distinct sharp peak for two keyblocks from the similar texture region of iris, and the peak drops when the keyblocks are from two dissimilar regions of iris. Thus, the height of the peak is used to determine the similarity for pairing the keyblocks.

Chapter 5

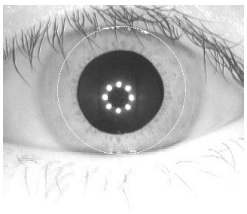
Experimental Results and Analysis

This chapter gives the details of results obtained and the analysis of the results obtained during iris segmentation and feature extraction.

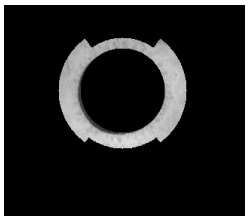
5.1 Iris Segmentation

In this section, the localisation performance of the proposed approach is compared with the approach in [6]. As localisation is the the most significant and fundamental step in iris identification, there should not be any error in the iris boundary detection as far as possible. Also the final image should be free from the noise due to eyelids. Therefore, the eyelids are removed by using a predefined mask. The system has been tested on both cooperative [15] as well as non-cooperative [14] iris databases. Table 5.1 shows the mis-localisation error in the proposed approach and the approach used in [6]. From the table it can be inferred the proposed approach performs better than the segmentation approach proposed in [6]. Figure 5.1 shows the localisation performance of both the approaches approach on a sample image from CASIA database. The session id, individual id, image instance and eye information (left or right) is given under the displayed result. For an image of CASIA database *S1001L01*, *S1* denotes $sessionid = 1$, *001* denotes $individulaid = 1$, *L* denotes the left eye and *01* denotes $imageinstance = 01$.Though the outer iris circle is not localised correctly but annular region contains sufficient information to extract features.

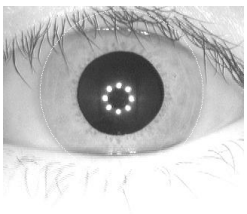
From Figure 5.1, it can be seen that the approach used in [6] crops a majority of



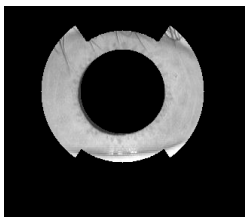
(a) Iris radius detection using [6]



(b) Annular iris using [6]



(c) Iris radius detection using proposed approach



(b) Annular iris using proposed approach

Figure 5.1: Localisation performance using [6] approach and proposed approach on image no S1001L01 of CASIA database

Table 5.1: Mis-localisation percentage of Sector Based approach of [6] and proposed approach

Database	**Sector Based [6]**	**Proposed**
CASIA	7.29	4.2
UBIRIS	15.53	4.5

the area of iris that could contain some significant feature points. The reason behind this is that, this approach considers the sum along the complete circle while detecting the outer iris boundary. In most of the cases, the upper and the lower half of the outer iris boundary is covered by the eyelids. Hence, the edge of the eyelids is detected as the outer iris boundary and thus crops a significant portion of iris. This drawback is addressed in the proposed approach and can be clearly seen in the Figure 5.1.

However, the proposed masking approach fails to extract noise independent annular iris if the degree of occlusion by upper and lower eyelids exceeds $\frac{1}{2}$ and $\frac{1}{4}$ respectively. Few such cases are shown in Figure 5.2. In such cases the iris is completely hidden inside the eyelids. So even the adaptive eyelid detection could not recover the hidden features from the iris. However, there is still a need to develop an adaptive masking approach eyelid because the proposed approach masks the iris even if there

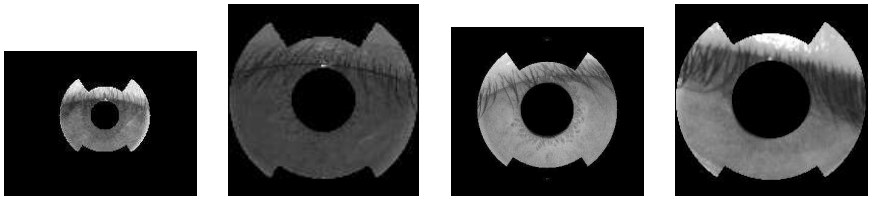


Figure 5.2: Failure to generate noise independent annular iris due to greater degree of occlusion by upper eyelid

is no occlusion by the upper and the lower eyelids.

