ADVANCING THE TECHNOLOGY OF SCLERA RECOGNITION

Sinan Husam Mahdi Alkassar

A Thesis Submitted for the Degree of
Doctor of Philosophy at Newcastle University

School of Electrical and Electronic Engineering
Faculty of Science, Agriculture and Engineering
January 2018

DECLARATION

I hereby declare that this thesis is my own work and effort and that it has not been submitted anywhere for any award. Where other sources of information have been used, they have been acknowledged.

Newcastle upon Tyne, January 2018

Sinan Alkassar

CERTIFICATE OF APPROVAL

I confirm that, to the best of my knowledge, this thesis is from the student's own work and effort, and all other sources of information used have been acknowledged. This thesis has been submitted with my approval.

Professor Jonathon A. Chambers

To my greatest supporters that is my wonderful parents, my beloved wife,
Shahad, and my lovely sons, Mohanad and Adam.
— Sinan

ACKNOWLEDGEMENTS

I would like to express my genuine appreciation and gratitude to my supervisors Prof. Jonathon A. Chambers, Prof. Satnam Dlay, and Dr Lok Woo for their help, support, and guidance through my PhD studies. They have been and always will be a source of motivation and inspiration and my role model as a researcher.

I am also grateful to my family, my parents, brothers, sister, and all my friends in Iraq and UK for their support and encouragement to finish my studies successfully.

In addition, I am also thankful to all my colleagues and staff within the School of Electrical and Electronic Engineering at Newcastle University for their help and support to make my times here wonderful both socially and educationally.

I also would like to show my gratitude to Prof. Arun Ross for his advice and support regarding my PhD project when we met at the International Conference on Biometrics (ICB) 2016.

Finally, I would like to thank the Ministry of Higher Education and Scientific Research in Iraq and the Iraqi Cultural Attaché in London for providing the opportunity to complete my PhD degree and their continuous support and encouragement.

To all I say, thank you very much.

ABSTRACT

Emerging biometric traits have been suggested recently to overcome some challenges and issues related to utilising traditional human biometric traits such as the face, iris, and fingerprint. In particular, iris recognition has achieved high accuracy rates under Near-InfraRed (NIR) spectrum and it is employed in many applications for security and identification purposes. However, as modern imaging devices operate in the visible spectrum capturing colour images, iris recognition has faced challenges when applied to coloured images especially with eye images which have a dark pigmentation. Other issues with iris recognition under NIR spectrum are the constraints on the capturing process resulting in failure-to-enrol, and degradation in system accuracy and performance. As a result, the research community investigated using other traits to support the iris biometric in the visible spectrum such as the sclera.

The sclera which is commonly known as the white part of the eye includes a complex network of blood vessels and veins surrounding the eye. The vascular pattern within the sclera has different formations and layers providing powerful features for human identification. In addition, these blood vessels can be acquired in the visible spectrum and thus can be applied using ubiquitous camera-based devices. As a consequence, recent research has focused on developing sclera recognition. However, sclera recognition as any biometric system has issues and challenges which need to be addressed. These issues are mainly related to sclera segmentation, blood vessel enhancement, feature extraction, template registration, matching and decision methods. In addition, employing the sclera biometric in the wild where relaxed imaging constraints are utilised has introduced more challenges such as illumination variation, specular reflections, non-cooperative user capturing, sclera blocked region due to glasses and eyelashes, variation in capturing distance, multiple gaze directions, and eye rotation.

The aim of this thesis is to address such sclera biometric challenges and highlight the potential of this trait. This also might inspire further research on tackling sclera recognition system issues. To overcome the above-mentioned issues and challenges, three major contributions are made which can be summarised as 1) designing an efficient sclera recognition system under constrained imaging conditions which include new sclera segmentation, blood vessel enhancement, vascular binary network mapping and feature extraction, and template registration techniques; 2) introducing a novel sclera recognition system under relaxed imaging constraints which exploits novel sclera segmentation, sclera template rotation alignment and distance scaling methods, and complex sclera features; 3) presenting solutions to tackle issues related to applying sclera recognition in a real-time application such as eye localisation, eye corner and gaze detection, together with a novel image quality metric.

The evaluation of the proposed contributions is achieved using five databases having different properties representing various challenges and issues. These databases are the UBIRIS.v1, UBIRIS.v2, UTIRIS, MICHE, and an in-house database. The results in terms of segmentation accuracy, Equal Error Rate (EER), and processing time show significant improvement in the proposed systems compared to state-of-the-art methods.

PUBLICATIONS

Journal publications:

1. **S. Alkassar**, W. L. Woo, S. S. Dlay, and J. A. Chambers, "Robust sclera recognition system with novel sclera segmentation and validation techniques," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 3, pp. 474-486, 2017.
2. **S. Alkassar**, W. L. Woo, S. S. Dlay, and J. A. Chambers, "Sclera recognition: on the quality measure and segmentation of degraded images captured under relaxed imaging conditions," *IET Biometrics*, vol. 6, no. 4, pp. 266-275, 2017.

Conference publications:

1. **S. Alkassar**, W. L. Woo, S. S. Dlay, and J. A. Chambers, "A novel method for sclera recognition with images captured on-the-move and at-a-distance," in *4th Int. Workshop on Biometrics and Forensics (IWBF)*, 2016, pp. 1-6.
2. **S. Alkassar**, W. L. Woo, S. S. Dlay, and J. A. Chambers, "Enhanced segmentation and complex-sclera features for human recognition with unconstrained visible-wavelength imaging," in *4th Int. Conf. on Biometrics (ICB)*, 2016, pp. 1-8.
3. **S. Alkassar**, W. L. Woo, S. S. Dlay, and J. A. Chambers, "Efficient Eye Corner and Gaze Detection for Sclera Recognition Under Relaxed Imaging Constraints," in *24th European Signal Processing Conf. (EUSIPCO)*, 2016, pp. 1-5.

CONTENTS

I Thesis Chapters 1

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS

ACRONYMS

AHE	Adaptive Histogram Equalisation
CDF 9/7	Cohen-Daubeches-Feauveau 9/7 wavelet
CLAHE	Contrast-Limited Adaptive Histogram Equalisation
D-LDP	Dense-Local Directional Pattern
D-SIFT	Dense-Scale Invariant Feature Transform
DCT	Discrete Cosine Transform
DDCDWT	Double-Density Complex Discrete Wavelet Transform
DET	Detection Error Tradeoff
DoG	Difference of Gaussian
dpi	dots per inch
DT-CWT	Dual-Tree Complex Wavelet Transform
DWT	Discrete Wavelet Transform
EER	Equal Error Rate

EGBU	Excellent, Good, Bad, and Ugly
F-G	Frontal-Gaze
FAR	False Accept Rate
FAST	Features from Accelerated Segment Test
FBI	Federal Bureau of Investigation
fm	focus measure
FMR	False Match Rate
FN	False Negative
FNMR	False Non-Match Rate
FP	False Positive
FRR	False Reject Rate
GAR	Genuine Accept Rate
GBU	Good, Bad, and Ugly
GLCM	Gray Level Co-occurrence Matrix
GWT	Gabor Wavelet Transform
HS	Harris-Stephan
IPs	Interest-Points
IUWT	Isotropic Undecimated Wavelet Transform
KLDA	Kernel Linear Discriminant Analysis
KPCA	Kernel Principal Components Analysis
L-G	Left-Gaze
LDA	Linear Discriminant Analysis
MLBP	Multiresolution Local Binary Pattern
NIR	Near-InfraRed

OLBP	Oriented Local Binary Pattern
PCA	Principal Component Analysis
PIN	Personal Identification Number
R-G	Right-Gaze
RANSAC	Random Sample Consensus
ROC	Receiver Operating Characteristic
ROI	Region Of Interest
SURF	Speed Up Robust Features
SVMs	Support Vector Machines
STRADS	Sclera Template Rotation Alignment and Distance Scaling
TN	True Negative
TP	True Positive

Part I

Thesis Chapters

1

INTRODUCTION

1.1 MOTIVATION

The significant developments in the state-of-the-art communication technology throughout the last decade have led to an enormous growth in the number of mobile and Internet users. It has been reported in 2014 via digital analysts at GSMA Intelligence [1] which is a website which provides definitive data and analysis for the mobile industry, that for the first time, the number of mobile devices has crossed the number of humans reaching 7.22 billion devices. These devices like tablets and smartphones are increasing five times faster than human beings in the world. In addition, the Statista website [2] has reported that in 2016, an estimated 62.9 % of the world population already own a mobile device and the number of mobile phone users will reach over 5 billion in 2019 as shown in Fig. 1.1. Furthermore, developing countries have witnessed a rapid increase in the number of mobile and Internet users such as in India, where 142 million mobile service contracts have been registered in 2011 [3] and the number is increasing rapidly. In addition, the number of mobile phone users in India has grown dramatically from 524.9 million in 2013 to 684.1 million in 2016 and estimated to reach 813.3 million by 2019 [4].

As a consequence, there has been an increased number of cybercrimes in recent years. This is due to the development of new communication infrastructures and their affordable prices. This was supported by the BBC News article [5] which reported the annual figures of cybercrimes and fraud scale in a survey conducted in England and Wales. Such crimes include bank and credit account fraud, unauthorised access to personal information, and hacking. The report suggests changing the policing strategies and with cooperation from industry, it is mandatory to develop new methods for fighting cybercrimes. Moreover, Bill Laberis [6] has summarised 20 cybercrime statistics providing clear evidence that nothing is bulletproof when it is related to com-

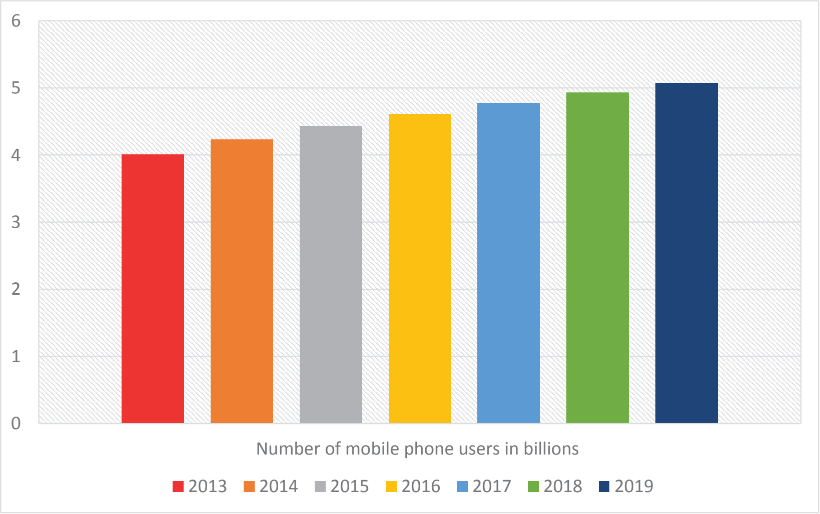
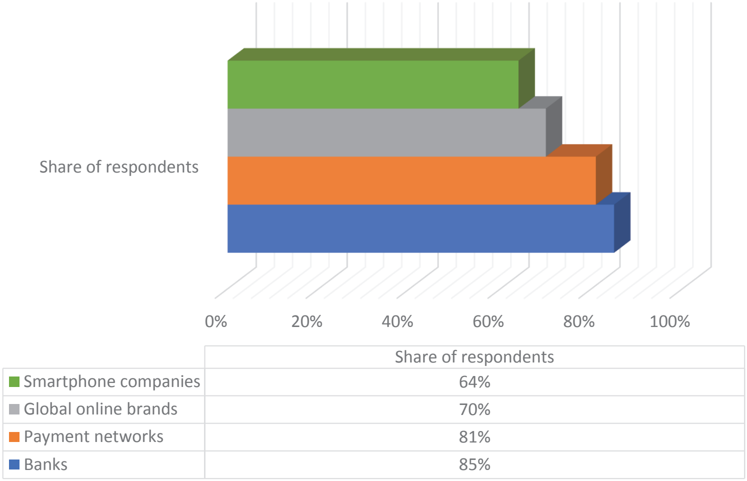


Figure 1.1: Estimated numbers of mobile phone users in the worldwide from 2013 to 2019 [2].

promising data. Examples of the reported statistics are that the global cost of these crimes will reach $2 trillion by 2019, more than 29 million attempts of identity theft were exposed across financial, government, education, and health care sectors and the global spending to tackle cybercrime issues exceeded $80 billion in 2016. Therefore, developing new methods to overhaul cyber security methods and adopting non-traditional methods to stop cyber offenders are essential.

The research community has therefore proposed adopting biometric traits as one of the efficient solutions to tackle the above-mentioned problems. Moreover, a survey on consumer trust in companies and institutions to store their biometric data for identification purpose to process payments conducted in Great Britain in 2016 [7] is shown in Fig. 1.2. Approximately 85% of consumers trust in banks to store their biometric information providing evidence that the markets, as well as consumers, depend on using automatic authentication and recognition systems based on biometric traits. As a result, much pressure has been placed on the research community to develop biometric systems even further. For instance, automatic fingerprint recognition has been utilised in the past four decades to prevent financial and identity fraud crimes [8] and yet, it is still a developing research area.

On the other hand, automatic iris recognition has been suggested in 1986 by Flom and Safir [9, 10] where they registered a patent but



	Share of respondents
🟩 Smartphone companies	64%
⬜ Global online brands	70%
🟧 Payment networks	81%
🟦 Banks	85%

Figure 1.2: A survey shows the amount of trust that consumers have in institutes and companies to store their biometric information [7].

without designing an actual algorithm for their patent. After 8 years of introducing the automated iris recognition idea, John Daugman designed a new practical system based on the integro-differential operator as a method for segmentation, Gabor filter for feature extraction, and Hamming distance as a score generator [11]. Amongst all biometric modalities, the iris code can be compared efficiently resulting in less complexity for the matching process. Therefore, it has been used widely in airports, border control, identity check, ATM, and recent mobile devices [12].

Commonly, imaging an iris in current commercial iris biometric systems is performed in the Near-InfraRed (NIR) spectrum as the disparity between the pupil and the iris is maximised. Recent research however has examined applying iris recognition in the visible spectrum [10]. This is due to the fact that the traditional NIR iris sensors have a restricted capturing distance and constrained imaging. Proenca [13] suggested that using visible wavelength illumination for iris recognition is more suitable when the capturing is on-the-move and at-a-distance. In addition, the Noisy Iris Challenge Evaluation (NICE) competitions part I and II [14, 15] were performed for eye images under visible wavelength illumination and less constrained conditions such as minor user cooperation and dynamic imaging domains and the results were promising. However, as the system design aims to employ the user non-cooperative scenario such as using his/her mobile camera for iris recognition, many issues have arisen [16].

Proenca [10] has reported four foremost factors in addition to the typical noise factors that affect iris biometric quality in the visible wavelength which can be summarised as 1) illumination; 2) blur; 3) perspective; and 4) occlusions. These four main factors may include occlusion by eyelashes, low resolution, illumination variation, reflections from glasses, and pose. This has led to asking what if the iris biometric fails under visible wavelength illumination. As a result, the challenges facing iris recognition under visible wavelength illumination have inspired many researchers to look for alternative or supplementary types of traits in the eye such as the sclera and the periocular regions [17, 18].

This thesis opens new research directions in the development of novel and promising sclera biometric systems. This study, by introducing multiple systems for different imaging conditions, is likely to inspire other research teams to further investigate the encouraging sclera biometric technology. The technology in this thesis involves designing two sclera recognition systems under constrained and relaxed imaging conditions respectively. These systems have remarkable segmentation, blood vessel enhancement, feature extraction, and template registration methods compared to state-of-the-art sclera biometric techniques. In addition, several issues related to the use of the sclera biometric in a real-time scenario have been investigated and efficient solutions have been proposed. The ultimate goal of this thesis is to introduce the sclera biometric to the mobile security industry by evaluating sclera images captured via different mobile cameras showing both strong and weak aspects of the design.

1.2 STATEMENT OF ORIGINALITY

The major contributions of this thesis can be summarised as follows:

- In Chapter 2, a comprehensive view is provided on the merits and drawbacks of vascular pattern recognition within human organs such as the retina, hand, finger and sclera. In addition, a survey on recent sclera recognition systems is presented which might be a ground truth for future studies on sclera recognition systems inspired to tackle issues and address their challenges. Further more, eye datasets under visible wavelength illumination which are utilised in this thesis are listed and compared along with their specifications;

- In Chapter 3, the performance of sclera recognition under constrained imaging conditions is studied. This has led to quantification of some challenges and introducing a robust method for sclera recognition suitable for eye images captured with minimised noise factors [19]. The proposed method includes efficient eye occlusion for sclera validation, adaptive sclera shape segmentation, enhanced vessel map extraction, and robust sclera feature extraction with new template registration and matching techniques;

- In Chapter 4, sclera recognition under relaxed imaging constraints is discussed where several noise factors are added. For instance, images captured at-a-distance and on-the-move with natural illumination conditions acquired by the user. Moreover, challenges related to sclera segmentation and blood vessel enhancement are reviewed providing an insight for further studies. In addition, novel fusion method for sclera segmentation which is adapted to noise factors is designed, efficient geometrical sclera template rotation alignment method which is invariant to eye rotation is proposed, robust sclera scaling method to minimise the effect of capturing a sclera image on-the-move and at-a-distance is introduced, and robust blood vessel enhancement and feature extraction methods based on what are denoted as *complex-sclera* features via the Dual-Tree Complex Wavelet Transform (DT-CWT) are proposed [20, 21].

- In Chapter 5, practical eye localisation method which enhances the AdaBoost eye detection is suggested and in particular, for real-time applications. In addition, novel fusion method for eye corner detection is proposed which is adapted to noise factors. Furthermore, effective eye gaze detection method which is invariant to eye rotation is introduced based on the detected eye corner points along with sclera image quality measure and classification techniques which are proposed based on sclera image focus measure [22, 23].

- Operation with sclera images captured at-a-distance, on-the-move, and under natural illumination conditions using ubiquitous camera-based devices such as a DSLR camera, mobile camera, or a webcam is feasible. The thesis provides a comprehensive evaluation for sclera recognition and paves the way for future research to tackle issues reported in this thesis.

1.3 THESIS ORGANISATION

This thesis is composed of six chapters, where two main challenges linked to sclera recognition in controlled and unconstrained imaging conditions will be discussed and tackled in Chapters 3 and 4. While in Chapter 5, some issues related to sclera recognition in real-time scenarios will be covered.

Chapter 1 "Introduction" presents the motivations, goals, and structure of this thesis.

Chapter 2 "Background and Literature Review" provides comprehensive background knowledge on biometric systems, in particular, blood vessel-based biometric systems and reviews the literature on sclera recognition in detail and presents datasets relevant to this thesis.

Chapter 3 "Robust Sclera Recognition System Under Constrained Imaging Conditions" presents a new sclera recognition system which is suitable for an indoor environment application and tackles some problems related to sclera segmentation, vessel enhancement, and template registration.

Chapter 4 "Sclera Recognition Under Relaxed Imaging Conditions: Key issues, Challenges, and Resolutions" focuses on sclera recognition when image conditions are challenging and the capturing routine is achieved in a relaxed fashion. It proposes new solutions for segmentation, alignment, scaling, and feature extraction.

Chapter 5 "Overcoming Challenges in Sclera Recognition Systems Under Various Imaging Conditions for Real-Time Applications" discusses some issues associated with the real-time implementation of sclera recognition such as eye localisation, image quality measure, and eye gaze and corner detection and proposes practical solutions to tackle these issues.

Chapter 6 "Conclusions and Future Work" summarises the advancement achieved in this study regarding sclera recognition and shows the advantages and challenges sclera recognition stands. Moreover, this chapter lists potential future work which can further develop sclera recognition.

2

BACKGROUND AND LITERATURE REVIEW

2.1 INTRODUCTION

Biometric systems play a vital role as a powerful platform for modern identification system designs. Such systems have many applications and are used in forensics, crime scene investigation, together with secure logging into a building, mobile device, ATM, or bank account. The ability to distinguish an individual through human characteristics highlights the benefit of using biometric traits as personal identifiers instead of a traditional password, Personal Identification Number (PIN), or smart card [24, 25]. Extensive research on biometric systems has shown the importance of using a biometric trait as well as revealing many challenges and issues. These challenges are particularly related to designing a solid and user-friendly system that is secure and can be acceptable to people who will use it on a daily basis. Many factors must also be considered when designing a biometric system including universality, distinctiveness, permanence, collectability, acceptability, and performance [26].

On the other hand, blood vein based biometric systems have attracted a lot of interest recently as a new biometric trait to overcome some problems in typical biometric systems such as fingerprints, the face, and iris which are the dominant traits in the market. For instance, using a human face for recognition can be easily spoofed because of its accessibility compared with other biometric traits given the ability of hackers to imitate face features easily [27]. Therefore, using hidden attributes of these blood vessels for recognition, which are relativity difficult to acquire, can provide a solution to minimise hackers' attacks. Extracting the blood vessel network from a human organ and using its discriminant features to identify or verify a human has achieved promising results in terms of recognition accuracy and error rates. Examples of blood vein based biometric systems are those based on the retina, finger vein, palm vein, and sclera blood vessels [28, 29, 30, 17, 31].

Consequently, sclera recognition has been proposed recently as a method to tackle issues in iris recognition under visible wavelength spectrum. Iris recognition mainly performs better in the Near-InfraRed (NIR) spectrum even for darkly pigmented irises revealing a rich feature texture compared to iris recognition under visible wavelength [32, 33]. As a result, and because most of the recent devices having a normal camera which captures RGB images, the need to overcome these issues has led to propose the sclera as a supporting biometric trait to ocular recognition.

This chapter will first review biometrics in terms of definition, characteristics, traits, recognition rates, design challenges, and application. The quantitative analysis used to evaluate proposed sclera recognition systems will then be presented. Then, an overview of recent blood vessel biometric systems will be introduced. Next, sclera recognition will be defined and recent work on system design will be discussed. Finally, datasets employed in this research will be discussed in terms of the acquisition device, capturing scene, image quality, and the challenges these datasets present for sclera recognition.

2.2 BACKGROUND

2.2.1 *Biometrics Overview*

Biometrics can be defined as the science of identifying an individual based on behavioural, physical, or chemical properties belonging to that individual [26]. The importance of this field of science is the large scale need for identity management systems which can serve in many applications. For instance, some high-risk applications such as logging into high-security infrastructure buildings, managing money transaction in banks, and controlling military devices need an accurate powerful identity system which does not rely only on a password or PIN which can easily be spoofed. Logging into a mobile device or a PC also requires a robust recognition system to avoid compromising an individual's privacy. Providing a system which prevents imposters from spoofing any identity using a biometric trait which is better than the traditional methods such as token-based or knowledge-based mechanisms [26] is therefore important.

Examples of human biometric traits used currently in the market are shown in Fig. 2.1. Some biometric traits are physical such as the iris, face, DNA, fingerprint, finger knuckles, and hand vein [32, 34, 35, 8, 36, 37, 38] while some are behavioural traits like a keystroke, signature, or individual's gait [39, 40, 41, 42]. Whereas for the voice, it can be considered as a combination of physical and behavioural characteristics [26, 43].

Typically, a biometric system consists of acquiring the data from any given biometric trait, extracting a distinctive feature set, creating an enrolled dataset, and applying a matching process to identify a certain individual. Therefore, any biometric system in general consists of four main modules which can be defined as: 1) sensor module; 2) quality assessment and feature extraction module; 3) matching and decision-making module; and 4) system database module [26].

The sensor module includes the device that reads or captures biometric data from a person. Some traits require a special sensor to collect the data such as the fingerprint sensor whereas other traits require an imaging device to capture gait, the iris, or face with NIR or visible wavelength illumination. The quality assessment measure and feature extraction can be summarised as examining the captured biometric data in terms of quality and performance in order to improve biometric data quality. This involves for instance checking image conditions for traits that requires image capturing or checking the quality of feature extracted from these traits. If the collected biometric data passed the quality test, the system extracts the features and creates a user template. Next, the matching and decision module is responsible for comparing the test template against one or all the stored templates in the database to generate a matching score and thus a decision. While for the database module, the database can be represented as a repository where all the enrolled users have their biometric data stored. Then, the stored templates are generally subjected to an encryption process to preserve the privacy of users and prevent imposters hacking any identity.