5.2 Iris Recognition

In this section, the performance of all the feature extraction algorithms discussed in Section 4 are compared with each other on various performance measures discussed earlier. All of these algorithms (SIFT, SURF and F-SIFT) were applied on CASIA V3 database [15] and compared with the proposed algorithm S-SIFT. Figure 5.3 shows the ROC curve for all the four algorithms. Genuine and Imposter score distributions using Dual stage approach is given in Figure fig:rocsift.

From the Figure 5.3 it is evident that the proposed work does not improve the accuracy. All the existing methods perform better than the proposed algorithm. Among the existing algorithms F-SIFT has the best performance. The reason for the poor performance of S-SIFT may be:

- **Inappropriate matching algorithm:** POC only uses the phase information to match two different keyblocks. It may so happen that two different regions with similar textures can be matched by POC.
- **Insufficient utilization of S Transform:** S Transform gives time, frequency and phase information of all the points in the image. However, in S-SIFT only the phase information of the S Transform is being utilized.
- **Smaller window size across the keypoints:** Due to high computational complexity of S Transform we have to compromise with a smaller size of the window. In the experiment the windows size was taken as 4×4. The smaller

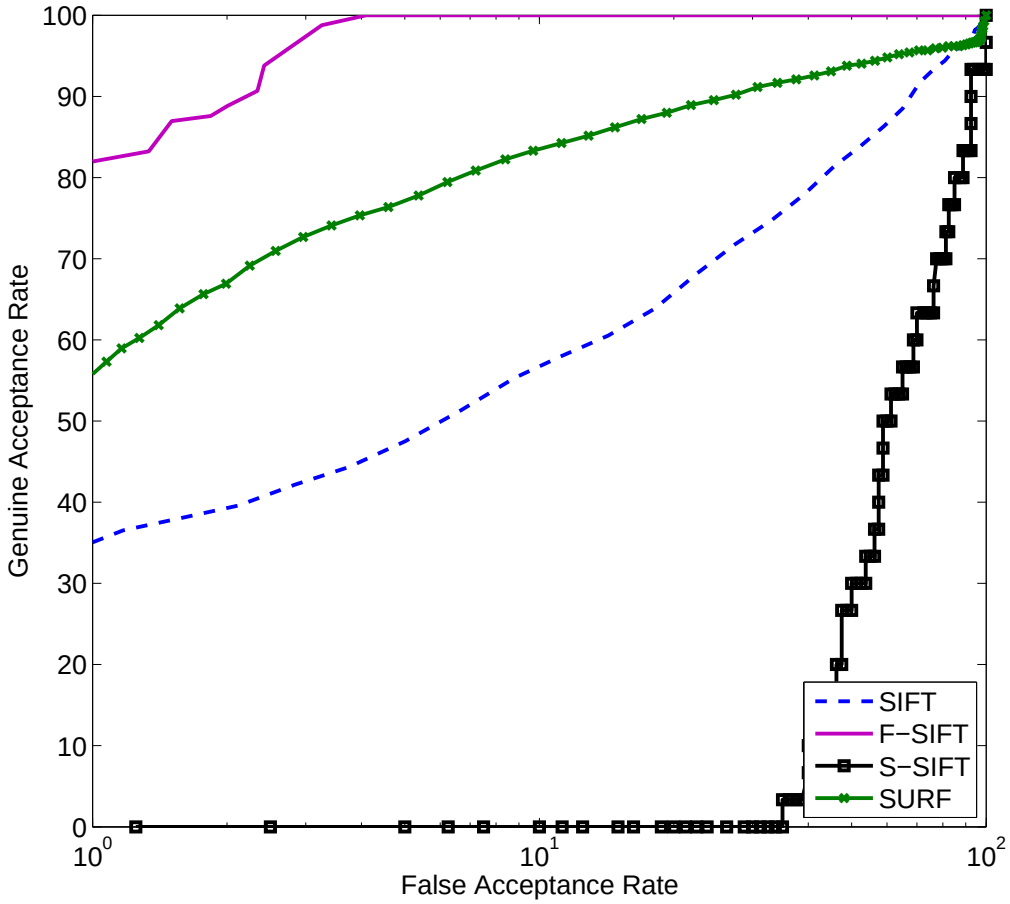
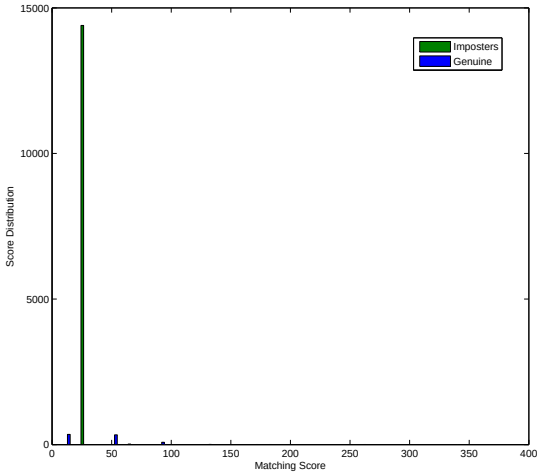


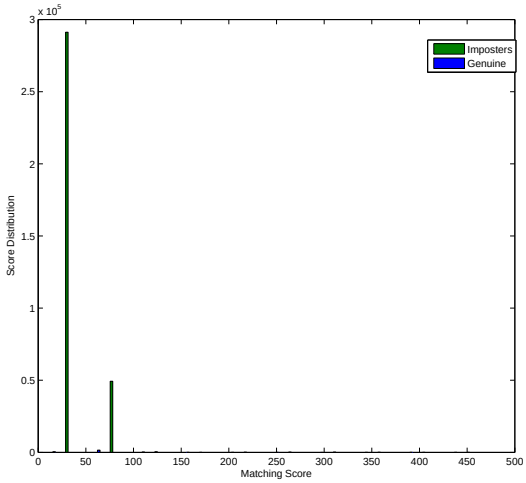
Figure 5.3: ROC curve for annular iris image

window may not be properly able to provide a unique and distinguishable descriptor.

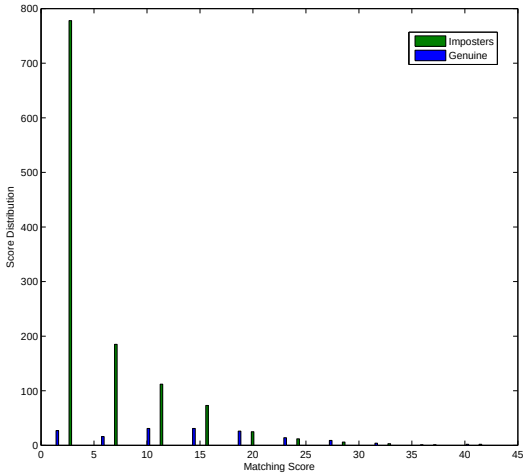
- **False pairing of keypoints:** It may so happen that the keypoints are being paired to the wrong keyblocks. This could be reduced by adopting a region based matching where the keypoint in one region can only be matched to the other keypoint if it lies in that particular region.



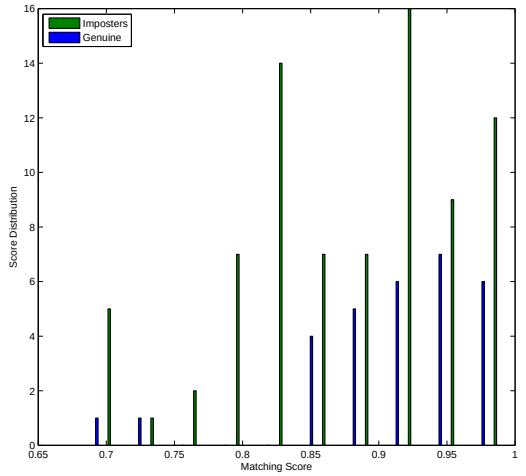
(a) SIFT



(b) SURF



(a) F-SIFT



(b) S-SIFT

Figure 5.4: Genuine and Imposter score distribution curve

Chapter 6

Conclusions

In this thesis, a new sector based algorithm for iris segmentation was proposed. The proposed algorithm was implemented and compared with the existing sector based approach using publicly available databases. The proposed approach outperforms the existing approach with a large margin. Further, a new approach for local feature extraction was proposed in the thesis. In addition to this, three other local feature extraction approach viz. Scale Invariant Feature Transform (SIFT), Speeded-Up Robust Features (SURF) and F-SIFT were implemented. These three algorithms along with the proposed algorithm were tested and compared using a publicly available database. Though the proposed approach failed to meet the accuracy of the existing algorithms, several reasons of the failures were analysed and reported. The research can be carried out in future to take care of the causes that led to the failure of S-SIFT.