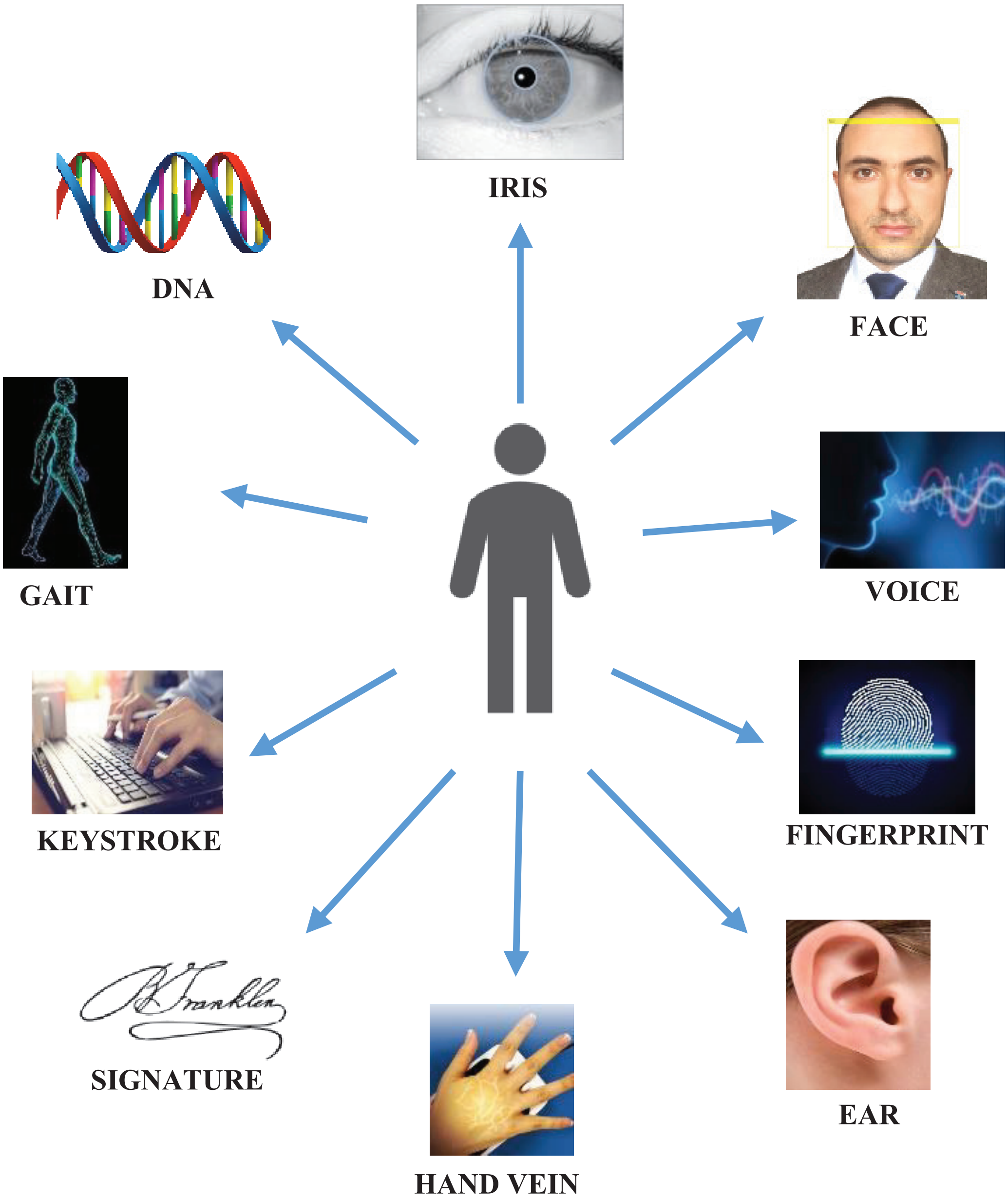


Figure 2.1: Examples of human biometric traits utilised in recent identification management systems.

2.2.2 *Characteristics of Biometric Traits*

As there are many biometric traits available when designing a security or an identity management systems, there are therefore factors and characteristics that should be considered to choose a suitable biometric trait for a particular purpose. These characteristics can be divided into two categories: standard quality measures and practical quality requirements [44, 26]. The standard quality measures for all biometric traits can be defined as

i. Universality: the biometric trait must be possessed by all users.

ii. Distinctiveness: for any two individuals, the biometric trait must have distinctive features which are sufficient to distinguish between them.

iii. Permanence: the characteristics of the biometric trait should be time-invariant.

iv. Collectability: it is possible to collect and measure the characteristics of the biometric trait using suitable devices and quantitative methods respectively.

While for practical quality requirements, particularly in a real-time application, there are some issues that affect the biometric trait selection process during the design which can be summarised as

i. Performance: how the chosen biometric trait performs in terms of recognition accuracy, error rates, and processing time.

ii. Acceptability: the acceptability of a particular trait to the regular user of an identity management system.

iii. Circumvention: the level of difficulty for the system to be spoofed by an imposter.

A comparison conducted by Jain et al. [44] based on the aforementioned requirements is shown in Table 2.1. It is noteworthy that this comparison is achieved based on authors' perception. It is evident that there is no single biometric trait that can outperform other traits in terms of operational environments. Each biometric has its advantages and disadvantages and hence the decision to select which biometric

Table 2.1: Comparison of many biometric traits in terms of universality, distinctiveness, permanence, collectability, performance, acceptability, and circumvention [44].

Biometric trait	**Universality**	**Distinctiveness**	**Permanence**	**Collectability**	**Performance**	**Acceptability**	**Circumvention**
DNA	H	H	H	L	H	L	L
Ear	M	M	H	M	M	H	M
Face	H	L	M	H	L	H	H
Facial thermogram	H	H	L	H	M	H	L
Fingerprint	M	H	H	M	H	M	M
Gait	M	L	L	H	L	H	M
Hand geometry	M	M	M	H	M	M	M
Hand vein	M	M	M	M	M	M	L
Iris	H	H	H	M	H	L	L
Keystroke	L	L	L	M	L	M	M
Odor	H	H	H	L	L	M	L
Palmprint	M	H	H	M	H	M	M
Retina	H	H	M	L	H	L	L
Signature	L	L	L	H	L	H	H
Voice	M	L	L	M	L	H	H

H: High, M: Medium, and L: Low.

trait is suitable is a design problem. For instance, DNA has a degree of distinctiveness which can be used in forensics while it is currently hard to employ DNA for logging on a mobile device where a fingerprint would be more efficient.

2.2.3 *Biometric System Architecture*

The biometric system basically is a pattern recognition system where features extracted from physical, behavioural, or chemical traits are utilised to recognise a person [45]. All biometric systems involve three stages, regardless of the technology used, which can be defined as the enrolment, verification, and identification stages. These stages are depicted in Fig. 2.2.

In the enrolment stage, the biometric data initially are collected from the user using a certain enrolment biometric reader such as a fingerprint sensor, face 3D camera, or iris sensor. Next, a mechanism for

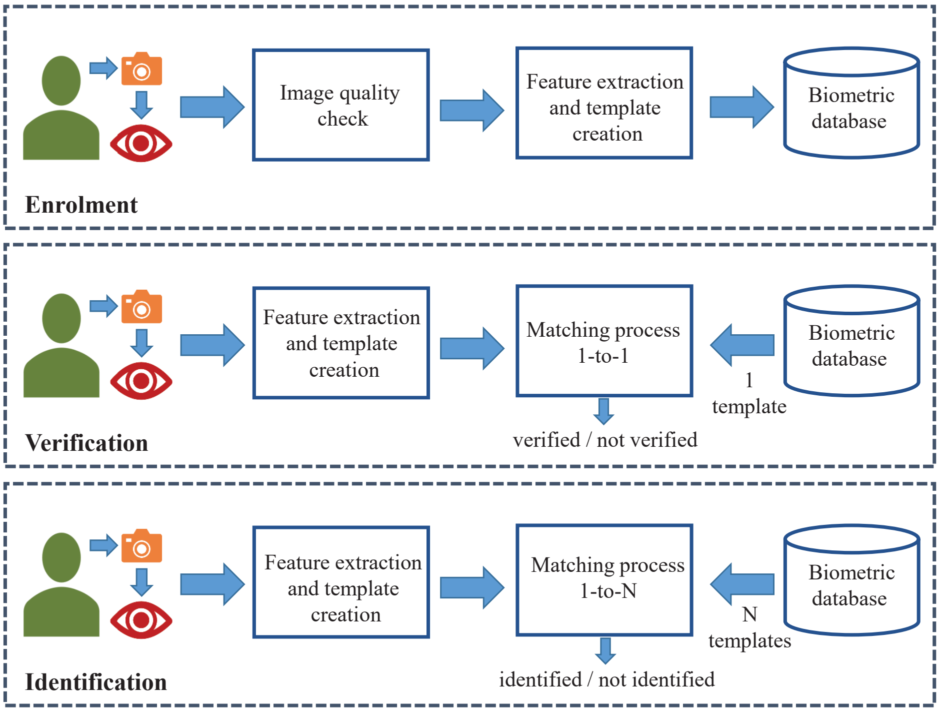


Figure 2.2: General biometric system architecture operating in the enrolment, verification, and identification stages.

checking the collected data quality to make sure the successive stages can reliably process the acquired biometric data is designed. Then, a feature extraction process is applied to extract the most distinctive features and generate a compact but expressive user template. All the enrolled templates are stored in a central storage called the biometric database.

There are two types of biometric enrolment which can be defined as positive and negative enrolment [42]. Positive enrolment is used for verification and positive identification is where a biometric database of eligible users is constructed. While for negative enrolment, which is basically utilised in the negative identification, the collected biometric data could be collected in a voluntary or covert fashion.

In the verification stage, the main purpose of the system is to confirm the identity that a person claims to be. A user submits biometric data through a biometric reader and the same mechanism for feature extraction that is used in the enrolment stage is utilised to create the user biometric template. Then, the system retrieves the claimed identity template from the database and a matching process is initiated and a result of match or non-match is returned [44]. Therefore, the verification process is a one-to-one process as shown in Fig. 2.2. The verification is normally used for positive recognition where the goal is to prevent multiple users from using the same identity [46].

The verification problem can be formally expressed as: for claimed identity I, a certain user submitted a feature vector $\mathbf{u}_Q$, the system will determine if the enrolled feature vector $\mathbf{u}_I$, corresponding to I, and $\mathbf{u}_Q$ belong to class c_1 or c_2. Class c_1 indicates a match for the claimed user (genuine user) whereas class c_2 indicates no-match (an imposter). This process can be defined as

$$\left(\mathbf{u}_I, \mathbf{u}_Q\right) \in \begin{cases} c_1 & \text{if } S\left(\mathbf{u}_I, \mathbf{u}_Q\right) \geqslant th \\ c_2 & \text{otherwise} \end{cases}, \tag{2.1}$$

where S is the function which calculates the similarity degree between $\mathbf{u}_I$ and $\mathbf{u}_Q$, and th is the threshold which determines the sensitivity of the system.

On the other hand, the aim of the identification stage is to recognise an identity of a user, where the system is user agnostic, by searching the entire database to find a match [44]. The system collects user biometric data and initiates a comparison on a one-to-N basis where N refers to the total number of enrolled templates. Therefore, the identification process is complex and requires higher processing time compared to the verification stage as N can be dozens, hundreds, thousands, or even millions of biometric records such as in an iris-based system. The identification stage is normally used for negative recognition where the aim is to prevent one person from using multiple identities [46] and it is widely used to identify criminals or dead bodies.

Similarly, the identification can be formally expressed as: for N identities, given an input subject feature vector $\mathbf{u}_Q$, the system will determine the identity I_k of that user where $k \in \{1, 2, \cdots, N, N+1\}$. If I_k is determined within the enrolled identities N, then the subject is identified. Otherwise, I_k will be equal to $N+1$ which refers to no identity can be determined for that subject. This can be expressed as

$$\mathbf{u}_Q \in \begin{cases} I_k & \text{if } \max_k \left\{S\left(\mathbf{u}_I, \mathbf{u}_Q\right)\right\} \geqslant th, k = 1, 2, \cdots, N \\ I_{N+1} & \text{otherwise} \end{cases}. \tag{2.2}$$

2.2.3.1 *Biometric Quantitative Analysis Protocols*

As any biometric system mainly depends on biometric readers to collect data from individuals, it is therefore expected that there will be some noise in that data. This is due to illumination variation, biometric reader malfunction, inconsistent user interaction, and changing biometric characteristics. Thus, it is seldom possible to collect two totally identical templates from the same user. Therefore, if a comparison between query and enrolled templates produces a perfect match, then it is reasonable to suspect that the biometric system is under attack [44]. The variance between two feature sets of the same individual is referred to as intra-class variation whereas the variance between two feature sets originating from different individuals is referred to as inter-class variation. The smaller (non-zero) intra-class variation and larger inter-class variation for a feature set, the better.

Measuring the similarity between two biometric feature sets will determine two scores denoted as genuine score or imposter score depending on the degree of similarity [44, 26]. The genuine score is generated when two feature sets belonging to the same subject are compared whereas the imposter score is generated when two feature sets belonging to different subjects are compared. The similarity measure depends mainly on a threshold as shown in Fig. 2.3 where four distributions are generated which can be summarised as: for the enrolled feature vector $\mathbf{u}_I$ belonging to identity I, a matching process is initiated with a query feature vector $\mathbf{u}_Q$ resulting in two hypotheses defined as

- H_1: $\mathbf{u}_Q$ does not belong to the same identity I as $\mathbf{u}_I$;
- H_2: $\mathbf{u}_Q$ does belong to the same identity I as $\mathbf{u}_I$;

Based on these two hypotheses, two associated decisions are defined as

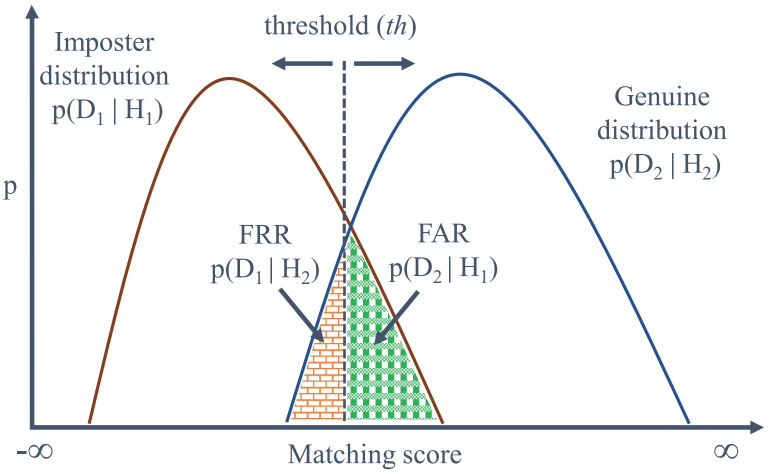


Figure 2.3: The distribution of similarity scores generated from genuine and imposter matching scores.

- D_1: the query subject identity claim is rejected;
- D_2: the query subject identity claim is accepted;

As a consequence, four distributions are generated which can be defined as

1. The genuine distribution defined as $p\left(D_2 \mid H_2\right)$ generating the True Positive (TP) scores;
2. The imposter distribution defined as $p\left(D_1 \mid H_1\right)$ generating the True Negative (TN) scores;
3. The false match distribution defined as $p\left(D_2 \mid H_1\right)$ generating the False Positive (FP) scores;
4. The false non-match distribution defined as $p\left(D_1 \mid H_2\right)$ generating the False Negative (FN) scores.

The last two probabilities represent error rates in the aforementioned hypothesises defined as the False Match Rate (FMR), which is also denoted as False Accept Rate (FAR), and the False Non-Match Rate (FNMR), which is also referred to as False Reject Rate (FRR). These two rates are key metrics for evaluating biometric system performance. FAR is defined as the part of imposter matching scores exceeding the threshold th while FRR is the part of genuine matching score falling below th [47]. The Genuine Accept Rate (GAR) is defined as the fraction of genuine matching scores which exceeds th. FAR, FRR, and GAR are defined as

$$\mathrm{FAR} = \int_{\mathrm{th}}^{\infty} p\left(S\left(\mathbf{u}_I, \mathbf{u}_Q\right), \mid H_1\right) dS = \frac{\mathrm{FP}}{\mathrm{FP}+\mathrm{TN}} \times 100\%, \tag{2.3}$$

$$\mathrm{FRR} = \int_{-\infty}^{\mathrm{th}} p\left(S\left(\mathbf{u}_I, \mathbf{u}_Q\right), \mid H_2\right) dS = \frac{\mathrm{FN}}{\mathrm{FN}+\mathrm{TP}} \times 100\%, \tag{2.4}$$

$$\mathrm{GAR} = 1 - \mathrm{FRR}. \tag{2.5}$$

The relationship between FRR and FAR, as both are functions of threshold th, can be represented using the Detection Error Tradeoff (DET) [48]. FRR and FAR are inversely proportional as shown in Fig. 2.4 where different operating points are selected based on the nature of the biometric system application. For high-security applications such as logging into a bank account or physically logging into nuclear facilities which could expose sensitive information to hackers and attackers, the system design will consequently increase FRR. Whereas, for forensics and criminal applications, it is preferable to widen the area of suspects and hence, increase FAR. The point where both FRR and FAR are equal can be used in civilian applications and is denoted as Equal Error Rate (EER).

EER denotes the error rate at threshold th where $\mathrm{FAR}(\mathrm{th}) = \mathrm{FRR}(\mathrm{th})$ in the DET curve and it is widely used to measure the performance of a biometric system. A lower EER value means both FRR and FAR values are low and hence, a better performance for that particular biometric system. In addition, another graphical method to evaluate the performance of a biometric system is through plotting the FAR on the x-axis versus the GAR on the y-axis curve is called the Receiver Operating Characteristic (ROC) [49]. ROC is a 2D graph which clarifies the diagnostic capability of a biometric classifier as th varies and vitalizes its performance [50]. Fig. 2.5 shows the performance of four assumed biometric classifiers at a certain threshold value. If a classifier ROC value approaches the bottom left corner, that means no positive classification in terms of TP and FP is committed whereas moving

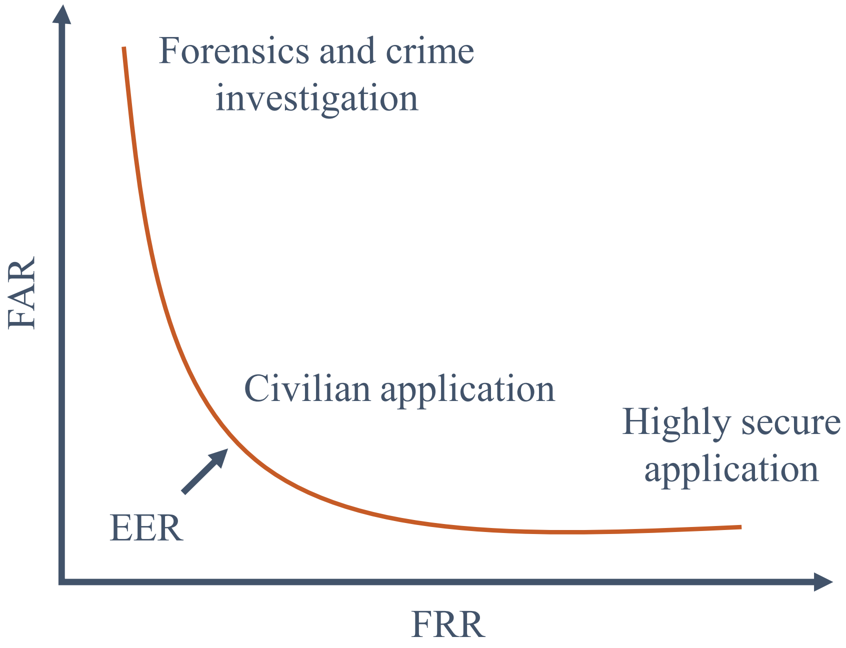


Figure 2.4: DET curve showing FRR versus FAR.

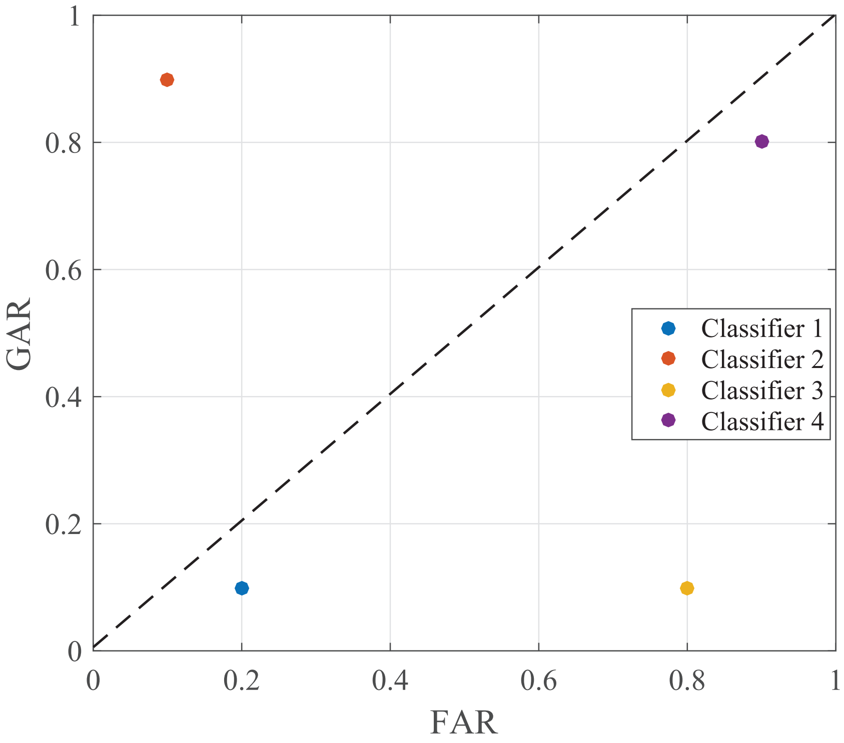


Figure 2.5: ROC plot of four assumed biometric classifiers.

towards the top right corner means the opposite strategy. In contrast, the more ROC value goes towards the top left corner, the better. As a result, classifier 2 has the best performance compared to other assumed classifiers showed in Fig. 2.5.

2.2.4 *Recent Blood Vessel Based Biometric Systems*

In this subsection, an overview of recent blood vessel based biometric systems is presented. The strong properties of blood vessels in various human parts such as the retina, fingers, palms, and the sclera means that it is difficult to forge and damage these vessels as they have internal features intrinsic to the human tissue. In addition, the capturing process of these vascular patterns cannot be achieved without the subject's consent. Therefore, vein patterns have been investigated as potentially powerful biometrics.

2.2.4.1 *Retina*

The retina is the inner and third tissue layer of the eye which is sensitive to light. The width of the retina layer is 0.5 mm and it consists of three layers of neural cells. The main function of the retina is to create an image of the visual world through the optics of the eye [51]. The fundus image in Fig. 2.6 shows the blood vessels in the retina which consist of veins and arteries. The veins are normally darker and slightly wider.

The blood vessel network within the retina has a unique vascular pattern which was discovered in 1935 while studying eye disease. Therefore, the pattern of blood vessels has been utilised to identify an individual [52]. There has been much research on retina identification [53, 54, 55]. However, due to the difficulty of capturing retinal images which requires using a special microscope with an attached camera and user cooperation, it is not widely used in the market.