Bibliography

[1] H. Mehrotra. Iris Identification using Keypoint Descriptors and Geometric Hashing. Master's thesis, NIT Rourkela, 2010.

[2] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[3] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-up robust features (surf). *Computer Vision and Image Understanding*, 110(3):346 – 359, 2008.

[4] H. Bay, T. Tuytelaars, and L. V. Gool. Surf: Speeded up robust features. In *9th European Conference on Computer Vision*, Graz Austria, 2006.

[5] H. Mehrotra, B. Majhi, and P.K. Sa. Unconstrained iris recognition using f-sift. In *Information, Communications and Signal Processing (ICICS) 2011 8th International Conference on*, pages 1–5, 2011.

[6] H. Mehrotra S. Bakshi and B. Majhi. Real-time iris segmentation based on image morphology. In *Proceedings of the 2011 International Conference on Communication, Computing & Security*, ICCCS '11, pages 335–338, New York, NY, USA, 2011. ACM.

[7] K.W. Bowyer, K. Hollingsworth, and P. J. Flynn. Image understanding for iris biometrics: A survey. *Computer Vision and Image Understanding*, 110(2):281–307, 2008.

[8] R. Bolle and S. Pankanti. *Biometrics Personal Identification in Networked Society*. Kluwer Academic Publishers, Norwell, MA, USA, 1998.

[9] A. K. Jain, A. Ross, and S. Prabhakar. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14:4–20, 2004.

[10] J. Daugman. The importance of being random: statistical principles of iris recognition. *Pattern Recognition*, 36(2):279 – 291, 2003.

[11] S. Sirohey, A. Rosenfeld, and Z. Duric. Eye tracking. Technical Report CAR-TR-922, Center for Automation Research, University of Maryland, College Park, 1999.

[12] J. L. Wayman. Error rate equations for the general biometric system. *IEEE Robotics & Automation Magazine*, 6(1):35–48, 1999.

[13] M. Golfarelli, D. Maio, and D. Maltoni. On the error-reject trade-off in biometric verification systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):786–796, 1997.

[14] H. Proenca and L. A. Alexandre. UBIRIS: A noisy iris image database. In *13th International Conference on Image Analysis and Processing - ICIAP 2005*, volume LNCS 3617, pages 970–977. Springer, 2005.

[15] CASIA Database. http://www.cbsr.ia.ac.cn/english/IrisDatabase.asp.

[16] L. Ma, T. Tan, Y. Wang, D. Zhang, and C. Boyce. Critique: Efficient iris recognition by characterizing key local variations. *IEEE Transactions on Image Processing*, 13(6), 2008.

[17] R. G. Stockwell, L. Mansinha, and R. P. Lowe. Localization of the complex spectrum: the s transform. *Signal Processing, IEEE Transactions on*, 44(4):998–1001, 1996.

[18] J. Daugman. Biometric personal identification system based on iris analysis. U.S. Patent No. 5,291,560, 1994.

[19] A.L. Yuille, D.S. Cohen, and P.W. Hallinan. Feature extraction from faces using deformable templates. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 104–109, 1989.

[20] J. Daugman. How iris recognition works. In *International Conference on Image Processing 2002*, volume 1, pages 33–36, 2002.

[21] R.P. Wildes. Iris recognition: an emerging biometric technology. *Proceedings of the IEEE*, 85(9):1348–1363, 1997.

[22] Y. Huang, S. Luo, and E. Chen. An efficient iris recognition system. In *International Conference on Machine Learning and Cybernetics*, volume 1, pages 450–454, 2002.

[23] Y. Liu, S. Yuan, X. Zhu, and Q. Cui. A practical iris acquisition system and a fast edges locating algorithm in iris recognition. In *20th IEEE Conference on Instrumentation and Measurement Technology*, volume 1, pages 166–168, 2003.

[24] H. Sung, J. Lim, J. Park, and Y. Lee. Iris recognition using collarette boundary localization. In *17th International Conference on Pattern Recognition*, volume 4, pages 857–860, 2004.

[25] Q. Tian, Q. Pan, Y. Cheng, and Q. Gao. Fast algorithm and application of hough transform in iris segmentation. In *International Conference on Machine Learning and Cybernetics*, volume 7, pages 3977–3980, 2004.

[26] G. Xu, Z.F. Zhang, and Y.D. Ma. Automatic iris segmentation based on local areas. In *International Conference on Pattern Recognition*, volume 4, pages 505–508. IEEE Computer Society, 2006.

[27] A. Zaim, M. Quweider, J. Scargle, J. Iglesias, and R. Tang. A robust and accurate segmentation of iris images using optimal partitioning. In *18th International Conference on Pattern Recognition*, volume 4, pages 578–581, 2006.

[28] T.A. Camus and R. Wildes. Reliable and fast eye finding in close-up images. In *16th International Conference on Pattern Recognition*, volume 1, pages 389–394, 2002.

[29] B. Bonney, R. Ives, D. Etter, and D. Yingzi. Iris pattern extraction using bit planes and standard deviations. In *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, volume 1, 2004.

[30] H. Proenca and L.A. Alexandre. Iris segmentation methodology for non-cooperative recognition. *IEE Proceedings on Vision, Image and Signal Processing*, 153(2):199–205, 2006.

[31] M. Tucceryan. Moment-based texture segmentation. *Pattern Recognition Letters*, 15(7):659–668, 1994.

[32] S.J. Pundlik, D.L. Woodard, and S.T. Birchfield. Non-ideal iris segmentation using graph cuts. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 1–6, 2008.

[33] X. Liu, K.W. Bowyer, and P.J. Flynn. Experiments with an improved iris segmentation algorithm. In *Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, pages 118–123, 2005.

[34] Z.A. Sun, T.N. Tan, and Y.H. Wang. Robust encoding of local ordinal measures: A general framework of iris recognition. In *ECCV Workshop on Biometric Authentication*, pages 270–282, 2004.

[35] L. Ma, T. Tan, Y. Wang, and D. Zhang. Efficient iris recognition by characterizing key local variations. *IEEE Transactions on Image Processing*, 13(6):739–750, 2004.

[36] P. Yao, J. Li, X. Ye, Z. Zhuang, and B. Li. Iris recognition algorithm using modified log-gabor filters. In *Proceedings of the 18th International Conference on Pattern Recognition*, pages 461–464, 2006.

[37] D. M. Monro, S. Rakshit, and D. Zhang. DCT-based Iris Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4):586–595, 2007.

[38] Christel Loic, L. Torres, and M. Robert. Person identification technique using human iris recognition. In *Vision Interface*, pages 294–299, 2002.

[39] R.C. Gonzalez and R.E. Woods. *Digital Image Processing (3^{rd} Edition)*. Prentice Hall, 2007.

[40] L.V. Birgale and M. Kokare. Iris recognition using discrete wavelet transform. In *International Conference on Digital Image Processing*, pages 147–151, 2009.

[41] L. Masek and P. Kovesi. Matlab Source Code for a Biometric Identification System based on Iris Patterns. The School of Computer Science and Software Engineering, The University of Western Australia., 2003.

[42] SIFT for Matlab. http://www.vlfeat.org/~vedaldi/code/sift.html.

[43] M. Brown and D. Lowe. Invariant features from interest point groups. In *In British Machine Vision Conference*, pages 656–665, 2002.

[44] C. Valgren and A. J. Lilienthal. SIFT, SURF and Seasons: Long-term outdoor localization using local features. In *Proceedings of the European Conference on Mobile Robots (ECMR)*, pages 253–258, 2007.

[45] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 511–518, 2001.

[46] P. Simard Eon, P. Y. Simard, P. Haffner, and Y. Lecun. Boxlets: a fast convolution algorithm for signal processing and neural networks. In *Advances in Neural Information Processing Systems*, pages 571–577. MIT Press, 1999.

[47] Y. Fang, J. Cheng, K. Wang, and H. Lu. Hand gesture recognition using fast multi-scale analysis. In *Fourth International Conference on Image and Graphics*, pages 694–698, 2007.