Generally, a retina recognition system includes first capturing the fundus image using a special retina scanner device. Then, blood vessel segmentation is achieved through pixel thresholding, local contrast and morphological enhancing methods. Next, feature generation is achieved using blood vessel patterns like angles, diameter, and location of these blood vessels and finally a matching process is initiated to

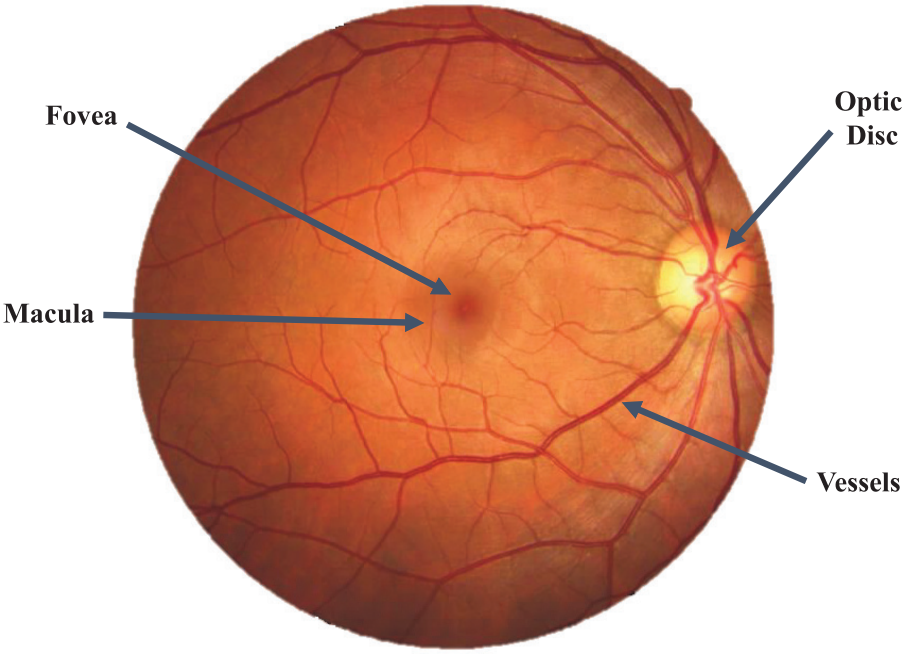


Figure 2.6: Human retina anatomy showing the fovea, macula, optic disc, and blood vessels.

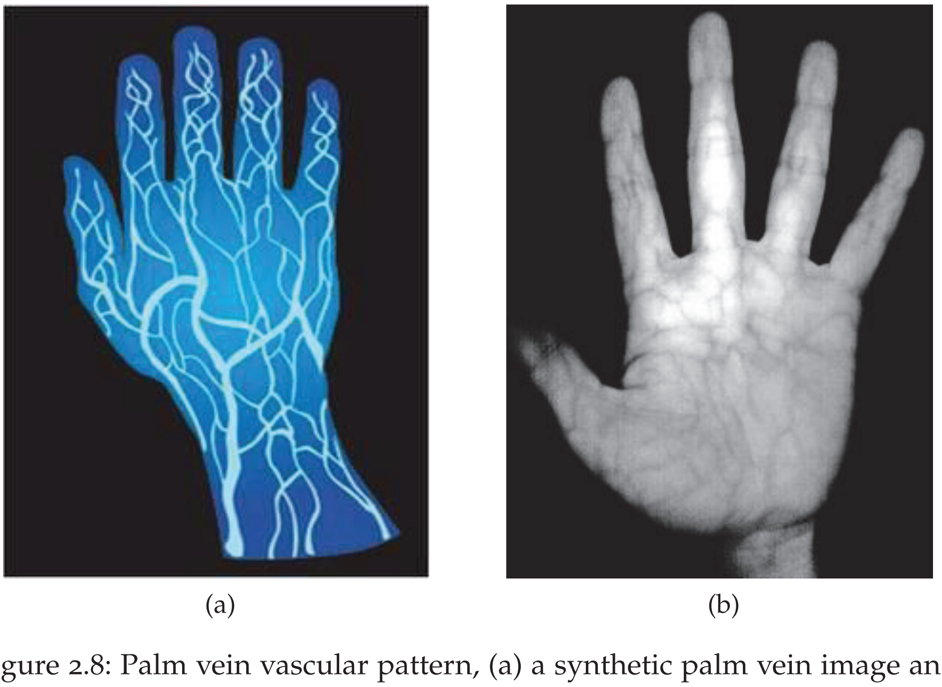
acquire a similarity measure. An example of blood vessel segmentation is shown in Fig. 2.7.

2.2.4.2 *Palm Vein*

The palm vein pattern has been extensively studied as a biometric trait recently due to its distinctive features [56, 57]. As shown in Fig. 2.8, the palm has a complex network of blood vessels which can be used for personal identification purposes [52, 58]. The acquisition of palm vein images is achieved by using NIR illumination in order to capture the vascular pattern with contact-free or constrained imaging setup. A palm vein identification system design is depicted in Fig. 2.9 which shows that the main tasks are to extract the Region Of Interest (ROI) and to enhance blood vessels area. Normally, finger locations are used to normalise and extract the palm ROI while for image enhancement, contrast stretching and histogram equalisation methods are applied as the NIR images have low contrast and appear darker in general [59].

The techniques used for palm vein identification may be grouped in two divisions which are vessel line/curve matching and holistic approaches using subspace learning [59]. An example of the first

Figure 2.7: An example of retina blood vessel segmentation which includes thresholding, local contrast and morphological enhancing methods.



(a) (b)

Figure 2.8: Palm vein vascular pattern, (a) a synthetic palm vein image and (b) a real image captured using NIR light source.

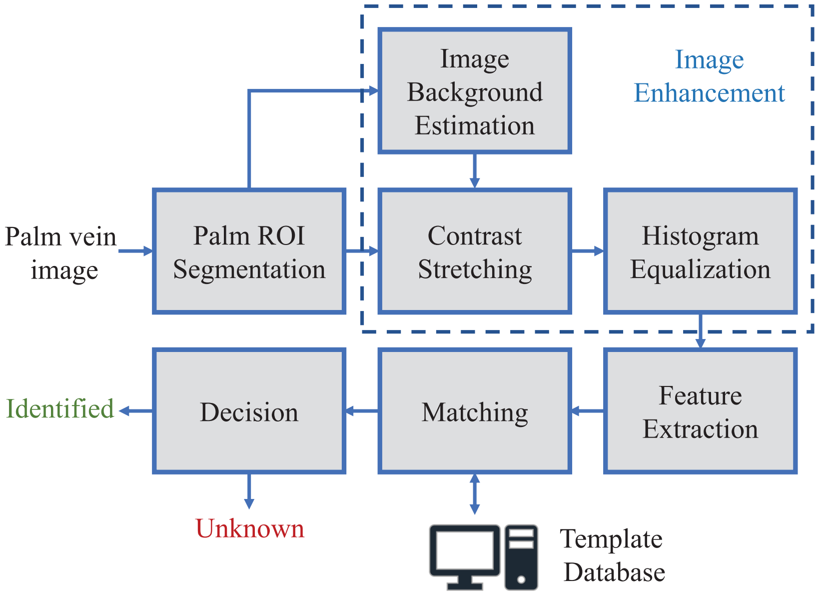


Figure 2.9: An example of retina blood vessel segmentation which includes thresholding, local contrast and morphological enhancing methods.

approach is applying some spatial domain filters such as Gabor and orthogonal Gaussian filters and extracting line and curve features from the output of these filters. Whereas for the holistic subspace learner approach, the whole palm image is projected in order to create training data such as via Principal Component Analysis (PCA) [60].

2.2.4.3 *Finger vein*

In recent years, there has been an increasing interest in finger vein recognition systems. This is due to the fact that the finger vein as compared to the typical fingerprint has advantages which can be summarised as: 1) the surface conditions of the finger has no effect on the blood vessels; 2) it is hard to forge or spoof a system using these blood vessels; 3) the capturing setup using infrared illumination along with contactless imaging can be considered as non-invasive imaging; 4) these vein patterns are stable and can be captured with low-resolution cameras [61]. Examples of finger blood vessels and the segmentation process output are shown in Fig. 2.10.

As in the palm vein capturing setup, the finger vein pattern is captured using NIR light with a wavelength (700-900 nm) [63]. As the

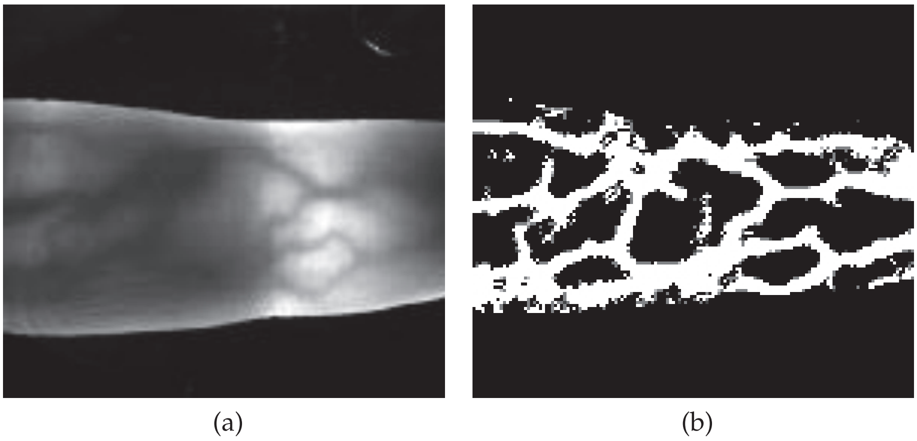

(a) (b)

Figure 2.10: An example of the vascular pattern within the human index finger. a) NIR image of finger blood vessels and b) blood vessel segmentation output [62]

haemoglobin absorbs more NIR radiation than muscles and bones, the vascular pattern can be therefore captured efficiently. However, and as the NIR light is absorbed by other tissues, this could lead to scattering of the light through the non-vessel areas and hence the contrast of these vessels can be reduced in the captured image [63]. Therefore, much research has been focused on improving the quality of the captured image with the optimum goal being to increase the recognition accuracy of the finger vein vascular pattern [64, 65, 66].

2.2.4.4 *Sclera blood vessels*

Recent research on biometrics has shown an increased interest in new human traits rather than the typical biometrics [67]. In particular, human recognition system based on the vascular pattern in the sclera [17]. The sclera can be defined as the white and opaque outer protective part of the eye. It consists of four tissue layers: the episclera, stroma, lamina fusca and endothelium [68] which surround the iris. The iris is the coloured tissue around the pupil. The structure of the human eye is shown in Fig. 2.11 containing the pupil, iris, periocular, and sclera regions. The sclera has visible blood veins which are randomly distributed in different orientations and layers making them a promising factor for the improvement of an eye recognition system [69] and ultimately, suitable to be employed as a biometric trait as it can be

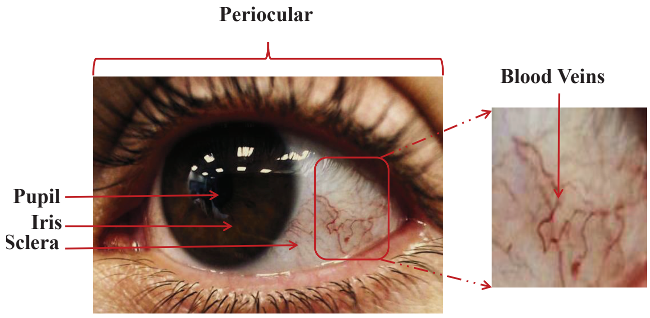


Figure 2.11: Eye anatomy consisting of pupil, iris, sclera and periocular regions.

imaged at a distance with visible wavelength light [17]. Acquisition of the sclera can be achieved using a normal RGB coloured camera device such as a DSLR, compact, webcam, or mobile camera giving an advantage to the sclera over the iris which has a better recognition performance in the NIR spectrum.

A typical vascular pattern recognition system based on the eye is shown in Fig. 2.12 which consists of iris segmentation and eyelids border detection, sclera segmentation and validation, vessel enhancement and mapping, feature extraction and template registration, database creation, classification and decision-making. Each part of the system design has its challenging issues which recent research on sclera recognition has endeavoured to resolve (Table 2.2).

Sclera segmentation methods can be classified in terms of human intervention into three kinds defined as: 1) high accurate manual segmentation where sclera regions are manually cropped as the work could focus on other sclera recognition design aspects; 2) semi-automated

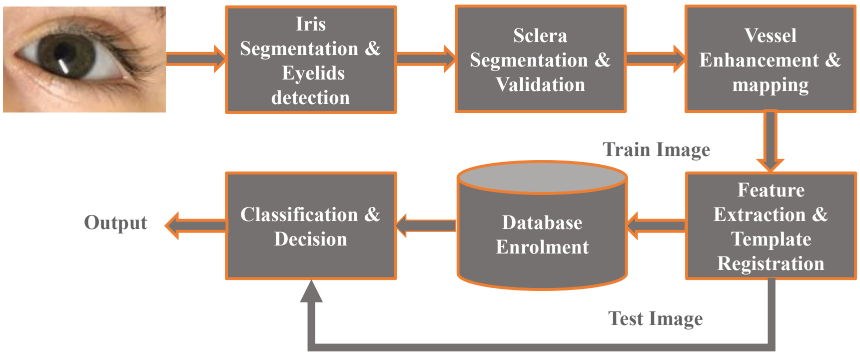


Figure 2.12: Typical sclera recognition system design.

segmentation where an automated segmentation algorithm is applied first, then a human intervention is used for further refinement; and 3) automated segmentation where a set of processes is designed to achieve the segmentation and sclera refinement without any human intervention. However, the automated segmentation could suffer from accuracy degradation especially when noise factors are introduced. As a result, much research has been conducted to overcome the segmentation accuracy degradation (Table 2.2). Also, two strategies have been suggested to segment the sclera based on either pixel values or sclera contour shape. Examples of manual and automated segmentation results on the sclera area are shown in Fig. 2.13. It is evident that the manual segmentation approach achieves higher accuracy than the automated segmentation approach but it is not applicable in a real-world application.

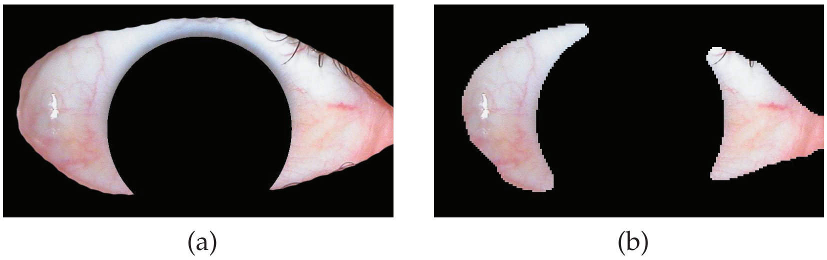

(a) (b)

Figure 2.13: Examples of manual and automated sclera segmentation methods. a) a manual segmentation result and b) is the result of an automated segmentation method.

On the other hand, vessel enhancement is a crucial step in the design as the vascular pattern in the sclera is not distinguished completely and hence preprocessing steps are required to enhance vessel contrast. This includes selecting the best channel from one of the colour spaces, for instance, the green component of the RGB colour space, and enhancing vessel separation from its background via applying a certain technique to enhance vessel pixel intensities such as histogram equalisation and spatial filtering. Most of the work in sclera recognition literature summarised in Table 2.2 has proposed vessel enhancement by first applying Contrast-Limited Adaptive Histogram Equalisation (CLAHE) on the green channel of the RGB image. This is because using the green channel has shown a better contrast between the vascular pattern and the background [70]. However, extra processing steps were added by researchers to improve vessel enhancement.

Similar to the hand or palm vein-based recognition system for feature extraction, two approaches are adopted in sclera vein feature extraction. These approaches can be defined as 1) holistic approach where the enhanced blood vessel image is projected into a linear or non-linear subspace learning method such as PCA resulting in a lower dimensionality feature vector; 2) the second approach focuses on converting the vascular pattern into a binary single-pixel skeleton and extracting vessel features such as vessel position, angle, rotation, and distance to a reference point to create the user template.

Finally, a similarity distance measure is utilised to classify the vascular pattern features in the enrolment stage and to compare features extracted from two biometric templates either in the verification or identification mode. The matching process normally encapsulates a decision process where a decision based on these measures is made which is whether to accept or reject a user. The next section will present a literature survey on recent work on sclera recognition and the best achieved results in terms of identification, matching, and error rates.

Table 2.2: Recent literature related to vascular pattern recognition in the sclera

Authors	Year	Sclera segmentation	Vessels enhancement	Feature Extraction	Database(subj./img.)	Learning-based	The best performance
Derakhshani et al. [17]	2006	Manual	CLAHE, region growing method	Hu's invariant moments, minutiae	In-house[a] DB (6/24)	No	100% identification rate
Derakhshani et al. [71]	2007	Manual	CLAHE	CDF 9/7	In-house DB(50/1200)	Yes 6 TIPU[b]	4.3% EER
Crihalmeanu et al. [72]	2009	Semi-automated (pixel thresholding)	CLAHE, selective enhancement filter	Affine transform	In-house DB(50/134)	No	25% EER
Thomas [73]	2010	Automated (pixel thresholding)	Gabor filters bank	Line segments description	UBIRIS.v1 (241/1205) IUPUI (45/178)	Yes 2 TIPU	3.38% EER 4.78% EER
Zhou et al. [74]	2010	Automated (pixel thresholding)	Gabor filters bank	Line segments description	UBIRIS.v1 (241/1805)	-	5% EER
Tankasala et al. [75]	2011	Automated (pixel thresholding)	CLAHE	GLCM	In-house DB(50/1200)	No	11.9% EER
Crihalmeanu and Ross [76]	2011	Automated (pixel thresholding)	CLAHE, selective enhancement filter	Affine transform	In-house DB(49/1536)	No	0% EER
Zhou et al. [77]	2011	Automated (pixel thresholding)	Gabor filters bank	Line segments description	IUPUI (44/352)	-	70% matching
Oh and Toh [78]	2012	Automated (pixel thresholding)	histogram equalisation low-pass filtering	Angular grid, MLBP	UBIRIS.v1 (241/1205)	No	0.47% EER
Gottemukkula et al. [79]	2012	Manual	CLAHE	Iris and sclera features	In-house DB(50/194)	Yes 2 TIPU	4.01% EER
Crihalmeanu and Ross [80]	2012	Automated (pixel thresholding)	CLAHE, selective enhancement filter	SURF	In-house DB(103/3280)	-	0.8% EER
Zhou et al. [31]	2012	Automated (pixel thresholding)	Gabor filters bank	Line segments description	UBIRIS.v1 (241/1805) IUPUI (44/352)	Yes 2 TIPU	4.09% EER 11.89% EER
Tankasala et al. [81]	2012	Manual	CLAHE	LDA	UBIRIS.v1 (40/160)	Yes 2 TIPU	8.69% EER

Continued next page

Table 2.2: (continued) Recent literature related to vascular pattern recognition in the sclera

Authors	Year	Sclera segmentation	Vessels enhancement	Feature Extraction	Database(subj./img.)	Learning-based	The best performance
Das et al. [82]	2013	Automated (shape contour)	AHE, Haar wavelet filters bank	D-SIFT	UBIRIS.v1 (241/1205)	Yes 3 TIPU	99.34% Acc[c]
Tankasala et al. [83]	2013	Semi-automated (pixel thresholding)	CLAHE	Contourlet transform	In-house DB(40/400)	Yes 5 TIPU	0.16% EER
Das et al. [84]	2013	Automated (shape contour)	AHE, Haar wavelet filters bank	OLBP	UBIRIS.v1 (241/1151)	No	0.52% EER
Lin et al. [85]	2014	Automated (pixel thresholding)	Gabor filters bank	Line segments description	UBIRIS.v1 (241/1168)	Yes 2 TIPU	3.05% EER
Das et al. [86]	2014	Automated (shape contour)	AHE, Descrete Meyer wavelet filters bank	D-LDP	UBIRIS.v1 (241/1350)	Yes 5 TIPU	3.95% EER
Tankasala and Doynov [87]	2015	Semi-automated (pixel thresholding)	DoG	Shearlet transform	CIBIT-I DB(50/-)	Yes	0.29% EER
Gottemukkula et al. [88]	2016	Automated (pixel thresholding)	CLAHE	HS, FAST, SURF descriptors	UBIRIS.v1 (241/1205) In-house DB(226/-)	Yes -	0.82% EER 0.04% EER

[a] In-house DB: private database created by authors.

[b] TIPU: Training images per user.

[c] Acc: Accuracy.

2.3 LITERATURE REVIEW

This section will present the recent literature on sclera recognition, which is also summarised in Table 2.2, in terms of sclera segmentation, vessel enhancement, feature extraction, template registration, matching process, and the best-reported results. It is noteworthy that blood vessel recognition within the eye has been named differently in recent research as sclera, ocular, and conjunctival vasculature recognition. However, they all have similarities in terms of extracting discriminative features from these vessels in order to recognise an individual.

2.3.1 *Recent Literature*

The reported results summarised in Table 2.2 represent best results in the recent literature in terms of EER, identification rate, and matching rate between two sclera templates. Although different recognition techniques, databases (in-house, public), and number of images are proposed in the recent work, these results however provide an overview of the advancement and challenges related to designing an efficacious sclera recognition system.

The first work on sclera recognition was conducted by Derakhshani et al. [17]. They utilised manual segmentation, which is inefficient, for the ROI which is a rectangular scleral region of the off-angle iris image, as this was an initial study to check the feasibility of using the vascular pattern within the eye. After applying CLAHE on the green component of the RGB image, they used a selective line enhancement method [89] which locates blood vessel lines and curves while suppressing other objects. The vessel feature matching approach was used where a skeleton binary network is first created using a growing method which is an iterative labelling method. There are only two labelling classes wherein each pixel is classified as either blood vessel or background and which uses Otsu's thresholding method for initial binarization [90]. After that, they proposed two levels of classification which can be defined as coarse-level matching and detailed matching levels. The coarse-level matching is the first step of classification where the aim is to reduce the search time for large-scale databases and creates a subset consisting of most probable matches. First Hu's invariant moments [91] are used to extract vessel features and the evaluation between two vectors is achieved by a Euclidean separation metric. On the other hand, for the detailed matching, a minutiae detection method is applied. Finally, Derakhshani et al. [17] used an in-house database collected from 12 eyes belonging to six subjects (both right and left eyes) in two sessions resulting in 24 images in total using a commercial 5 MP Sony camera captured in constrained conditions. Left gaze eye direction is used and the best-reported result is 100% identification rate. The major criticism of their work is the manual segmentation, constrained imaging conditions where noise factors have been minimised, and the small number of participants.

Then, Derakhshani and Ross [71] also applied manual segmentation but with one exception that they utilised an intensity-based routine to enhance the segmentation furthermore. For vessel enhancement, applying CLAHE is achieved by dividing each right and left sclera side into 8×8 tiles and choosing 256-bin histograms for the CLAHE resulting in a better local enhancement. This routine is also used in [81]. After that, a special type of the Discrete Cosine Transform (DCT) which is the Cohen-Daubeches-Feauveau 9/7 wavelet (CDF 9/7) is utilised together with a neural network classifier. The CDF 9/7 is an efficient biorthogonal wavelet used in image compression by the Federal Bureau of Investigation (FBI) for fingerprint images and in JPEG2000 image format [92]. First, the sclera segments are down-sampled into 100×200 pixels and a two-dimensional CDF 9/7 with eight levels is applied and only 512 components of the filtering process are retained using lossy compression. While for the matching process, a single hidden layer feed forward neural network is utilised as a classifier with 512 input nodes, 300 hidden nodes, 50 output nodes with bi-valued targets and hyperbolic tangent non-linearities. Then, the database used is an in-house database where 50 volunteers participated to capture their eye images with 4 gaze directions (left, right, up, and straight) using a Cannon 20D digital SLR camera with EFS 60 mm f/2.8 Macro and EF 70-200 mm f/4L USM telephoto lenses. The capturing distances are 30 cm with a macro lens, and 152 cm and 274 cm receptively with a telephoto lens. The number of images used is 1200 images for left and right gaze directions and the best results in terms of EER are (4.3%, 8.8%, and 9.2%) for capturing distances in cm (30, 152, and

274) respectively. Their method is also applied in constrained imaging capturing with semi-automated segmentation.

Next, a semi-automated technique was proposed in [72] and also used in [87] which involves using the K-means clustering method, which suffers from a computational load with high dimensional data [93], with human intervention for further refinement. The K-means clustering method represents each pixel as a three-dimensional point in the Cartesian system based on the three spectral RGB values. Then, some parts of the eye which are erroneously segmented are manually removed for final refinement. Crihalmeanu et al. [72, 80] then removed specular reflections as these reflections of light could cause problems for image processing algorithms and hence resulting in spurious results before applying a similar technique for vessel enhancement suggested in [71]. Next, a holistic image feature is proposed where image registration using a local and global affine transformation suggested in [94] is utilised resulting in image alignment with contrast and brightness variation consideration. Their method was also used in [76]. Then, a cross-correlation between two sclera images is measured. The database used to evaluate their method was another in-house database captured using Sony CyberShot DSC F717 with 5 MP and 50 volunteers. The total number of images used were 100 images from session 1 and 34 images from session 2 for the left eye captured at three distances 1 ft, 9 ft and 12 ft and for the left gaze direction and EER equals 25%. Despite the small number of images used and the semi-automated sclera segmentation suggested in [72], the EER was dramatically high (25%).

After that, Thomas [73] suggested an automatic method for sclera segmentation which has also been utilised in [74, 77, 95, 85]. It involves creating two maps based on the sclera properties and eyelid skin using RGB and HSV colour spaces after down-sampling the eye image. Then, a convex hull is applied to refine the sclera area. Next, iris and eyelid borders are detected and refined producing the final sclera mask which is up-sampled to the image original size. Then, he suggested a different approach for vessel enhancement which was also used in [74, 31, 77, 95, 85] where a bank of Gabor filters is utilised with only even filters and multiple directions to enhance blood vessels. Then an adaptive thresholding is utilised to create the final skeleton binary network of these blood vessels. Next, a new method for a vessel line descriptor which is also used in [74, 77, 31, 95] is proposed where three quantities including line angles to pupil centre, line distance from pupil centre, and the dominant orientation of the line segment are used as features for the vascular pattern. For template registration, the Random Sample Consensus (RANSAC) method was used where sets of points representing the centre of the line segments were registered first for the training templates. Then, the one which minimises the distance between training templates was chosen as the optimal template. Euclidean distance between centre points and angles of line segments belonging to the registered template and target template was calculated. To evaluate his proposed work, Thomas [73] utilised two databases which are publicly available UBIRIS.v1 [96] and in-house IUPUI multiple wavelength databases. He used high quality 1205 and 178 images for the UBIRIS.v1 and the IUPUI databases respectively and the best EER result is 3.38% for the UBIRIS.v1 database. His method however is lacking eye localisation, sclera segmentation validation, complexity evaluation, and only evaluated on constrained imaging databases.

Tankasala et al. [75, 83] applied the K-means clustering suggested in [72] for the sclera region classification. Then, a convex hull was applied to remove specular reflection holes and misclassified iris pixels from the sclera region. The CLAHE on the green component is utilised in [75, 79, 83, 88]. Then, Tankasala et al. [75, 81] considered using the Gray Level Co-occurrence Matrix (GLCM) [97] as a feature extraction method. GLCM measures the co-occurrence of grey levels in an intensity image by considering pixel pairs separated by distance and angle vectors. Then, a square matrix is generated with size equal to the number of selected grey levels which is eight. Then, statistical features such as energy, contrast, correlation, and homogeneity are calculated from the GLCM square matrices. For feature classification, linear and non-linear methods are investigated in [75, 87] which are Linear Discriminant Analysis (LDA) [98] and a neural network respectively. The LDA reduces the dimensionality of GLCM features whereas the employed neural network, which is a non-linear data-driven classification method, generates a predictor from the enrolled feature vectors. Matching scores from the two sclera left and right sides are averaged in [75] using the sum-rule method to calculate the overall sclera matching score. Then, both sclera and iris matching scores are fused similarly to evaluate the bimodal system performance. Finally, the same in-house database collected by authors in [71] is used in [75] and the best result is 11.9% EER for using LDA and a neural network score fusion. The result is evidence that their system performs ineffectively although they fused matching scores from two methods and used high-quality in-house images.

On the other hand, and for investigating multi-spectral images in ocular vascular pattern recognition, Crihalmeanu and Ross [76] used NIR spectral illumination to segment the sclera. The segmentation process consists of three complex stages defined as 1) coarse sclera segmentation where a normalised sclera index is first generated from the NIR and green spectral spaces; 2) pupil region segmentation; 3) final sclera segmentation which involves iris segmentation where K-means clustering is used to segment the iris for final sclera mask refinement. For vessel enhancement, the L*a*b colour space [99] converted from RGB is used where CLAHE is applied on the L* component. The enhancement process is utilised on smaller tiles of the images adding more processing time. Then, the image is converted back to RGB and the green component is extracted. Finally, a selective enhancement filter for lines is used. After that, they created an in-house database using the Redlake MS3100 multi-spectral camera for coloured and NIR images. The lighting of the capturing device is an ophthalmologist slit lamp with a mount consisting of a chin rest and flexible moving camera arm to control image focus resulting in a complicated capturing procedure. The resolution of the captured images is $1035 \times 1373 \times 3$. Then, videos of left and right eyes with left and right gazes are recorded and 8 images per eye per gaze are extracted for 49 subjects resulting in a total of 1536 images. The capturing is conducted under a controlled environment. The best result is 0.0% EER on the left eye with left gaze due to constraints on imaging acquisition and small subject number.

Zhou et al. [77] utilised the sclera segmentation and vessel enhancement methods suggested in [73]. However, a matching score fusion was proposed for multi-angle (front and left) sclera images where four different fusion methods were investigated. They utilised the IUPUI database and selected 352 images for 44 subjects captured under the green wavelength only. They achieved 70% matching rate between two compared templates.

The saturation component in the HSV colour space is reversed and partitioned into two sub-images based on the iris centre for enhanced illumination direction adaptability in [78, 100]. Then, a pixel thresholding process is used to binarize both sub-images after applying histogram equalisation and low-pass filtering. Then, Oh and Toh [78, 100] utilised a histogram equalisation process and low-pass filtering with a binary threshold for vessel enhancement. Next, an angular grid is used to provide candidate feature locations. For matching binary features, they used normalised Hamming distance as binary features are extracted from the vascular pattern using the Multiresolution Local Binary Pattern (MLBP) [101]. Similarly, the UBIRIS.v1 dataset is used in [78] with only 1205 best quality images and the best EER is 0.47% making their method unrealistic for noisy images captured under a real-life scenario. In addition, Oh et al. [100] combined periocular and sclera features at the score level and evaluated their system using the UBIRIS.v1 database. Both capturing sessions were utilised and the best EER was around 5% even after using a multimodal biometric system.

Gottemukkula et al. [79] studied fusing iris and sclera vascular patterns to improve the ocular biometric in the visible spectrum and hence, a manual segmentation was applied as it was not the focus of the proposed work. Then, two feature vectors representing the mean and variance values for each pixel block are extracted from sclera blood vessels with tile size defined either 160 or 260 depending on the position of the pixel block, and Daugman's method [32] was applied to extract iris features. After that, they utilised a set of similarity scores for the vascular pattern and dissimilarity score for the iris and used score fusion to present the final result. Once the fusion is performed, the proposed multimodal biometric system is evaluated by collecting an in-house database for 50 subjects with entire face captured and a resolution of 2832×4256 using a Nikon D3S FX 12.1 MP and micro Nikon 105 mm lens. The ambient lighting was blocked and a front ring flash was used with additional pre-flash red LEDs to reduce the dilation of the eye for capturing the iris and the sclera simultaneously. They captured 194 images under very constrained imaging and achieved 0.53% EER with controlled imaging illumination.

Crihalmeanu and Ross [80] developed their segmentation method in [72] by adding preprocessing steps including applying image denoising through the Double-Density Complex Discrete Wavelet Transform (DDCDWT) [102] and eliminating specular light reflections. Then, three separate feature extraction and matching methods are suggested in which the first two techniques focus on extracting interesting points from the vascular pattern whereas the third technique deals with the entire sclera image. The first technique is achieved via applying Speed Up Robust Features (SURF) [103]. While the second and third techniques are described in [17] and [72] receptively. The Euclidean distance is used to measure the similarity between interesting point pairs where only pairs with a distance greater than 0.1 are counted. Subsequently, Crihalmeanu and Ross [80] extended their in-house database collected in [76] with an increased number of 103 subjects and number of images equal to 3280 using the exact imaging setup. Three methods were used for feature extraction and the best results produced via applying the SURF was 0.8% EER.

Zhou et al. [31] utilised two methods for coloured and greyscale images. The first method for coloured images is described in [73] whereas the second method uses Otsu's thresholding method [90] where the eye image is divided into left and right sections based on the iris centre and the thresholding process is applied separately. No fusion process has however been applied between the masks resulted from the RGB and greyscale spaces. Therefore, suggesting sclera segmentation for greyscale colour space was unnecessary. Two databases were utilised to evaluate the proposed work which are the UBIRIS.v1 and IUPUI multiple wavelength databases. They replicated a similar work proposed in [73] but an increased number of images for both databases was used. 1805 images from session 1 and 2 for 241 subjects and 352 images from 44 subjects from the UBIRIS.v1 and IUPUI databases respectively were employed. The best result using high-quality session 1 images from the UBIRIS.v1 database is 4.09% EER.

Tankasala et al. [81] evaluated their work proposed in [75], where a bimodal ocular biometric system is designed, using the UBIRIS.v1 database. Therefore, for sclera segmentation, and as the captured eye images in the UBIRIS.v1 database have only front gaze direction, two maximum rectangle areas on the left and right sclera starting from iridial limbic boundaries are manually cropped. Another drawback is they performed a visual quality assessment to select the best iris and sclera images from the UBIRIS.v1 database. Only 160 images for 40 subjects were utilised in the evaluation with two terms of capturing time which were short-term for images from the same session and long-term for images from different sessions. EER of 8.69% was the best-reported result using only the sclera and EER of 4.52% was achieved after fusing sclera and iris matching scores.

Tankasala et al. [83] applied the Laplacian pyramid decomposition with a decomposition level equal to 4 to extract low-pass and band-pass versions of an input image. Then, a directional filter bank is utilised to divide the sub-bands into vertical and horizontal directions. After that, four statistical features (median, mean, variance, and entropy) are extracted from RGB, HSV, and YCbCr colour spaces. Then, they created an in-house database with 40 volunteers using Canon T2i DSLR with a macro ring LED illumination and two extended lateral lighting pads. Two capturing sessions were performed in still and video modes and the total number of images was 400 including both eyes and two gaze directions (left and right). The best-reported result was when using still images and it equals to 0.16% EER.

In [95], a quality-based fusion is adopted for iris and sclera matching scores where only the best score of either modality is chosen based on their quality thresholds. Session 1 images in the UBIRIS.v1 database were used to evaluate the work and 94.07% GAR is achieved using the sclera as a single modality and was improved after applying quality score fusion on the sclera and iris matching scores to 96.42%. However, no reported EER was included.

The red channel is considered for sclera segmentation in [82, 84] as the blood vessels are less prominent and thus avoiding the impediment of contour growing. Then, Adaptive Histogram Equalisation (AHE) is applied with 2×2 tiles to reduce vessel contrast. After that, a low-pass bank of Haar reconstruction filters is used to eliminate all the vessels in the sclera. Finally, a time-adaptive active contour-based method is utilised for the right and left sides of the sclera. Whereas for vessel enhancement, Das et al. [82, 84] applied AHE with a window size of 42×42 in the green channel. Then, a bank of 2D decomposition Haar wavelets is used for further vessel contrast. Feature extraction is performed based on the Dense-Scale Invariant Feature Transform (D-SIFT). D-SIFT descriptors of the training images are merged to form a bag of words representing the final training module. Then, the K-means clustering method is applied on the training module to generate a codebook with a vocabulary size equal to 1024. For feature classification, Support Vector Machines (SVMs) are utilised for mapping the extracted features into a higher space of dimensionality. Then, a linear plane which separates data in that higher space is investigated by the SVMs using three kernel functions which are linear, polynomial, and Radial Basis Function (RBF). Likewise, Das et al. [82] utilised session 1 from the UBIRIS.v1 database to evaluate their proposed work. They reported a recognition accuracy of 99.34% using 3 images for training and 2 for testing using only high-quality images.

The Oriented Local Binary Pattern (OLBP) [104], which is an extension of the MLBP which can express the orientation information more explicitly, is utilised in [84] to generate a binary representation of the vascular pattern. Then, image registration is achieved by subtracting the image centre from the iris centre to attain an independent image translation. Only session 1 from the UBIRIS.v1 database was utilised and the best EER reached was 0.52%.

Fuzzy C-means clustering [105] is used for sclera segmentation in [86, 106] which allows for one-pixel data to belong to more than one cluster and its fuzzy partitioning is achieved through an iterative optimisation. The AHE was applied first on the green channel. Then, a bank of Discrete Mayer Wavelets [107] is employed to enhance the image. Next, Dense-Local Directional Pattern (D-LDP) method [108] is proposed for feature extraction as it computes the edge response values in multiple directions. Das et al. [86, 106] used multi-session context using the UBIRIS.v1 database and achieved 3.95% EER.

Meanwhile for feature registration, and as the line descriptor proposed in [73] is time-consuming, Lin et al. [85] therefore proposed an improved descriptor based on the Y shape which can improve the coarse registration of two sclera templates. In addition, the new descriptor is used as a filter for some non-matching pairs. They proposed two coarse-to-fine matching stages. The Y descriptor is used as a coarse matching whereas a fine matching is achieved through the Weighted Polar Line (WPL) descriptor which can automatically align two templates. Lin et al. [85] also utilised the UBIRIS.v1 session 1 images and reported a 3.05% EER.

Radu et al. [109] proposed a new sclera segmentation method only based on multiple classifier system. The process involves extracting specific features (colour, edges, shape) from the eye image first to achieve higher recognition accuracy using three different feature extraction methods and concatenating their features into a single vector. Then, three linear classifiers which are Bayes classifier with linear boundary, LDA, and the regularised logistic regression are used with non-linear feed forward neural network to classify each pixel into sclera, non-sclera labels.

Gottemukkula et al. [88] utilised the blue channel and applied a median filter to remove the imprints of any remaining blood vessels. Then, light boundary suppression is adopted to enhance edges. After that, morphological operations are performed to eliminate some regions of the eyelids. Finally, a convex hull is applied to smooth the final sclera mask. However, no segmentation accuracy or processing time was reported.

Tankasala and Doynov [87] proposed a new vessel enhancement method based on the Difference of Gaussian (DoG) method. Then, a multi-scale multi-directional Shearlet Transform is proposed as it involves an additional shear operation to control the direction of the wavelet. Two in-house databases were collected by the authors in [87] named CIBIT-I which were captured using a Canon 20D DSLR camera and 50 volunteers with left and right gaze directions and at three distances: 9ft, 5ft, and 1ft. While the second database was CIBIT-II which involved 35 volunteers and their images were captured using an iPhone 4s back facing camera. The best result of EER was 0.29% acquired from the CIBIT-I database.

Subsequently, combination of three interesting point detectors is utilised in [88] which are the Harris-Stephan (HS) [110], Features from Accelerated Segment Test (FAST) [111], and SURF detectors. Next, a two-stage matcher is designed where the first matcher is based on the RANSAC method and generates matching score 1 and registers the verification template via a geometric transformation matrix for the second matching stage. While the second matcher using Gabor kernels performs global matching to produce matching score 2 which will be fused with matching score 1 for final recognition decision. After that, one publicly available database which is the UBIRIS.v1 and one in-house database were utilised to evaluate the proposed work. The best EER achieved is 0.82% on the UBIRIS.v1 using only session 1 and 0.04% on the in-house database.

Finally, most of the aforementioned work on sclera recognition has not reported the accuracy of the proposed segmentation method, has suggested a complex recognition system with a high computational load using either single or multimodality, and used an in-house or high-quality database providing an unrealistic solution for a real-world application.

2.3.2 *Limitations and Challenges*

Several issues and challenges remain for recent sclera recognition work as shown in Table 2.2, which may affect the system performance. These are: **1)** sclera segmentation with pixel thresholding could be affected by the noise and distortion present in sclera images; **2)** the effects of the sclera boundary on the convergence of the sclera shape contour; **3)** occluded or partial occluded and noisy images are discarded manually; **4)** the enhancement of blood vessels and the feature extraction algorithm should be invariant to non-linear blood vessel movement [31]; **5)** robust user template registration and an efficient matching procedure are required; **6)** previous evaluations of sclera segmentation schemes have not used databases with unconstrained eye image conditions or heavily noise eye images; **7)** most of the manual, semi-automated and automated sclera segmentation methods have not exploited more than one colour space; **8)** eye rotation alignment method has not been developed which could affect the blood vessel angles and position matching; **9)** multiple eye gaze directions were not considered when extracting the shape contour during segmentation; **10)** blood vessel enhancement and feature extraction methods have not considered the lack of shift invariance, poor directional selectivity in discrete wavelets, together with redundant and heavy computations for Gabor wavelets.

In addition, real-time challenges in recent sclera recognition systems can be summarised as: **i)** Images used are captured under controlled environments where most of the noise factors have been minimised such as capturing distance, blurring, gaze direction, pose direction, capturing on-the-move, and natural illumination conditions. As a result, reported results in terms of system performance and error rates may not be realistic in real life applications. **ii)** Eye localisation, eye corner positions, and gaze directions were manually annotated and labelled which is not efficient for unsupervised applications. **iii)** Various image qualities have not been evaluated in terms of sclera segmentation and recognition rates. **iv)** The only work on sclera image quality measure was conducted by Zhou et al. [156] where only images with constrained imaging conditions were considered. In addition, the quality measure involves discussing the occlusion cases and counting extracted features which not necessarily measure the image quality rather than feature extraction method quality. **v)** The ocular recognition using a smartphone proposed by Gottemukkula et al. [88] has used only one mobile device and the dataset acquisition was inside a lab environment with controlled illumination conditions and capturing distance.

2.4 DATASETS UTILISED IN THIS THESIS

The following section explains in detail the datasets employed in this research and their properties and challenges.

2.4.1 *The UBIRIS.v1 database*

The UBIRIS.v1 database [96] is the most utilised data in the recent literature in eye biometrics including both the iris and the sclera due to the availability to the public, population variety, many noise factors, and being captured under visible wavelength illumination. This feature in particular has attracted researchers to investigate iris and sclera recognition under visible wavelength spectrum.

The main propose of the UBIRIS.v1 database is to evaluate the iris recognition under different types of noise and simulating the acquisition of the eye image with minimal user collaboration. The first work on the sclera using the UBIRIS.v1 database was by Thomas [73] as he compared sclera and iris under visible wavelength illumination and the results were promising.

The acquisition device utilised is a Nikon E5700 camera 71 mm focal length, 4.2 F-number, 1/30 sec exposure time, and ISO-200 speed settings with RGB colour representation. The captured images have a dimension of 2560×1704 pixels with 300 dots per inch (dpi) horizontal and vertical resolution and 24-bit depth and have been saved in JPEG format [73]. Image dimension was reduced into two size and colour versions which are 800×600 24 bit colour, 200×150 24 bit colour, and 200×150 greyscale versions.

The UBIRIS.v1 database is collected from 261 individuals composed of 1877 RGB images in two capturing sessions with front-gaze iris direction. In the first session which consists of 1214 images from 241 users, noise factors such as reflections, luminosity and contrast were minimised as the images were captured inside a dark room with 20 cm capturing distance. While in the second session which is composed of 663 images and involved only 135 from the 241 users, capturing location was changed in order to introduce a natural luminosity factor which introduced more reflections, luminosity and focus problems.

The UBIRIS.v1 database includes the images captured by a vision system with or without minimising the collaboration of the subjects. Some images have a poor quality condition such as irritated sclera, blink, no-sclera area, closed eye, severe blurring, un-centred un-cropped eye-area and poor lighting. Therefore, images from both sessions have been classified manually into three quality labels (good, average, and bad) depending on three parameters which are: visible iris, focus, and reflections with the overall percentage of these three parameters for the good label being 56.48%.

Some examples of the UBIRIS.v1 database good and poor quality images in both sessions are shown in Figs. 2.14, 2.15, 2.16, and 2.17. As the Daugman method [32] for the iris segmentation achieved higher rate compared to other iris segmentation methods on the UBIRIS.v1 database [96], it is therefore employed in the proposed contributions in Chapters 3, 4, and 5.

2.4.2 *The UBIRIS.v2 database*

As the UBIRIS.v1 database received criticism over the realism of its noise factors and the constrained imaging and acquisition process,

Figure 2.14: Examples of good quality images in the UBIRIS.v1-session 1.

Figure 2.15: Examples of poor quality images in the UBIRIS.v1-session 1.

Figure 2.16: Examples of good quality images in the UBIRIS.v1-session 2.

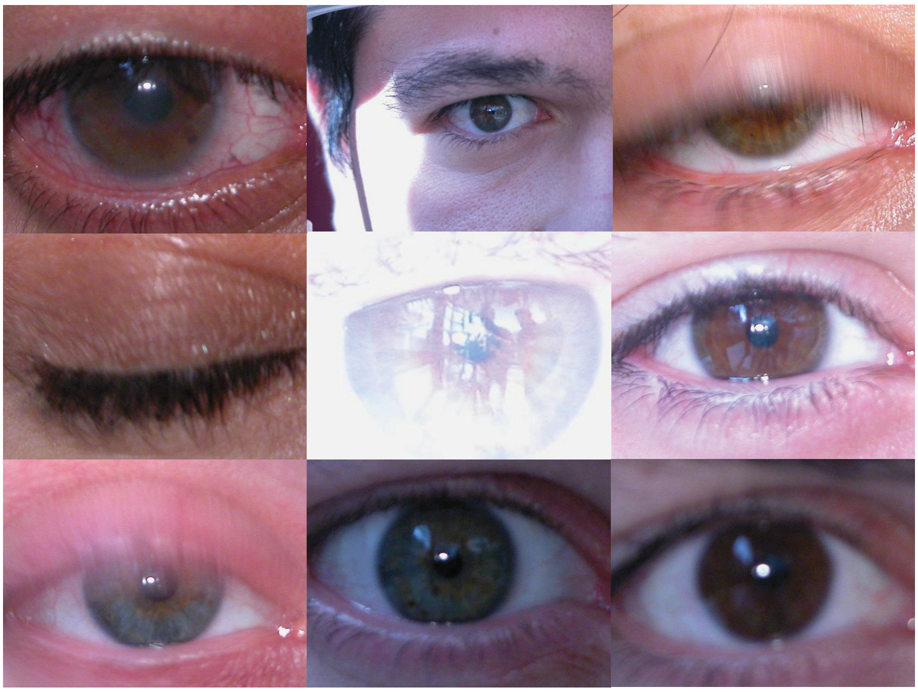


Figure 2.17: Examples of poor quality images in the UBIRIS.v1-session 2.

the development of a new database is therefore achieved named the UBIRIS.v2 database [112]. The major propose of establishing this database is to evaluate the iris biometric under visible wavelength with non-ideal imaging conditions and its feasibility at-a-distance and on-the-move.

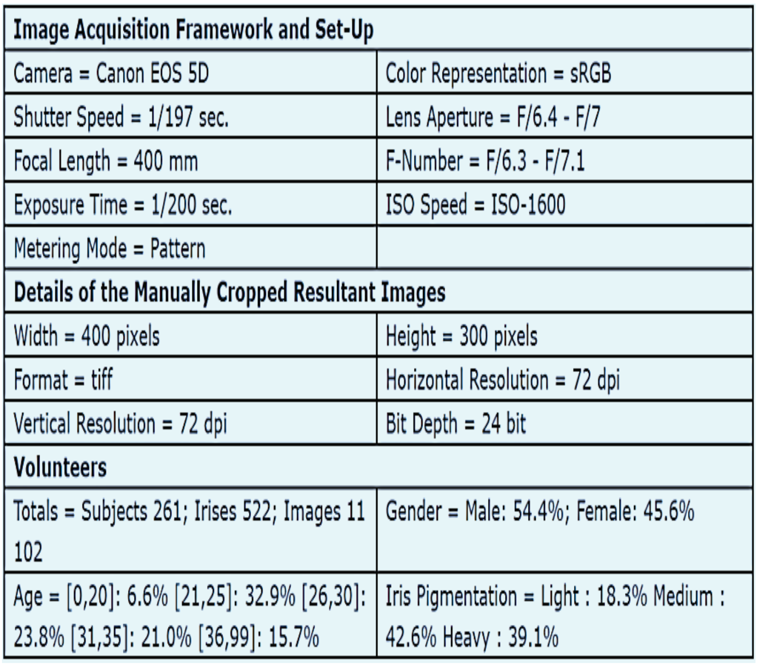


Image Acquisition Framework and Set-Up	
Camera = Canon EOS 5D	Color Representation = sRGB
Shutter Speed = 1/197 sec.	Lens Aperture = F/6.4 - F/7
Focal Length = 400 mm	F-Number = F/6.3 - F/7.1
Exposure Time = 1/200 sec.	ISO Speed = ISO-1600
Metering Mode = Pattern	
Details of the Manually Cropped Resultant Images	
Width = 400 pixels	Height = 300 pixels
Format = tiff	Horizontal Resolution = 72 dpi
Vertical Resolution = 72 dpi	Bit Depth = 24 bit
Volunteers	
Totals = Subjects 261; Irises 522; Images 11 102	Gender = Male: 54.4%; Female: 45.6%
Age = [0,20]: 6.6% [21,25]: 32.9% [26,30]: 23.8% [31,35]: 21.0% [36,99]: 15.7%	Iris Pigmentation = Light : 18.3% Medium : 42.6% Heavy : 39.1%

Figure 2.18: Imaging framework information for the UBIRIS.v2 database [112].

The framework setup, the acquisition device, and the number of subjects are shown in Fig. 2.18. Natural and artificial lighting conditions are utilised installed in a lounge and distance marks were fitted on the floor at distances from 4 to 8 meters with a 1-meter interval. The subject has to move slowly while looking at several lateral marks resulting in changing head pose and eye gaze directions. Two capturing sessions are conducted for 261 subjects collecting data for both eyes resulting in 11,102 images. Session 2 has a different location, luminosity and acquisition device to increase heterogeneity and only 221 subjects attended both sessions. For each subject, 3 images of 3 gaze directions (front, left, and right) were captured at each distance resulting in 15 images per subject per session. Examples are shown in Fig. 2.19.

Compared to the UBIRIS.v1 images, the dimension of the captured images in the UBIRIS.v2 is 400×300 pixels with 72 dpi captured on-

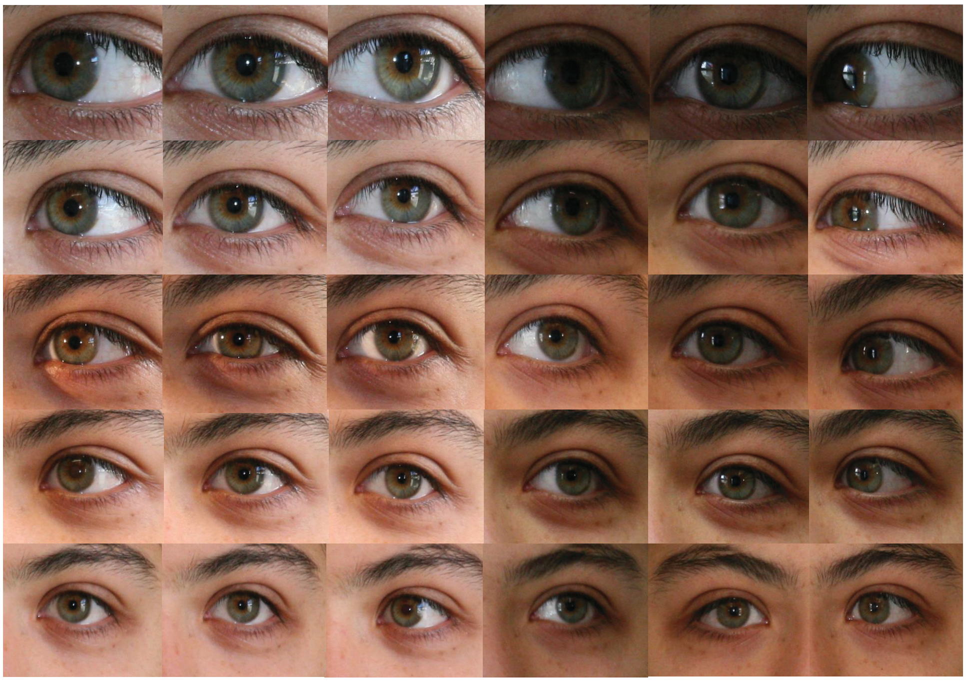


Figure 2.19: One subject image samples from the UBIRIS.v2 database captured at distances ranging from 4 to 8 meters in 2 sessions. The first three columns from the left represent session 1 whereas the next three columns are from session 2.

the-move and at-a-distance making these properties a distinguishable point for the UBIRIS.v2 database. However, no single research on the sclera has used the UBIRS.v2 database due to its challenging properties which means adding extra pre or post processing complexity to the system design. Some examples of the challenging cases of images in the UBIRIS.v2 database which are shown in Fig. 2.20 include multiple gaze directions, poorly focused eye, rotated eye image, blurred eye image, sclera obstruction due to eyelashes, sclera obstruction due to glasses, eye obstruction due to hair, poor lighting conditions, specular reflections, blinking, and no or insufficient sclera area.

2.4.3 *The UTIRIS database*

The UTIRIS database [113] is collected by a group of researchers to evaluate their work on iris recognition under visible wavelength illumination. The iris under visible illumination has the advantage of the visibility of the pigment melanin which provides a rich pattern

Figure 2.20: Examples of challenging eye images in the UBIRIS.v2 database including the off-angle sclera region, poor illumination, specular reflection, sclera obstruction by hair or glasses, and insufficient sclera area.

compared to NIR images. Therefore, to study the aforementioned phenomena, the UTIRIS database is collected under visible wavelength and NIR illumination.

For the acquisition of visible wavelength images, a Canon EOS 10D camera with macro lens is utilised whereas an ISG LightWise LW camera is used for acquiring NIR images. Images are captured from both eyes of 79 volunteers resulting in 158 classes with mostly front iris gaze direction and some deviated iris gaze directions. The visible wavelength images have a high dimension of 2048 × 1360. Daugman's method is used by authors in [113] to segment the iris resulting in high segmentation accuracy.

Although the UTIRIS visible wavelength images have a high resolution which means better segmentation and recognition accuracy due to more visible pattern, there is however an out of focus sclera region issue with some images. This is because of the necessity to extract a rich pattern from the iris during the capturing session resulting in high focus on the iris and blurred effect on the sclera using a macro lens. Therefore, a visual inspection is carried out to select images with non-blurred sclera region. Other issues in the UTIRIS database include rotated eye image, blinking, and poorly cropped image. Some examples of the high quality and blurred sclera region quality images are shown in Fig. 2.21.

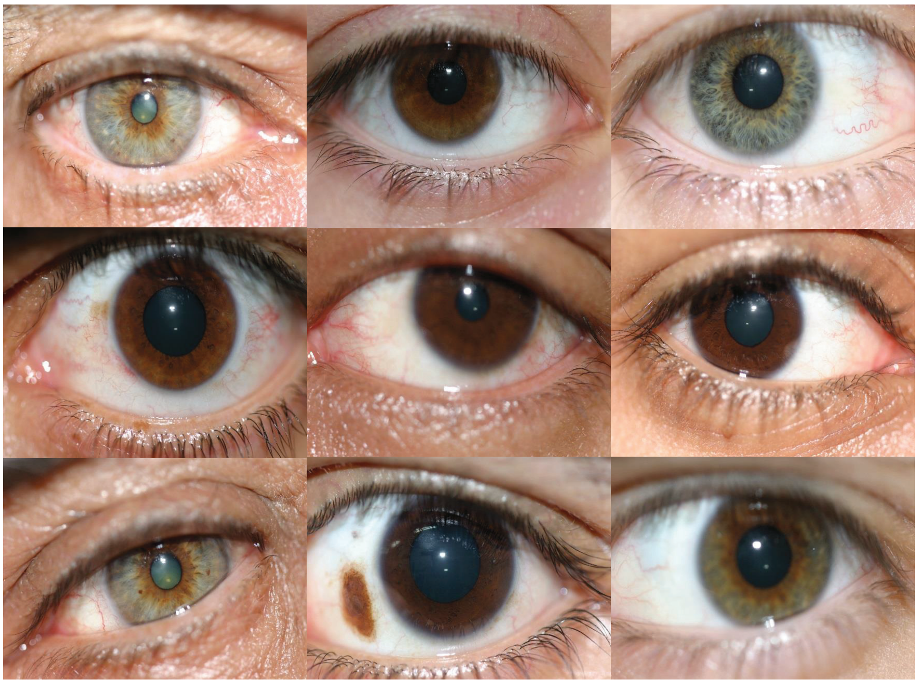


Figure 2.21: Examples of the UTIRIS database including high-quality images in the top row, different gaze directions images in the middle row, and poor quality sclera regions images in the bottom row.

2.4.4 *The MICHE-I database*

The Mobile Iris CHallenge Evaluation (MICHE)-I database was [114] established based on one goal: to evaluate iris recognition within an uncontrolled challenging environment represented by performing the capturing with a mobile device and its associated problems in real-time implementation. Some mobile device problems besides the typical noise factors in the aforementioned datasets involve illumination conditions, uncooperative subjects, varying capturing distance and unconstrained data collection. The uncooperative user scenario means that the particular user who captures eye images may not have the knowledge or expertise to ensure the captured image could achieve maximum segmentation accuracy. Hence, introducing a key issue to

Figure 2.22: Examples of the MICHE-I database images for 3 subjects captured by iPhone 5 cameras. The first row represents images captured using the anterior camera, the second row is for images from the posterior camera, the first three columns from the left captured in the indoor session, whereas the next three columns are captured in the outdoor session.

the eye biometric using a normal ubiquitous camera-based device which is the most prevalent case in daily use.

The acquisition framework protocol is user-dependent where the user is asked to act normally and capture eye images using a mobile phone (4 images per mobile or per camera if the mobile has anterior and posterior cameras) with indoor and outdoor sessions. The user for instance may keep his/her glasses or contact lenses during the capturing to simulate normal user engagement with a biometric application. Artificial lights combined with natural lights are used in the indoor session while only natural light is utilised for the outdoor session.

Three types of camera-based mobile devices are used for data acquisition which are the iPhone 5, Samsung Galaxy S4, and Samsung Galaxy II tablet. The iPhone 5 has a posterior camera with 8 MP 72 dpi and anterior camera with 1.3 MP 72 dpi whereas the Samsung Galaxy S4 has an anterior camera of 2 MP 72 dpi and posterior camera of 13 MP 72 dpi. In contrast, the Samsung Galaxy II tablet has only an anterior 0.3 MP low-resolution camera. The number of subjects who participated is 92 attending the first session and only 22 subjects out of 92 attending the second session resulting in a total number of images equal 3732. Examples of images captured from 3 subjects in indoor and outdoor sessions using the iPhone 5, Samsung Galaxy S4, and Samsung Galaxy II tablet cameras are shown in Figs. 2.22, 2.23, and 2.24 respectively.

Figure 2.23: Examples of the MICHE-I database images for 3 subjects captured by Samsung Galaxy S4 cameras. The first row represents images captured using the anterior camera, the second row is for images from the posterior camera, the first three columns from the left captured in the indoor session, whereas the next three columns are captured in the outdoor session.

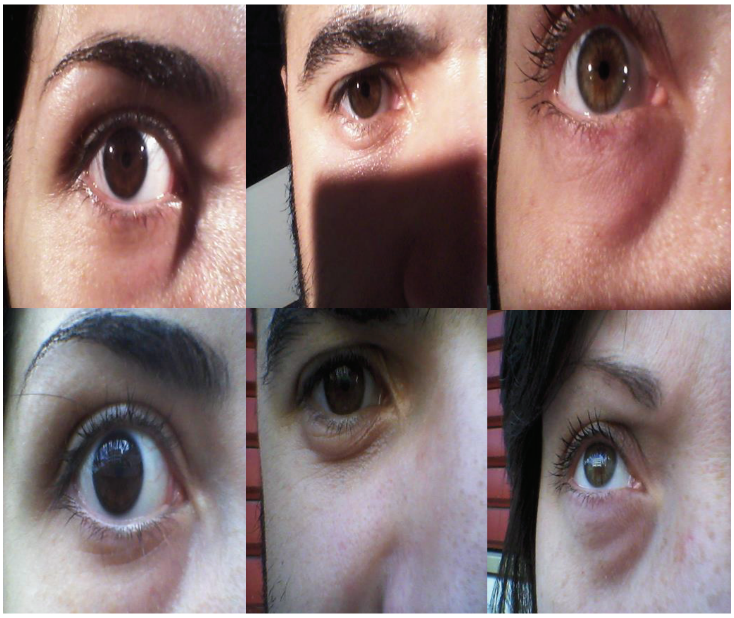


Figure 2.24: Examples of the MICHE-I database images for 3 subjects captured by Samsung Galaxy II tablet anterior camera. The first row represents images captured in the indoor session whereas the second row is for images from the outdoor session.

Different noise factors on the MICHE-I database in addition to the aforementioned challenges can be summarised as artificial and natural illumination sources, shadows, occlusions, blur, reflections, sclera obstruction due to eyelids, eyelashes, and hair, and finally low resolution for the anterior cameras of the three devices. Examples of noisy images from the MICHE-I database are shown in Fig. 2.25.

Figure 2.25: Examples of the MICHE-I database including high-quality images in the top row, different gaze directions images in the middle row, and poor quality sclera regions images in the bottom row.

2.4.5 *In-house database using HD Webcam*

As the proposed work in the thesis focuses mainly on sclera recognition under unconstrained imaging conditions, part of the proposed algorithm needed to be tested in a real-time scenario and hence an in-house database was collected locally in a lab area under natural room illumination. The acquisition device is a Logitech HD webcam c920 mounted on a tripod and the capturing distance is approximately 30 cm. The number of volunteers who participated was 6 using both eyes resulting in 12 eye classes. No constraints were forced on subjects as the user is required to sit in front of the camera and a video is recorded for 10 sec with 30 frames per sec capturing rate.

Examples of challenging issues in the captured frames are shown in Fig. 2.26 such as specular reflection, blurring, multi-angle sclera regions, unconstrained capturing distance, shadows, and the background noise. Due to the small number of subjects who volunteered,

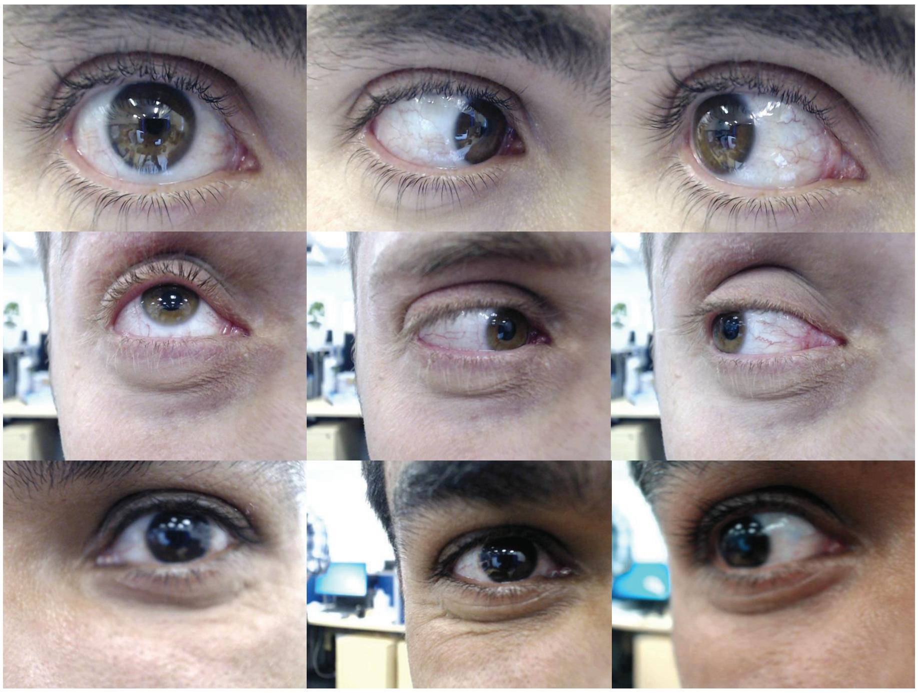


Figure 2.26: Examples of images for 3 subjects in the In-house HD Webcam database.

this database is only used to evaluate part of the proposed work in Chapter 5.

2.5 CHAPTER SUMMARY

In this chapter, an introduction to the biometric recognition system in terms of physical and behavioural traits, biometric characteristics, system design, recognition modes and performance evaluation was presented. Then, an overview of recent vascular-based recognition system was given. Next, the sclera recognition was explained and a literature survey was established on recent proposed work. Finally, a synopsis of the datasets which will be used in the proposed contributions in Chapters 3, 4, and 5 was introduced.

The next chapter will discuss briefly some issues in the recent work under controlled image capturing where noise factors are minimised and will propose a robust sclera recognition system. The new method includes a new sclera segmentation technique, an efficient sclera validation method through eye occlusion detection, blood vessel pattern enhancement, and novel template registration method.

3

ROBUST SCLERA RECOGNITION SYSTEM UNDER CONSTRAINED IMAGING CONDITIONS

3.1 INTRODUCTION

This chapter focuses on sclera recognition in applications where noise factors must be minimised. For instance, in indoor security applications, accessing a building or logging into a bank account, where the capturing process can be controlled and the illumination setting is fixed, specific system design is required. The design process can focus on improving the system performance, decreasing the complexity, and building an efficacious system which is user-friendly. Therefore, this chapter discusses the design of a robust sclera recognition system with a high accuracy. Normally, the sclera recognition system design consists of the following major steps which can be summarised as: **a)** sclera segmentation; **b)** blood vessel enhancement and feature extraction; and **c)** training the recognition system and matching procedure.

To mitigate limitation points (1-5) mentioned in subsection 2.3.2, the major contributions of this chapter are

- Occluded eye image detection;
- Adaptive sclera shape contour segmentation based on active contours without edges;
- Efficient image enhancement and vessel map extraction;
- Robust sclera feature extraction with template registration and matching.

The system comprises new sclera segmentation and occluded eye detection methods. In addition, an efficient method for vessel enhancement, extraction and binarization is proposed. In the feature extraction and matching process stages, an efficient method is additionally developed that is orientation, scale, illumination and deformation invariant. The obtained results using the UBIRIS.v1 [96] and UTIRIS [113] databases show an advantage in terms of segmentation accuracy and computational complexity compared with state-of-art methods.

3.2 SCLERA SEGMENTATION DESIGN

Sclera segmentation is the initial and the most challenging step in a sclera recognition system. The accuracy of the sclera recognition system could be degraded if the segmentation process fails to extract the correct sclera regions from an eye image. Some incorrect sclera segmentation scenarios include segmenting the sclera with some parts of the iris, eyelids and eyelashes. Table 2.2 shows sclera segmentation techniques developed from both manual and automated segmentation processes. In general, two strategies, sclera pixel thresholding and sclera shape contour, have been adopted for automated and semi-automated sclera segmentation each having their advantages and disadvantages. The sclera shape contour technique is the focus in this chapter as the sclera pixel thresholding involves multiple steps to remove noise and segmenting the sclera region, thus increasing the complexity and processing time which will be discussed in the next chapter. An adaptive approach for unsupervised fully automated sclera segmentation is proposed by using active contours without edges [115] with a novel occluded eye detection method.

3.2.1 *Iris Segmentation*

Iris centre estimation has an essential role in the proposed sclera segmentation method, although the sclera recognition system does not depend directly on the iris for the system implementation. However, locating the position of the iris centre within the eye image plays a crucial part in the proposed sclera segmentation. There is a significant amount of literature on iris segmentation [32, 116, 117, 118, 119] for which the iris region is modelled as having circular boundaries. However, it is not the research focus in this chapter to improve these methods but rather to extract the iris centre and radius. The integro-differential operator suggested by Daugman [32] which acts as a circular edge detector is used. To find the parameters (r, x_0, y_0), the following optimisation which maximises these parameters and uses the integro-differential operator is defined as

$$\arg\max_{(r,x_0,y_0)} \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r,x_0,y_0} \frac{I(x,y)}{2\pi r} ds \right|, \tag{3.1}$$

where $I(x,y)$ is the greyscale level of an eye image, (r,x_0,y_0) are the iris radius and centre coordinates, the symbol $*$ is the convolution operator, and $G_\sigma(r)$ is a Gaussian smoothing function of scale σ. The coloured image is first converted into greyscale format and down-sampled by factor of 0.25 to enhance the processing time. This process is shown in Fig. 3.1.

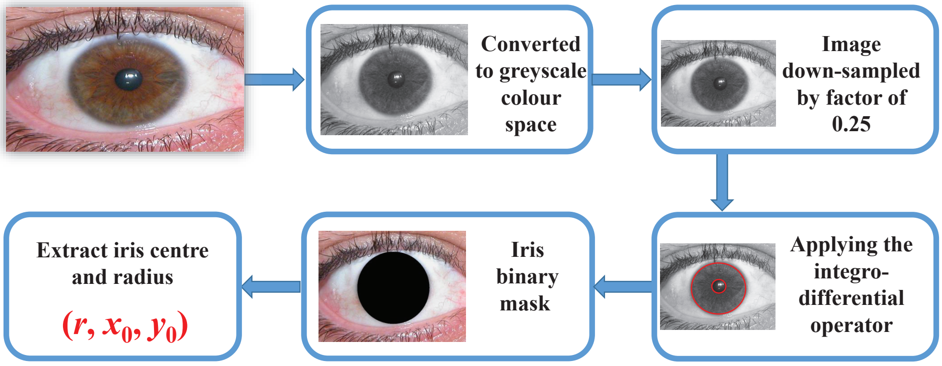


Figure 3.1: Iris segmentation process.

3.2.2 *Occluded Eye Detection for Sclera Validation*

A major challenge in the sclera segmentation process is the sclera validation where the segmented sclera is automatically verified without any supervision from the human. There are some factors such as an occluded eye image and small sclera area which will affect the automatic process and increase error rate. Many researchers have discarded these images manually or applied sclera validation after the sclera segmentation step. However, if the sclera image has an occluded eye, then validating the sclera after segmentation is inefficient. Therefore, a sclera validation process is proposed by isolating the poor samples in the enrolment and verification stages. The proposed method is applied before sclera segmentation to detect a partial or fully occluded eye image, to make a validation decision of a sufficient sclera region.

Based on the iris radius and centre coordinates (r, x_0, y_0), two arc areas of intensities are specified according to the following

$$\mathbf{Arc} = \mathrm{RGB}\left(x_0 + r\cos\theta, y_0 + r\sin\theta\right), \tag{3.2}$$

where $\theta \in [-\pi/3 : \pi/3] \cup [-2\pi/3 : 2\pi/3]$ with uniform increment steps of 0.1 degrees and is set in this range to check the status where eyelids are partially closed. Then, the RGB intensities of each pixel in these two arc areas have been classified by heuristic rules into skin or non-skin labels using the Colour Distance Map (CDM) proposed in [120]. First, the classification method will be illustrated by applying it on the right arc $\mathbf{Arc}_r$ and the same operation will be applied on the left arc $\mathbf{Arc}_l$. First, two skin clusters for natural and flash illumination conditions are defined as

$$CDM_1 = \begin{cases} & R > 95, G > 40, B > 20, \\ 1, & \max\left(R, G, B\right) - \min\left(R, G, B\right) > 15, \\ & |R - G| > 15, R > G, R > B \\ 0, & \text{otherwise} \end{cases} \tag{3.3}$$

$$CDM_2 = \begin{cases} & R > 220, G > 210, B > 170, \\ 1, & |R - G| \leqslant 15, B < R, B > G \ \cdot \\ 0, & \text{otherwise} \end{cases} \tag{3.4}$$

Then, the non-skin map is created based on these clusters as

$$\mathbf{s}_r(i) = \begin{cases} 1, & \text{if } CDM_1(i) \parallel CDM_2(i) = 0 \\ 0, & \text{otherwise} \end{cases}, \tag{3.5}$$

where the symbol $\parallel$ refers to the logical OR operator, 1 refers to a non-skin pixel and 0 to a skin pixel. The same operation will be repeated on $\mathbf{Arc}_l$ to produce $\mathbf{s}_l$. Next, from the $\mathbf{s}_r$ and $\mathbf{s}_l$ vectors, four sub-arcs

$\mathbf{up}_r$ with $\Theta \in [\pi/6 : \pi/3]$, $\mathbf{up}_l$ with $\Theta \in [5\pi/6 : 2\pi/3]$, $\mathbf{down}_r$ with $\Theta \in [-\pi/6 : -\pi/3]$ and $\mathbf{down}_l$ with $\Theta \in [-5\pi/6 : -2\pi/3]$ are used to define decision parameters for the non-skin pixels as

$$\mathbf{nsk}_{up}(i) = \mathbf{up}_r(i) \bigcap \mathbf{up}_l(i), \tag{3.6}$$

$$\mathbf{nsk}_{down}(i) = \mathbf{down}_r(i) \bigcap \mathbf{down}_l(i), \tag{3.7}$$

where $\bigcap$ is the logical AND operator to achieve the intersection between the two masks, and Θ is set within this range empirically to enhance the processing time. Then, the percentage of the non-skin pixels is calculated for both $\mathbf{nsk}_{up}$ and $\mathbf{nsk}_{down}$ as

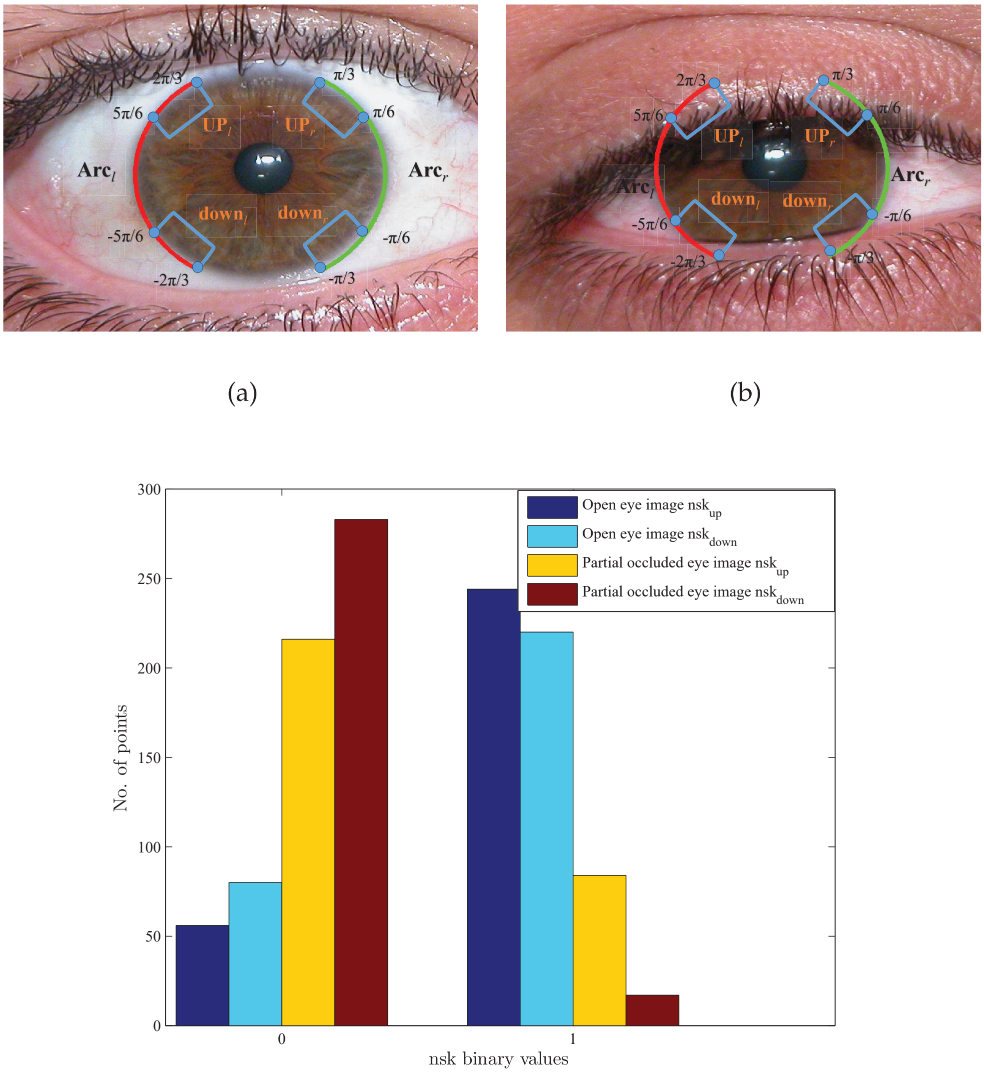
$$per_{\mathbf{nsk}_{up}} = \frac{\text{No. of 1s } (\mathbf{nsk}_{up})}{\text{Total No. of elements } (\mathbf{nsk}_{up})}, \tag{3.8}$$

$$per_{\mathbf{nsk}_{down}} = \frac{\text{No. of 1s } (\mathbf{nsk}_{down})}{\text{Total No. of elements } (\mathbf{nsk}_{down})}. \tag{3.9}$$

According to $per_{\mathbf{nsk}_{up}}$ and $per_{\mathbf{nsk}_{down}}$, the final decision to accept or reject the eye image is defined as

$$\text{Eye image} = \begin{cases} \text{Accept}, & \text{if } per_{\mathbf{nsk}_{up}} \geqslant 0.6 \bigcap per_{\mathbf{nsk}_{down}} \geqslant 0.3 \\ \text{Reject}, & \text{otherwise} \end{cases}, \tag{3.10}$$

where 0.6 and 0.3 are set empirically for optimum performance. The reason that the $per_{\mathbf{nsk}_{up}}$ value is set higher than $per_{\mathbf{nsk}_{down}}$ is that the upper eyelid physiologically is more mobile and contributes more for the eye closure process. After that, if the eye image is validated for a sufficient sclera area, then the iris binary mask is applied to segment the iris and the image is ready for the sclera segmentation. Otherwise, no further processing time is required and the image is rejected. The proposed validation process for an ideal eye and partial occluded eye images is shown in Fig. 3.2.



(a)

(b)

(c)

Figure 3.2: Proposed occluded and partially occluded eye image detection, (a) and (b) are an ideal eye image and partial occluded eye image respectively with the angles depicted where the four sub-arcs $\mathbf{up}_r$, $\mathbf{up}_l$, $\mathbf{down}_r$, and $\mathbf{down}_l$ are calculated, (c) the histograms of skin and non-skin pixels in the $\mathbf{nsk}_{up}$ and $\mathbf{nsk}_{down}$ vectors.

3.2.3 *Sclera Segmentation*

3.2.3.1 *Seed-based Initialization of Contours*

For a frontal-looking iris, two initial seeds for the left and right contours are initialised depending on the iris radius and centre coordinates (r, x_0, y_0). The centre positions of these seeds are calculated as

$$C_{rs}\left(x, y\right) = \left(x_0 + \left(1.35 \times r\right), y_0\right), \tag{3.11}$$

and

$$C_{ls}(x,y) = \big(x_0 - (1.35 \times r), y_0\big), \tag{3.12}$$

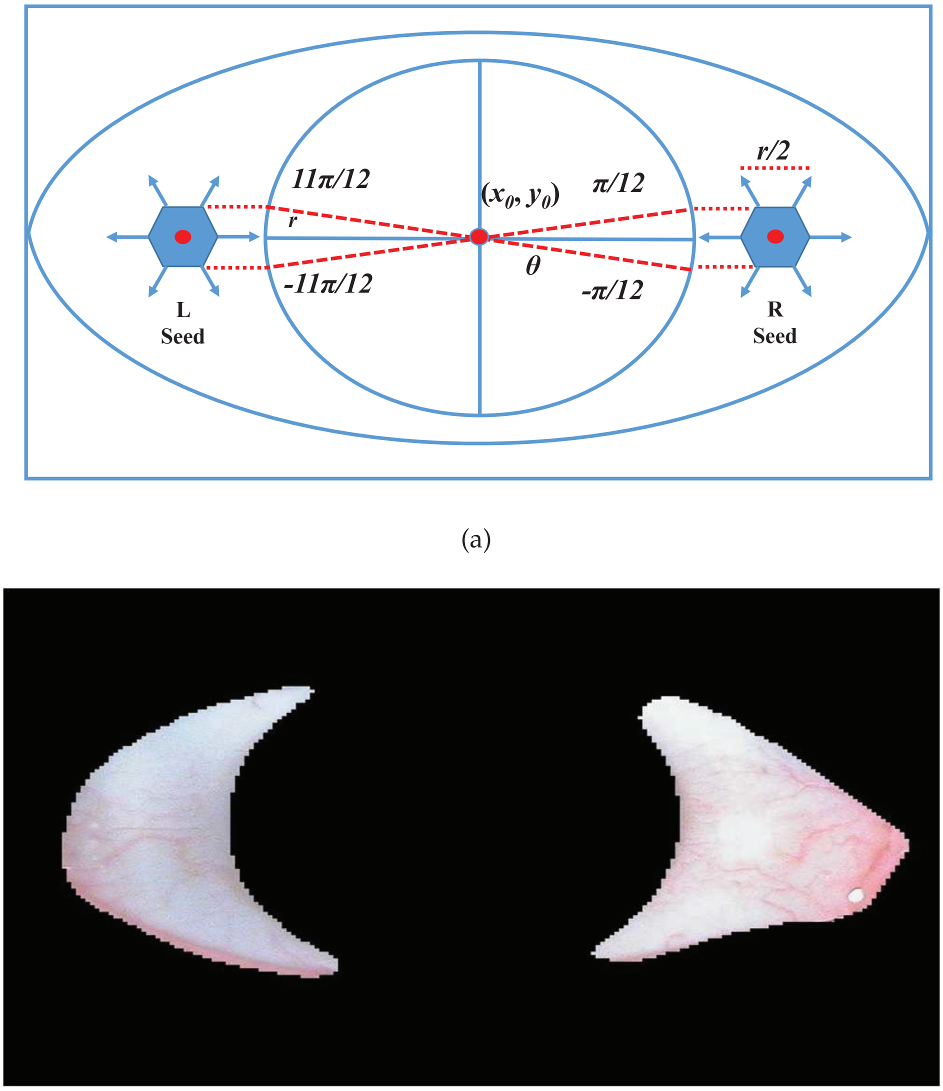
where $1.35 \times r$ is set to ensure that the right and left seed centres are outside the iris. The height of the initial contours is set with altitude from the point on the iris circumference with $\theta = \pi/12$ to the point with $\theta = -\pi/12$ and the width $(r/2)$ to make the contours converge inside the sclera regions. The initial seed positions are depicted in Fig. 3.3a.

3.2.3.2 *Active Contours Without Edges*

The basic idea in snakes or active contour models is to develop a curve subject to image forces in order to detect salient objects [121, 122, 123, 124]. For instance, a curve is initialised around an object to be detected and the curve will move towards that object until its boundary is detected. However, a major problem with these active contour models is that they have to rely on an edge function which depends on the image gradient. These models can detect only objects with edges defined by the gradients that stop the curve evolution. In practice, the discrete gradients are bounded. Therefore, the stopping function is never zero on the edges and the curve may pass through the boundary. Another problem is that if the image is contaminated with noise, it will require a strong isotropic smoothing Gaussian function to remove the noise at the cost of smoothing the edges as well.

In contrast, the active contours without edges model [115] has no stopping edge function i.e., the gradient of the image is not adopted for the stopping process, instead the Mumford-Shah segmentation techniques are adopted [125]. The advantage of this model is the ability to detect contours independent of gradients, for instance, objects with very smooth boundaries or with discontinuous boundaries.

Summarising the development in [115], let an eye image in the greyscale I_0 be an open subset of $\mathbb{R}^2$ with its boundary equals ∂I_0, and the evolving curve C in I_0 representing the boundary of sclera ω of I_0 i.e., $\omega \subset I_0$ and $C = \partial\omega$. Then, the terms $\mathtt{inside}(C)$ and $\mathtt{outside}(C)$ will equal the area regions of ω and $I_0 \backslash \bar{\omega}$ respectively, where the symbol $(\backslash)$ denotes removal of the ensuing term and $\bar{\omega}$ is the



(a)

(b)

Figure 3.3: Segmentation process for the sclera, (a) initial seeds for the right and left contours for iris-centred eye, (b) sclera segmented template.

closed region of ω. Assuming that I_0 has two regions of approximately piecewise-constant intensities I_0^i and I_0^o, then the sclera to be detected is the region with value I_0^i. The fitting term of the active contour model is represented by the energy function $F(c_1, c_2, C)$ which is minimised if the curve C is on the boundary of ω and defined as

$$\begin{aligned} F(c_1, c_2, C) = F_1(C) + F_2(C) + \mu.\text{Length}(C) \\ + \nu.\text{Area}(\text{inside}(C)), \end{aligned} \tag{3.13}$$

where $F_1(C)$ and $F_2(C)$ equal

$$F_1(C) = \lambda_1 \int_{\text{inside}(C)} |I_0(x,y) - c_1|^2 \, dxdy, \tag{3.14}$$

$$F_2(C) = \lambda_2 \int_{\text{outside}(C)} |I_0(x,y) - c_2|^2 \, dxdy, \tag{3.15}$$

where $\mu \geqslant 0$, $\nu \geqslant 0$, $\lambda_1, \lambda_2 > 0$ are fixed parameters, and c_1, c_2, are the averages of I_0 inside and outside C respectively. Further details of the minimising procedure are in [115]. The parameters $\lambda_1, \lambda_2 = 1$ and $\nu = 0$ are fixed for best performance whereas the time step Δt is set to 0.1. μ is the controlling parameter of the evolving curve C. The smaller the value of μ, the more ability to detect as many objects as possible with different sizes. As μ increases, the curve C will detect only large objects. In this research, μ is set empirically to high value equal to 0.2 in order to not detect the blood veins inside the sclera rather the sclera boundary.

The values of $F_1(C)$ and $F_2(C)$ depend on the curve position in the eye image. As shown in Fig. 3.4, positioning the initial seeds outside the sclera area or between the sclera area and other eye parts could cause the curve to converge incorrectly on other eye parts. Therefore, it is more efficient design to initiate these seeds inside the sclera as $F_1(C)$ is already zero.

Then, the iris-segmented image is converted to the blue channel and down-sampled by factor = 0.2 to enhance processing time and the final sclera binary map is created. The initial contours will converge towards the sclera boundaries in all direction and will stop for a number of iterations $i = 200$. The final sclera template is shown in Fig. 3.3b.

3.3 SCLERA BLOOD VESSEL ENHANCEMENT AND MAPPING

The main purpose of the vessel enhancement is to isolate the blood vessels in the sclera from their background. This process has two stages. In the first stage, the green layer of the RGB image is extracted as it leads to better contrast between the sclera blood vessel and the

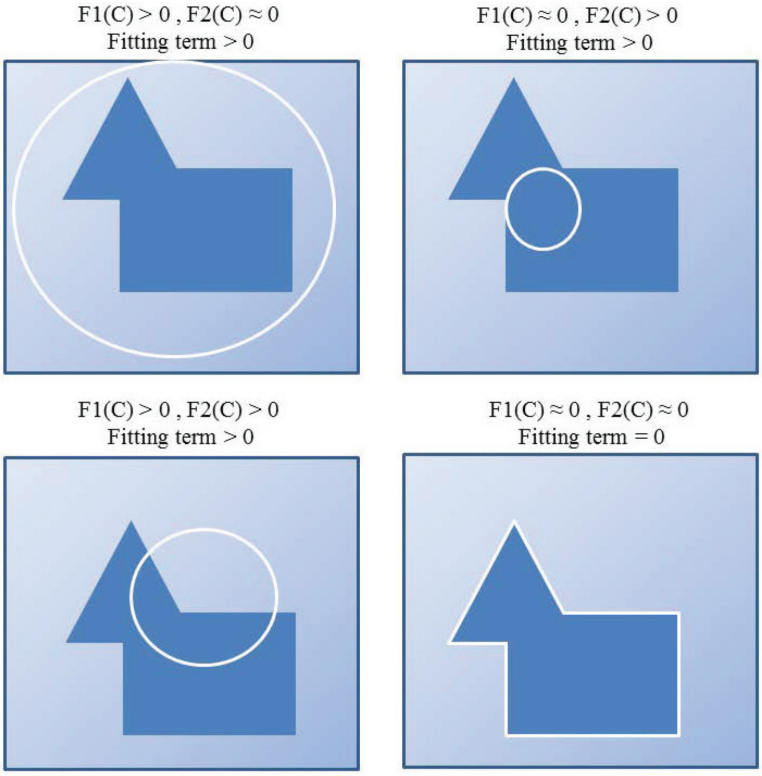


Figure 3.4: All the possible cases of the initial curve and their $F_1(C)$ and $F_2(C)$ values. Note that the fitting term is equal zero when the curve is on the boundary of the object.

background [70]. Then, CLAHE was applied to sclera regions as it will enhance the green layer of the coloured image [17].

For the vessel extraction, which is the second stage, the two-dimension Isotropic Undecimated Wavelet Transform (IUWT) is proposed as it is robust and well adapted to astronomical data and biology images where objects are more or less isotropic [126]. To extract $\mathcal{W} = \{w_1, \cdots, w_J, c_J\}$, where $w_{j=1,\dots,J}$ are the wavelet coefficients at scale j and c_J are the coefficients at the coarsest resolution, a subtraction between two adjacent sets of coefficient scales is applied as

$$w_{j+1}[k,l] = c_j[k,l] - c_{j+1}[k,l], \tag{3.16}$$

where

$$c_{j+1}[k,l] = \left(\bar{h}^{(j)}\bar{h}^{(j)} * c_j\right)[k,l], \tag{3.17}$$

and $\bar{h}^{(j)}[k] = \frac{[1,4,6,4,1]}{16}$ is the non-orthogonal Astro filter bank with $k = \{-2, \dots, 2\}$. At each scale j, one wavelet set is obtained which has the same resolution as the sclera image. This feature solves the dimensionality increment introduced by using Gabor filters with different scales and orientations and produces efficient processing time. The reconstruction process is obtained by co-addition of j wavelet scales as

$$c_0[k,l] = c_J[k,l] + \sum_{j=1}^{J} w_j[k,l]\,. \tag{3.18}$$

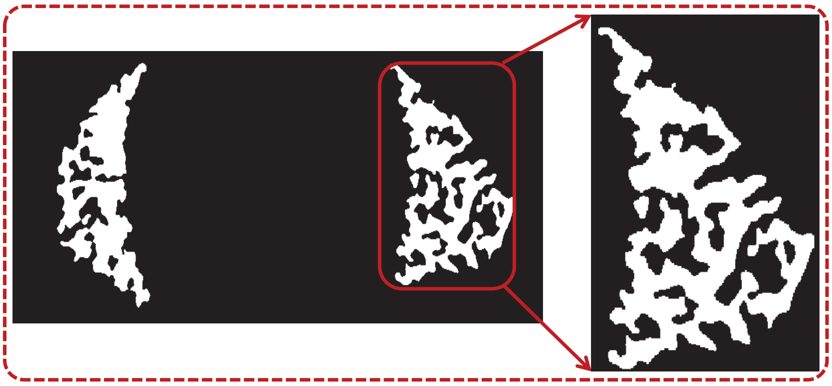
The segmentation process for the vessels can be initiated simply by adding the best wavelet levels with a thresholding process that represents the best vessel contrast. The threshold to detect the vessels is empirically set to identify the lowest 30% of the coefficients. It is likely to misclassify non-vessel pixels as vessel pixels. However, a cleaning process can be achieved simply by calculating the area of misclassified pixels and a threshold can be set to remove these undesired pixels. This is shown in Fig. 3.5a

The thickness variation in the sclera vessels due to the physiological status of a person [31] affects the recognition process. Therefore, these vessels must be transformed to a single pixel skeleton map. For the thinning process, a morphological thinning algorithm was applied and binary morphological operations are applied to remove the exterior pixels from the detected vascular vessels map and to create a one-pixel skeleton running along the centre of the vessels. The binary skeleton map of the sclera vessels is shown in Fig. 3.5b.

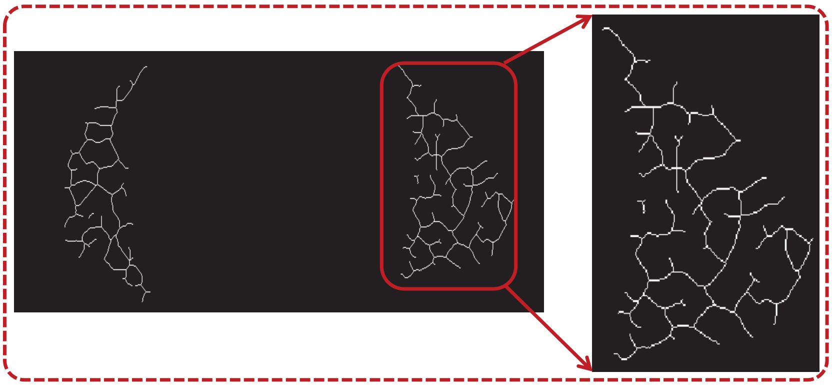
3.4 SCLERA FEATURE EXTRACTION AND TEMPLATE REGISTRATION

3.4.1 *Sclera Feature Extraction*

The feature extraction process in the sclera recognition system involves producing a reliable mathematical model in order to identify individuals. A new method for the sclera pattern feature extraction is proposed based on Harris corner and edge detection [110]. This algorithm detects the Interest-Points (IPs) represented by the corner response R, where some of the vessel corner formations can supply a significant 2D texture for pattern recognition.



(a)



(b)

Figure 3.5: The sclera blood vessels enhancement and mapping, (a) applying IUWT for the vessel extraction, (b) the morphological thinning process.

The corner response R is defined as

$$R = \det(A) - k\,\mathrm{tr}^2(A), \tag{3.19}$$

where det is the determinant of matrix A, k is the corner sensitivity factor set empirically to 0.04, tr is the matrix trace and A is the image structure matrix computed from the image derivatives as

$$A = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}, \tag{3.20}$$

where I_x and I_y are the partial derivatives in x and y respectively. The corner response value is positive when a corner region is detected,

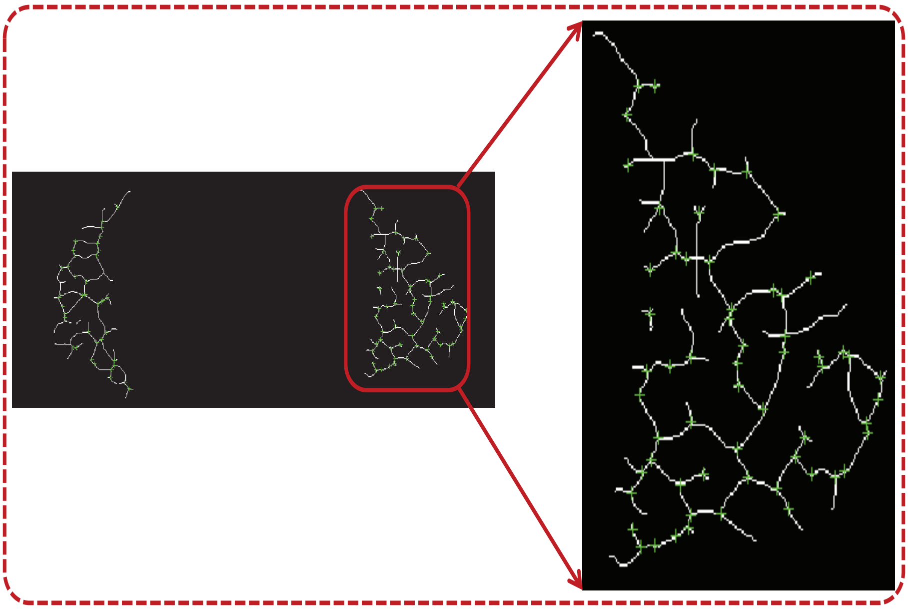


Figure 3.6: Applying Harris detector to extract IP features.

small if a flat region is detected or negative if an edge region is detected within the smooth circular Gaussian window. The steps for the IPs extraction process are as follows

i. compute the 1st derivatives in x and y;

ii. apply the Gaussian smoothing filter to remove the noisy response due to a binary window function;

iii. find the points with a large corner response function R when $(R > th_C)$;

iv. take the points of the local maxima of R.

where th_C is set to 0.01 for best performance. The extracted IPs in the sclera vessel map are shown in Fig. 3.6.

After the IP locations have been extracted, the characteristic information such as the magnitude, phase and orientation within a specific window surrounding each feature point are calculated. The two-dimensional monogenic signal method [127] is exploited to extract the local information of these IPs. An analytic signal is constructed by using the Riesz transform. The analytic signal is isotropic and therefore provides a split of identity where the information is orthogonally decomposed into geometric and energetic information for the blood veins IPs in term of local amplitude, local phase and local orientation. If the coordinates of the vertical and horizontal filters are $\mathbf{u} = (u_1, u_2)$, then the Riesz kernels in the frequency domain are multiplied with a Log-Gabor scale mask to calculate the spatial representation of the vertical and horizontal filters respectively as

$$H_1(\mathbf{u}) = i\frac{u_1}{|\mathbf{u}|}G(|\mathbf{u}|), H_2(\mathbf{u}) = i\frac{u_2}{|\mathbf{u}|}G(|\mathbf{u}|), \tag{3.21}$$

where $|\mathbf{u}| = \sqrt{u_1^2 + u_2^2}$ represents the radius of frequency values from the centre and $G(|\mathbf{u}|)$ is the monogenic scale space defined by using the Log-Gabor filter for a wavelength $1/f_0$ as

$$G(|\mathbf{u}|) = \exp\left(-\frac{(\log(|\mathbf{u}|/f_0))^2}{2(\log(k/f_0))^2}\right), \tag{3.22}$$

where $1/f_0 = 0.1$ and $k = 0.2$, which are set empirically, determine the ratio of the standard deviation of the Gaussian describing the log-Gabor filter's transfer function in the frequency domain to the filter centre frequency. If $f(I)$ represents the input image in the frequency domain, then the monogenic signal of input image $f_M(I)$ can be defined as

$$f_M(I) = f(I) + H(\mathbf{u}) * f(I), \tag{3.23}$$

where $*$ is the convolution operator. For the monogenic signal representation in the image domain, 3-tuple variable $\{p(I), q_1(I), q_2(I)\}$ is defined as

$$\begin{aligned} p(I) &= \mathbb{F}^{-1}(f(I)\ G_a(I)) \\ q_1(I) &= \mathbb{F}^{-1}(f(I)\ H_1(I)), \\ q_2(I) &= \mathbb{F}^{-1}(f(I)\ H_2(I)) \end{aligned} \tag{3.24}$$

where $\mathbb{F}^{-1}$ is the inverse Fourier transform, and G_a is the Log-Gabor filter of the image domain. The amplitude information is ignored as the vessel image is a binary form whereas the monogenic phase $\varphi(I)$, which will be in range $-\pi/2 \leqslant \varphi \leqslant \pi/2$, is calculated as

$$\varphi(I) = \tan^{-1} \frac{p(I)}{\sqrt{q_1(I)^2 + q_2(I)^2}}. \tag{3.25}$$

Finally, for each IP location, a square window patch with size equals $w^2 = 19 \times 19$ is stored along with the analytic signal information and the sclera template will consist of the following components

$$\mathbf{St} = \{\text{IP.locations}, \text{IP.phases}\}. \tag{3.26}$$

3.4.2 *Sclera Feature Template Registration*

The investigation of non-linear blood vein movement in the sclera carried out by Zhou et al. [31] has shown that these vessels move slightly as the eye moves. To overcome this limitation and produce an invariant-blood movement user template, a new method for user template registration is proposed. This method is initiated with the alignment of user sclera templates to a reference point. For three sclera training templates $\mathbf{St}_1$, $\mathbf{St}_2$ and $\mathbf{St}_3$, for instance, the reference points are $rf_1(x, y)$, $rf_2(x, y)$ and $rf_3(x, y)$ which represent the centre of the iris. Then, these templates are aligned as follows

i. let $rf_1(x, y)$ be the point where $\mathbf{St}_2$ and $\mathbf{St}_3$ are to be aligned;

ii. for $\mathbf{St}_2$, subtract ($rf_1(x, y)$,$rf_2(x, y)$) to extract the shift values (hr,vr) in the horizontal and vertical dimensions;

iii. apply a circular shift where $\mathbf{St}_2$ is shifted in both directions as $\mathbf{St}_2(x + hr, y + vr)$;

iv. repeat for $\mathbf{St}_3$.

If (hr,vr) are positive, $\mathbf{St}_2$ is shifted to the right and to the bottom, if (hr,vr) are negative, $\mathbf{St}_2$ is shifted to the left and to the top, if (hr) is positive and (vr) is negative, $\mathbf{St}_2$ is shifted to the right and to the top, and if (hr) is negative and (vr) is positive, $\mathbf{St}_2$ is shifted to the left and to the bottom.

Next, a new method for feature matching based on a local descriptor is proposed to generate putative matching pairs $\mathbf{p_m}$ between these sclera feature sets, and RANSAC [128] is used to select the homogeneous feature sets $\mathbf{f_{in}}$. RANSAC uses the minimum number of feature points to estimate the initial feature sets and proceeds to maximise the number of pairs within the set with consistent data points and thus decreases the processing time. The proposed local descriptor includes generating a correlation matrix between two sclera templates $(\mathbf{St}_1, \mathbf{St}_2)$. The correlation matrix holds the correlation strength of every feature set in $\mathbf{St}_1$ relative to $\mathbf{St}_2$. Then, search for $\mathbf{p_{m1}}$ pairs that correlate maximally in both directions; $\mathbf{p_{m1}}$ can be calculated as

i. set a radius for the correlation window $rc = \frac{w-1}{2}$;

ii. calculate the distances of $\mathbf{St}_1$.IP.location(i) to all $\mathbf{St}_2$.IP.locations;

iii. specify the pairs of points that are separated by a Euclidean distance $< d_{max}$;

iv. normalise the phase information of the selected pairs of points window to a unit vector form and measure the correlation using a dot product;

v. find $\mathbf{p_{m1}} = \arg\{\max|corr(\mathbf{St}_1, \mathbf{St}_2)\}$.

where d_{max} is set empirically to 50 for best performance.

Once the $\mathbf{p_{m1}}$ pairs for these templates were specified, then the process of finding $\mathbf{f_{in1}}$ for $\mathbf{p_{m1}}$ can be defined as follows

i. select randomly the minimum number of feature points $m = \arg\{\min(\mathbf{p_{m1}})\}$;

ii. normalise each set of points so that the origin is at the centroid and the mean distance from origin is $\sqrt{2}$;

iii. calculate Sampson error distances [129] between these sets and determine how many feature points from the templates fit with predefined Sampson error tolerance $< \epsilon$, where ϵ is empirically set to 0.002;

iv. if the ratio of the number of $\mathbf{f_{in1}}$ over the total number of $\mathbf{p_{m1}}$ exceeds a predefined threshold b, re-estimate the model parameters using all the identified $\mathbf{f_{in1}}$ and terminate;

v. otherwise, repeat steps i through iv for a maximum of N iterations.

The decision to stop selecting new feature subsets is based on the number of iterations N required to ensure that the probability $z = 0.99$ that choosing random samples has at least one set does not include $\mathbf{f_{out}}$ points, which are the non-homogeneous feature points. N can be set as

$$N = \frac{\log(1-z)}{\log(1-b)}, \tag{3.27}$$

where $b = 1 - \left(\frac{\text{No. of } \mathbf{f_{in}}}{\text{No. of } \mathbf{p_m}}\right)^s$ represents the probability that any selected feature points is a homogeneous feature points $\mathbf{f_{in1}}$ and $s = 8$ is the number of points needed to fit a fundamental matrix [128]. This method uses a minimum number of feature points to estimate the homogeneous pairs set and maximises this set with consistent feature points. After $\mathbf{f_{in1}}$, $\mathbf{f_{in2}}$ and $\mathbf{f_{in3}}$ have been extracted and grouped in one set, the final sclera user template $\mathbf{St}^f$ is created by removing the duplicated points. This is shown in Fig. 3.7. The angles ϕ between $\mathbf{St}^f$ points and the iris centre have been calculated for each point which will be used in the matching process to mitigate some outlier points being paired incorrectly.

3.5 SCLERA MATCHING AND DECISION PROCESS

The matching and decision process between two sclera templates is the final stage where the $\mathbf{St}^f$ of an enrolled user is compared in terms of local orientation and local phase with $\mathbf{St}^t$ of any test template. First, the putative feature sets $\mathbf{p_m}$ are extracted from $\mathbf{St}^f$ and $\mathbf{St}^t$. Then these pair sets are compared in term of the location and the angle to the iris centre For the decision process, a two stage decision method is proposed by using the Euclidean distance and the angle difference to

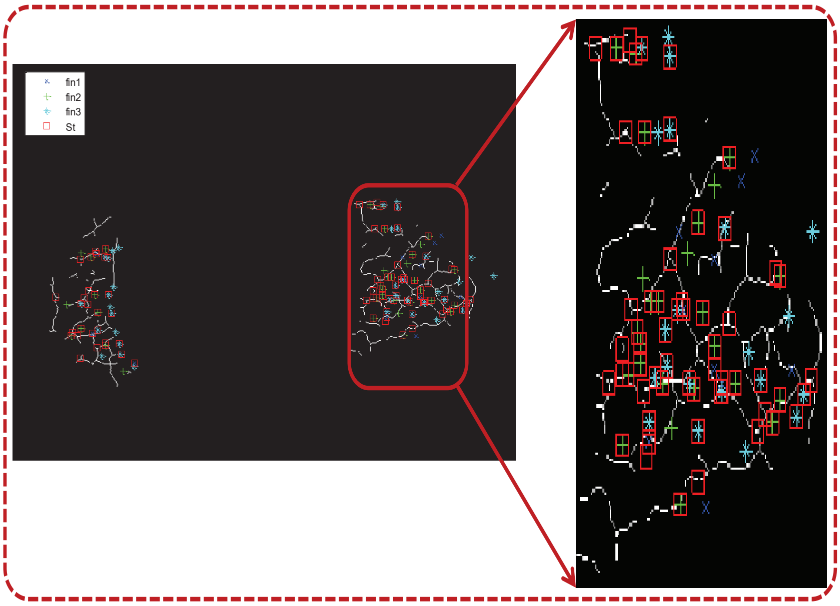


Figure 3.7: The most homogeneous feature sets $\mathbf{St}^f$ represented in the red square extracted from $\mathbf{f}_{\mathbf{in}1}$, $\mathbf{f}_{\mathbf{in}2}$ and $\mathbf{f}_{\mathbf{in}3}$.

classify the matching results. For each pair, two parameters are defined as

$$D_p = \begin{cases} 1, & \text{if } \operatorname{dist}\left(\mathbf{St}^f, \mathbf{St}^t\right) \leqslant th_p \\ 0, & \text{otherwise} \end{cases}, \tag{3.28}$$

$$D_\phi = \begin{cases} 1, & \text{if } |\phi_f - \phi_t| \leqslant th_\phi \\ 0, & \text{otherwise} \end{cases}, \tag{3.29}$$

where $th_p = 20$ and $th_\phi = 5$ degrees are set empirically for best performance. The decision to accept or reject each pair is shown in Algorithm 3.1.

Finally, the decision process is calculated as

$$D_f = \begin{cases} \text{Accept}, & \text{if } \frac{\text{no. of accepted pairs}}{\text{no. of } \mathbf{p_m}} \geqslant th_f \\ \text{Reject}, & \text{otherwise} \end{cases}, \tag{3.30}$$

where th_f is set empirically to 0.6.

Some examples of the initial matching process between two sclera templates are shown in Fig. 3.8, whereas examples of the final matching result are shown in Fig. 3.9. The matching between two sclera templates belonging to the same individual are shown in Fig. 3.8a and 3.9a, whereas Fig. 3.8b and 3.9b show the matching between two sclera templates belonging to different individuals. D_f is dramatically higher ($D_f = 89.51\%$) when matching two sclera templates for the same user than matching different user templates ($D_f = 49\%$).

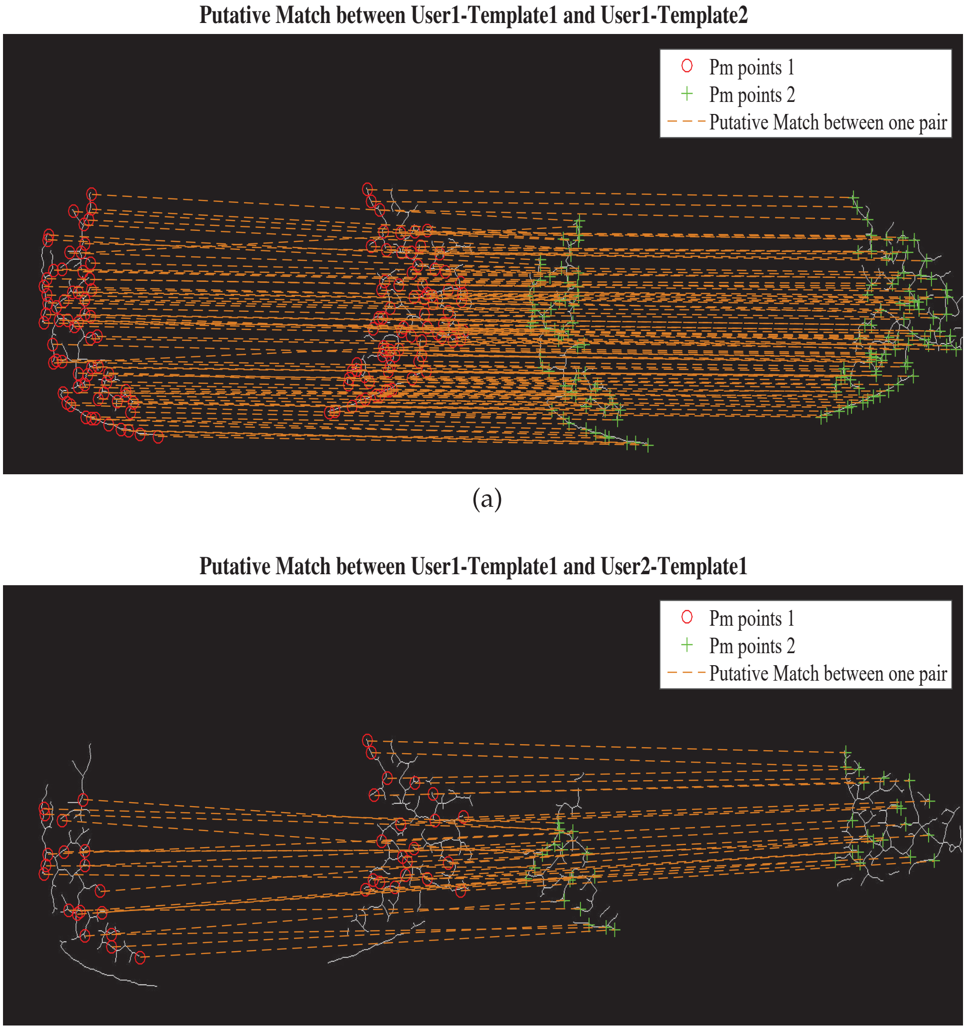
3.6 EXPERIMENTAL RESULTS AND DISCUSSIONS

3.6.1 *Experimental Results Using The UBIRIS.v1 database*

The UBIRIS.v1 database [96] is utilised to evaluate the performance of the proposed sclera recognition method. This database is composed of 1877 eye RGB images collected from 241 individuals in two sessions. More information about this database is in Chapter 2 and examples are shown in Fig. 3.10. The evaluation process has been achieved in single-session and multi-session contexts. For the single-session context, 3 images for training are used and 2 for testing per user whereas the

Algorithm 3.1 Decision process between two sclera template pairs

```
Input: Two binary inputs $(D_\phi, D_p)$ generated by applying (3.28) and (3.29)
Output: Decision to accept or reject a pair set
if ($D_\phi == 1$) then
    if ($D_p == 1$) then
        Pair accepted
    else
        Pair rejected
    end if
else
    Pair rejected
end if
```


Putative Match between User1-Template1 and User1-Template2

(a)

Putative Match between User1-Template1 and User2-Template1

(b)

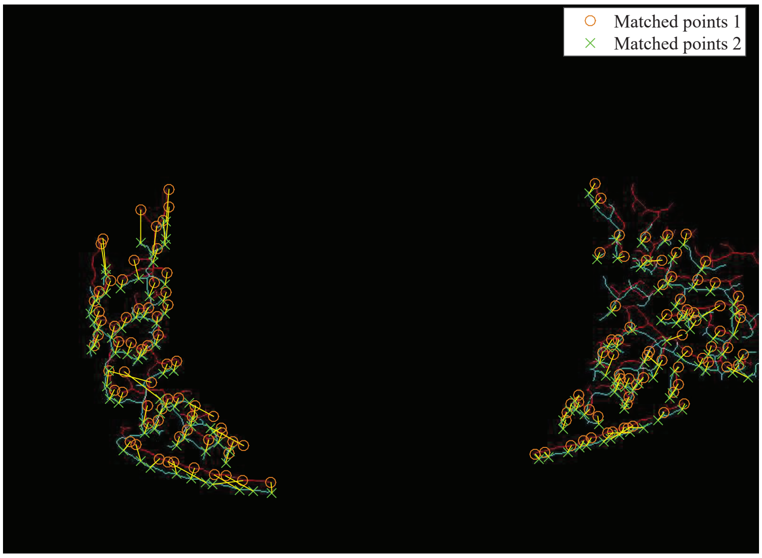
Figure 3.8: Examples of the initial matching process between two sclera templates, (a) initial matching process between two sclera templates belonging to the same individual, (b) initial matching process between two sclera templates belonging to different individuals.

multi-session scenario uses session 1 images for training and session 2 images for testing and vice versa.

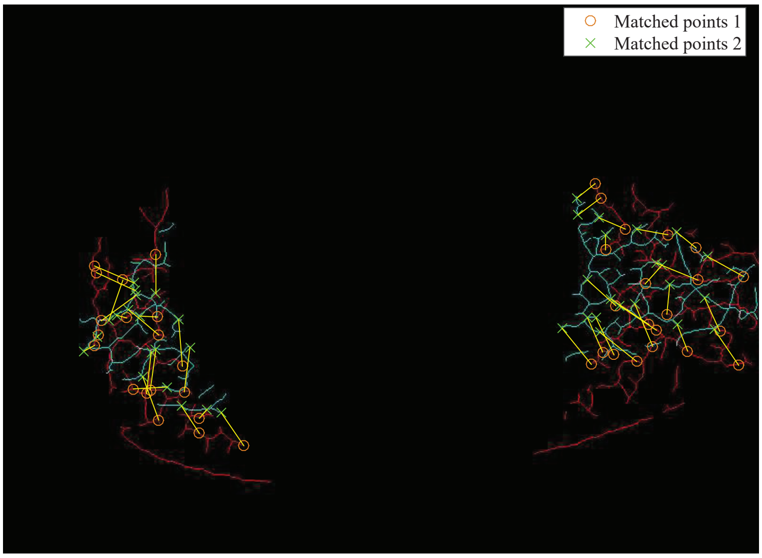
3.6.1.1 *Sclera Validation Method Evaluation*

The performance of the proposed sclera validation is measured by computing the Correct Sclera Validation (CSV) rate where the correct validation images are subjectively evaluated against the eye image decision in (3.10) and CSV is calculated as

$$\mathrm{CSV} = \frac{\text{Number of CAS} + \text{Number of CRS}}{\text{Total number of images}} \times 100\%, \qquad (3.31)$$



(a)



(b)

Figure 3.9: Examples of the final matching result between two sclera templates, (a) matching result between two sclera templates belonging to the same individual, (b) matching result between two sclera templates belonging to different individuals.

where CAS is the Correctly Accepted Sclera image and CRS is the Correctly Rejected Sclera image. As shown in Table 3.1, the proposed method has removed the partial or full occluded images and thus avoids expensive processing times and any human intervention. However, some limitations in the proposed occluded eye detection method can be concluded as follows 1) the detection process depends on the skin detection algorithm which is specifically designed for RGB images

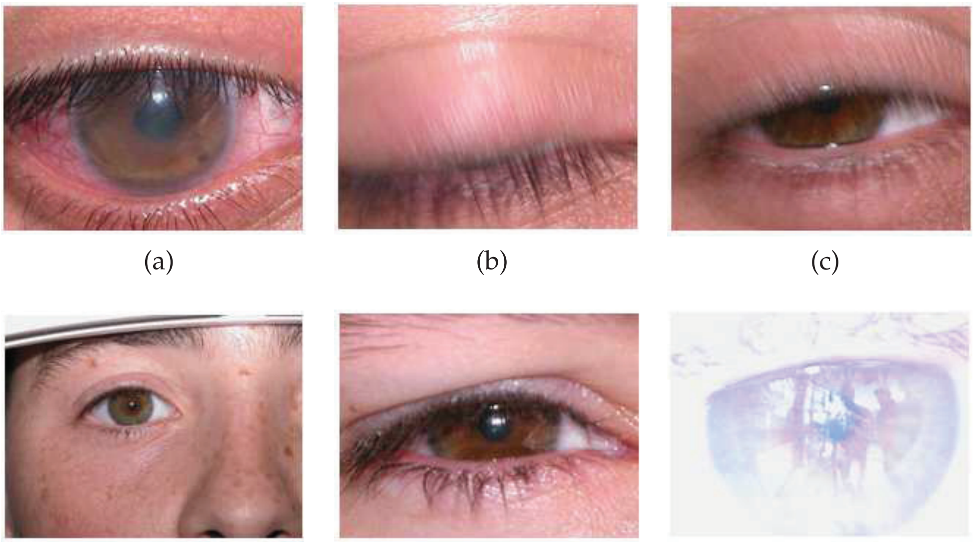
Table 3.1: The Correct Sclera Validation (CSV) rate on the UBIRIS.v1 database

CSV	**Session 1**	**Session 2**
	95.38%	89.45%

and thus, the inability to use it on greyscale images; 2) the process of finding $\mathbf{Arc}_r$ and $\mathbf{Arc}_l$ depends dramatically on the iris segmentation algorithm for the iris centre Therefore, the failure in extracting the iris centre will cause errors in the validation process.

3.6.1.2 *Active Contour Methods for Sclera Segmentation Comparison*

The proposed sclera segmentation method using active contours without edges is compared with state-of-the-art active contour models such as the geodesic [124], balloons [130] and the gradient vector flow [131] active contours in term of accuracy and complexity as shown in Table 3.2. Matlab (version R2013a) on a PC with Intel core i5 3.0 GHz processor and 8.0 GB RAM is used for implementing these algorithms. The proposed algorithm is tended to focus on computer-based



(a) (b) (c)

(d) (e) (f)

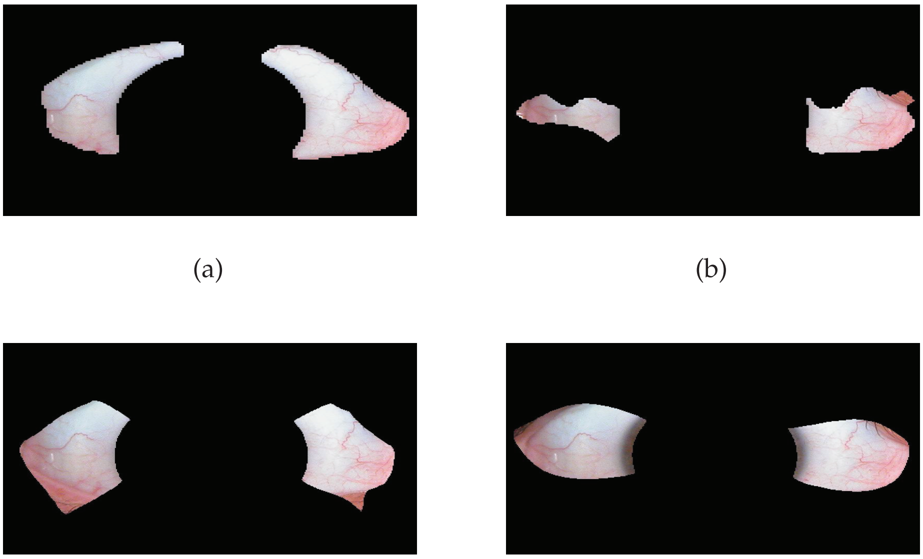
Figure 3.10: The UBIRIS.v1 database poor quality images. (a) an iterated sclera, (b) occluded eye, (c) partially occluded eye, (d) uncropped eye image, (e) insufficient sclera region and (f) high light exposure image.

application and thus the hardware complexity is not discussed. For a fair comparison, the number of iterations in all models is set to 200 and the sclera images are down-sampled with factor = 0.2. For the subjective evaluation results as shown in Fig. 3.11, the proposed method overcomes the active contour model problem by not relying on the edge function which depends on the image gradient to stop the curve evolution. These drawbacks appear severely in noisy images and in the case of irregular sclera shape. In term of accuracy, the proposed adapted sclera segmentation method achieved 98.65% for Session 1 and 95.3% for Session 2 which is significantly higher than other active contour methods. In addition, the proposed method presented significantly lower processing time among these methods (0.003s and 0.010s for Session 1 and 2 respectively).

In addition, a supervised evaluation method used by Proenca [132] is utilised to evaluate the proposed sclera segmentation method. First, 200 sclera images from both sessions are manually segmented and labelled. Then, the sclera border pixels with vector length = n is extracted. After

Table 3.2: Comparison of sclera segmentation using different active contour models in term of accuracy and complexity on the UBIRIS.v1 database

Session	**Active Contours Method**	**Correctly Segmented Sclera Images**	**Processing Time**
Session 1	Geodesic Active Contour[124]	68.02%	3.235 s
Session 2		52.3%	7.128 s
Session 1	Balloons Active Contour[130]	77.15%	3.366 s
Session 2		72.36%	5.402 s
Session 1	GVF Active Contours[131]	87.15%	4.818 s
Session 2		72.36%	6.310 s
Session 1	Proposed Active Contours without Edge	98.65%	0.003 s
Session 2		95.3%	0.010 s



(a) (b)

(c) (d)

Figure 3.11: The sclera regions extracted using different active contour models applied on the same image, (a) proposed active contours without edges, (b) the geodesic, (c) balloons and (d) the gradient vector flow active contours.

that, R^2 which is an objective measure for the good-of-fit is calculated as

$$R^2 = 1 - \frac{\sum_i^n (y_i - \hat{y_i})^2}{\sum_i^n (y_i - \bar{y})^2}, \tag{3.32}$$

where y_i are the desired sclera pixel response value resulted from the manual segmentation, $\bar{y}$ is the mean and $\hat{y_i}$ are the actual sclera pixel response value resulted from the above mentioned active contour segmentation methods. As shown in Fig. 3.12, the proposed method has superiority over the compared methods and hence, active contours without edge method is more reliable for sclera segmentation as it has no stopping edge function.

3.6.1.3 *A State-Of-The-Art Sclera Recognition Method Comparison*

The proposed sclera recognition method is compared with the state-of-the-art methods on the UBIRIS.v1 database as shown in Table 3.3. In addition, the ROC curves for these comparison scenarios are plotted

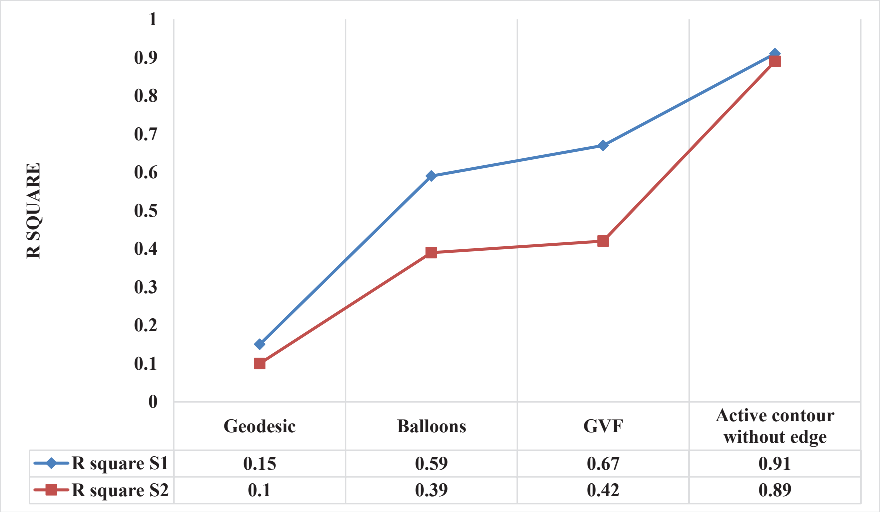


Figure 3.12: R square values for different active contour methods' evaluation at iteration $i = 200$.

in Fig 3.13. The comparison in terms of EER includes single-Session 1 and 2, Session 1 training-Session 2 testing and Session 2 training-Session 1 testing. In terms of human supervision, the authors in [133, 78, 31, 82, 85], for which their learning-based information is shown

Table 3.3: A state-of-the-art comparison of recent work on the UBIRIS.v1 database

Method	**Image discarded**	**EER(%) Single-session**		**EER(%) Multi-session**	
		S1	**S2**	**S1 training S2 testing**	**S2 training S1 testing**
Thomas et al. [133]	Yes (S2)	3.38	n/a	n/a	n/a
Oh and Toh [78]	Yes (S2)	0.47	n/a	n/a	n/a
Zhou et al. [31]	Yes	4.09	9.98	n/a	n/a
Lin et al. [85]	Yes	3.05	n/a	n/a	n/a
Das et al. [82]	Yes (S2)	0.66	n/a	n/a	n/a
Das et al. [86]	No	0.42	0.51	3.95	4.34
Proposed method	No	2.19	2.67	3.68	4.11

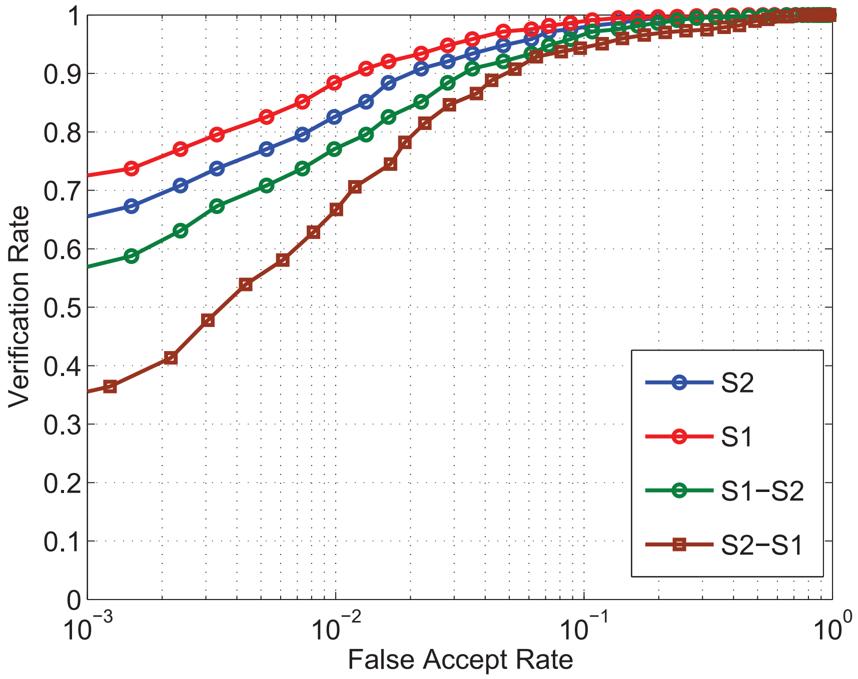


Figure 3.13: The UBIRIS.v1 ROC curves using single and multi-session comparisons.

in Table 2.2, have discarded Session 2 manually as this session has more noisy images which will affect system evaluation. While in [86], the authors stated that they used all the images even the occluded eye images but without including any occluded eye detection method. For the segmentation speed in term of active contours, [86] used a balloon-based active contour with segmentation speed of 0.14s.

In contrast, a method for occluded and partial occluded eye images is proposed which increases the robustness of the system. In addition, the segmentation speed of the proposed sclera segmentation is significantly lower (0.003s). While in terms of the non-linear blood vein movement effect, the methods in [78, 82, 86] do not discuss nor suggest any method to overcome this limitation whereas the proposed system has an adaptive user template registration to extract the most homogeneous features for a final user template.

3.6.1.4 *Sclera versus Iris Recognition Using The UBIRIS.v1 Database*

The proposed sclera recognition system is compared with the iris recognition systems using the UBIRIS.v1 database in order to analyse both iris and sclera recognition systems using visible wavelength

Table 3.4: EER comparison of iris and sclera recognition systems on the UBIRIS.v1 database

Modality	**Method**	**No. of Images**	**EER (%)**
Iris	Proenca and Alexandre [134]	Best 800	2.38
Iris	Duagman [134]	Best 800	3.72
Sclera	Zhou *et al.* [31]	Rand 800	3.83
Sclera	Zhou *et al.* [31]	Best 800	1.34
Sclera	Proposed system	Rand 800	2.31
Sclera	Proposed system	Best 800	1.23

images. For the iris recognition, two systems suggested by Proenca and Alexandre [134] and the traditional Daugman method were used. While for the sclera recognition system, the Zhou et al. [31] method was used in comparison. Similarly, 800 images for evaluating the proposed system are used since in these systems, they chose 800 images with the best quality and least noise to evaluate the accuracy. In addition, 800 randomly selected images are used to evaluate the proposed system. As shown in Table 3.4, the proposed system results are better (EER = 1.23%) compared to Daugman method (EER = 3.72%) and to Proenca and Alexndre method (EER = 2.38%) using iris in the visible wavelength. In addition, the proposed system results are promising compared to Zhou's method in both best and random 800 images selection using proposed sclera recognition system.

3.6.2 *Experimental Results Using The UTIRIS database*

The UTIRIS database [113] is also used to evaluate the proposed segmentation method accuracy. The UTIRIS database consists of two distinct sessions of visible-wavelength and NIR images for 79 individuals with five images from both the left and right eyes as shown in Fig. 3.14. More information about this database is in Chapter 2. The NIR images are discarded as the sclera image with visible blood vessels can be acquired in the visible-wavelength spectrum. The dimension

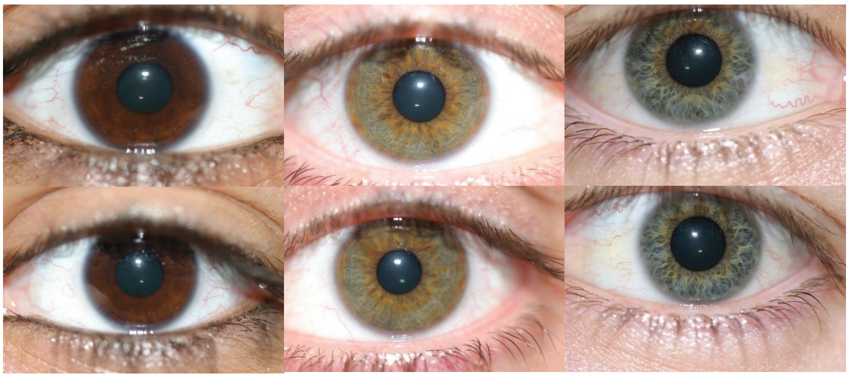


Figure 3.14: UTIRIS right and left eye images. The upper row represents the right eye images whereas the lower row represents the left eye images.

of RGB images is 2048 × 1360. Some of these images are off-angle iris position which have been discarded manually. First, the dimension of RGB images is resized to 600×800 using bilinear interpolation and the same evaluation process parameters are utilised as on the UBIRIS.v1 database. As shown in Table 3.5, the accuracy of the proposed active contours without edges method has been decreased as compared with the UBIRIS.v1 database because some UTIRIS database images are defocused on sclera regions as this database is created for evaluating the iris pigmentation role in an iris recognition system. Therefore, sclera regions are not considered important in their study. However, the processing time remains significantly lower compared to other active contours models. In addition, EER has been calculated (EER = 6.67%) and ROC curve is shown in Fig. 3.15.

Table 3.5: Comparison of sclera segmentation using different active contours models in term of accuracy and complexity on the UTIRIS database

Active Contours Method	**Correctly Segmented Sclera Images**	**Processing Time**
Geodesic Active Contour [124]	66.33%	3.145 s
Balloons Active Contour [130]	71.01%	3.708 s
GVF Active Contour [131]	82.44%	4.779 s
Proposed Active Contours without Edge	90.82%	0.004 s

3.6.3 *Computational Load of The Proposed System*

This section provides the computational load of the proposed system in each step as shown in Table 3.6. The properties of simulation program and PC are mentioned in the subsection 3.6.1.2. The computational complexities of each step are calculated by running the simulation program 20 times using both the UBIRIS.v1 and UTIRIS databases and the average processing time is recorded.

Table 3.6: Time complexity of the proposed system

Sclera Recognition Steps	**Complexity**
Iris Segmentation	1.97 s
Sclera Validation	0.24 s
Sclera Segmentation	0.007 s
Feature Extraction	0.29 s
User Template Registration	0.46 s

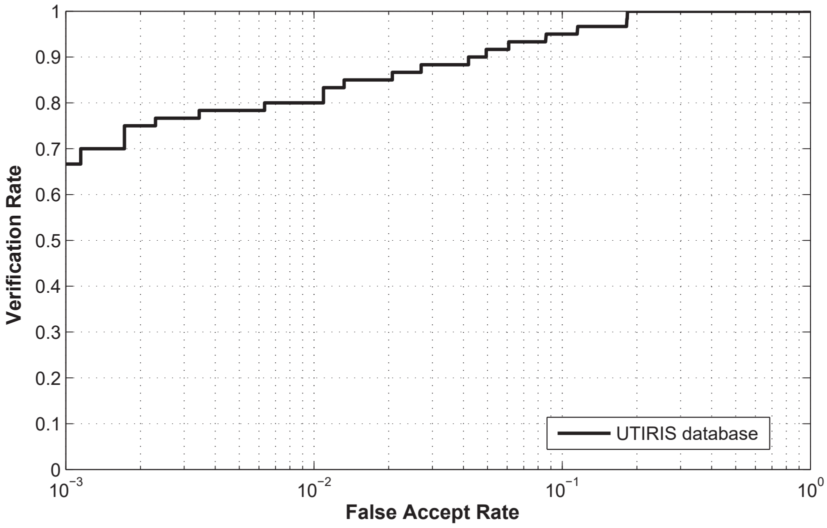


Figure 3.15: The UTIRIS database ROC curve.

3.6.4 *Empirical Parameters Tuning Analysis*

To analyse the tuning criteria for the system controlling parameters, these variables are summarised according to their location as follows:

- Occluded eye detection for sclera validation: this section has three controlling parameters which are Θ, $per_{\mathbf{nsk}_{up}}$ and $per_{\mathbf{nsk}_{down}}$. For Θ, increasing the angle width provides more accurate results at the expense of processing time. While for $per_{\mathbf{nsk}_{up}}$ and $per_{\mathbf{nsk}_{down}}$, the higher the values, the more sclera samples are rejected.

- Sclera segmentation: this section has two controlling parameters which are μ and the number of iterations i. Increasing i means more curve convergence with a possibility that the fitting curve might pass through the sclera border whereas μ is discussed in subsection 3.2.3.2.

- Sclera feature extraction and template registration: to find a corner point, two parameters need to be set which are k and th_c. The corner sensitivity is controlled by k where smaller value means more sharp corner is detected. Whereas $0 < th_c < 1$ determines the percentage of minimum accepted quality corners within a certain region. On the other hand, a higher $d_m ax$ value means more accepted pairs in $\mathbf{St}_1$ and $\mathbf{St}_2$ whereas the lower ϵ value, the more accurate sets of feature pairs are detected via the RANSAC.

- Sclera matching and decision process: finally, the decision process is determined through first detecting the pair sets depending on th_p and th_ϕ whereas th_f sets the final decision. Increasing the values of th_p and th_ϕ are more likely to produce false pair sets. While th_f determines FRR and FAR where the higher the th_f, the higher the FRR and the lower FAR and vice versa.

3.7 CHAPTER SUMMARY

This chapter has developed efficient model for a sclera recognition system and discussed the performance and reliability when eye images are captured under constrained imaging conditions. In particular, a novel occluded eye detection method has been proposed to discard the noisy images. Sclera segmentation has been achieved through an adaptive active contours without edges method. In addition, a new method for blood vein extraction and mapping has been suggested based on IUWT. For the non-linear movement of the blood veins, a new user template registration was proposed to overcome this effect and create a robust sclera feature template.

The results using the UBIRIS.v1 and UTIRIS databases showed that the proposed method has a lower processing time with high segmentation accuracy. In addition, the proposed method minimises the human supervision during the recognition process introducing more robustness. This chapter has not evaluated the use of off-angle sclera images which will be considered in the next chapters.

4

SCLERA RECOGNITION UNDER RELAXED IMAGING CONDITIONS: KEY ISSUES, CHALLENGES, AND RESOLUTIONS

4.1 INTRODUCTION

This chapter investigates the challenges of applying sclera recognition under relaxed imaging conditions (i.e., eye image captured in a real life scenario is used). Some challenges include uncontrolled human pose, multiple iris gaze directions, different eye image capturing distance, and variation in lighting conditions. Therefore, designing a sclera recognition system under these severe conditions is challenging. There has been much research on sclera recognition under constrained imaging conditions, yet no published research has investigated or overcome these conditions. This chapter provides practical solutions to mitigate the aforementioned noise factors and challenges.

To overcome challenging points (6-10) mentioned in subsection 2.3.2, the following contributions are proposed in this chapter which can be summarised as

- A novel fusion method for sclera segmentation which is adapted to noise factors.
- An efficacious sclera segmentation method based on fusing binary skin-based classifiers for constrained and unconstrained imaging conditions.
- An efficient sclera template rotation alignment method which is invariant to eye rotation.
- A robust sclera scaling method to minimise the effect of capturing a sclera image on-the-move and at-a-distance.
- Robust blood vessel enhancement and feature extraction methods based on what are denoted as *complex-sclera* features via Dual-Tree Complex Wavelet Transform (DT-CWT).

The chapter will first discuss the fusion of multiple eye skin and sclera area classifiers in order to achieve higher segmentation rate. In addition, a rotation and alignment scheme will be proposed to eliminate the effect of head pose rotation and multiple distance capturing. Next, a novel *complex-sclera* feature extraction method will be explained to overcome some issues in the discrete and Gabor wavelet. Then, the evaluation of the proposed methods will be introduced in the results and discussion section.

4.2 SCLERA SEGMENTATION USING MULTIPLE COLOUR SPACES

In this section, a new method for sclera segmentation is proposed where two strategies are adopted to extract the sclera white area and the skin area surrounding the eye. First, six classifiers are applied in a pixel-wise operation to create a map for the eyelid skin. Then, a sclera area classifier is applied to extract a second map. After that, a fusion process and refinement are used to merge the two maps and a novel method to remove spur pixels is proposed for final map extraction. This is depicted in Fig. 4.1.

First for the iris segmentation, the integro-differential operator is used to locate the circular pupil and iris regions [32]. Specular reflection removal is applied to remove the brightest points in the eye image. Then, for each image, iris centre coordinates $C_{iris}(x, y)$ and radius r are stored.

4.2.1 *Skin-based Pixel Classifiers*

Many binary skin-based segmentation algorithms have been developed to segment skin area on a human body depending on a different colour space such as RGB, HSV, HSI and YCbCr [120, 135, 136, 137, 138, 139]. These methods have achieved satisfactory results in terms of skin segmentation accuracy and complexity. However, these methods were tested only on limited databases with some having constrained imaging conditions. In addition, recent pixel-based sclera segmentation methods have used one colour space which is not robust to noise factors and unconstrained imaging necessitating more post-processing complexity. As a result, sclera segmentation design depending on

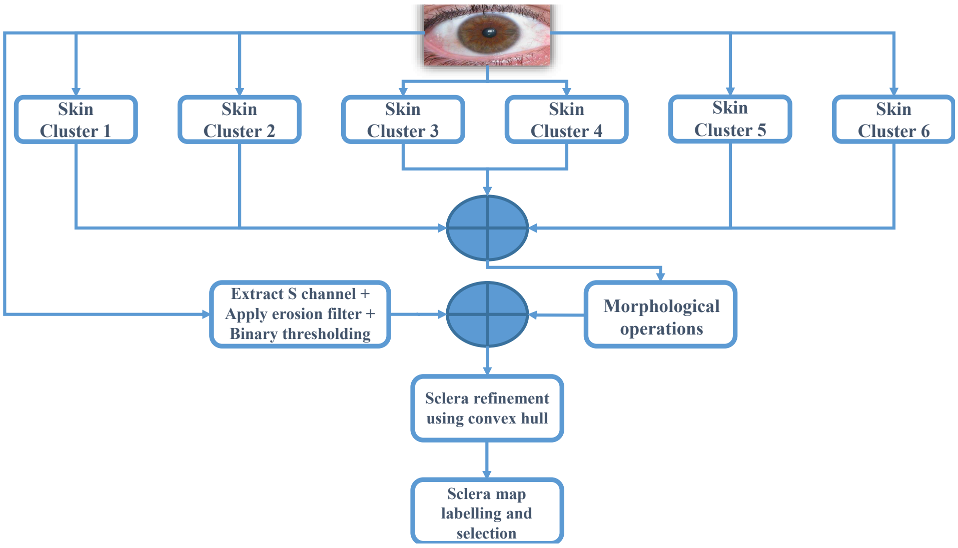


Figure 4.1: Proposed sclera segmentation algorithm.

multiple colour spaces will ensure the versatility of segmentation for constrained and unconstrained eye imaging but the complexity could be considerable. Therefore, a fusion method is proposed which processes each pixel of the eye image in the RGB, HSV, HSI and YCbCr colour spaces in order to mitigate the noise factors and to attain higher segmentation accuracy. First, a brief summary of each method according to their colour spaces is presented as:

1. **RGB** [120]: this method creates skin maps for processing two illumination conditions on images which are daylight and flash illumination resulting in the first skin map $\mathbf{s}_{k1}$. It has been explained in chapter 3, subsection 3.2.2.

2. **rgb** [135]: this method utilises the rgb components in the normalised form and a Restricted Covering Algorithm (RCA) as its selective learner. The combination rules which were found empirically to introduce the highest skin segmentation accuracy are defined as

$$\mathbf{s}_{k2}(x,y) = \begin{cases} & \frac{r}{g} > 1.185 \\ 1, & \frac{r.b}{(r+g+b)^2} > 0.10, \\ & \frac{r.g}{(r+g+b)^2} > 0.112 \\ 0, & \text{otherwise} \end{cases} \tag{4.1}$$

where the normalised r, g and b are defined as

$$r = \frac{R}{R+G+B}, \quad g = \frac{G}{R+G+B}, \quad b = \frac{B}{R+G+B}. \tag{4.2}$$

3. **HSV$_1$** [136]: to detect skin pixels, the V map is thresholded to discard the darker colours whereas H and S colour maps are selected to ensure that the segmentation for the red colours with the exclusion of the pure red and the very dark red which generally result from the small variations in lighting conditions. The thresholding rates are defined as

$$\mathbf{s}_{k3}(x,y) = \begin{cases} & V \geqslant 40 \\ 1, & 0.2 < S < 0.6 \\ & 0^\circ < H < 25^\circ \parallel 335^\circ < H < 360^\circ \\ 0, & \text{otherwise} \end{cases}, \tag{4.3}$$

where H and V values are $[0^\circ \to 360^\circ]$ and $[0 \to 100]$ respectively.

4. **HSV$_2$** [137]: the thresholds to find skin pixels are calculated by finding a set of planes through successive adjustments depending on the segmentation results. The segmentation process is defined in Algorithm 4.1. In this algorithm, the colour map H values are $[-180^\circ \to 180^\circ]$, while S and V values are $[0 \to 100]$.

5. **HSI** [138]: this method depends on the HSI colour space to detect skin pixels as the colours in the HSI are stable under different lighting conditions. The classification of skin pixels is defined in Algorithm 4.2. The intensity I values are $[0 \to 255]$, S values are $[0 \to 255]$ and H values are $[-180^\circ \to 180^\circ]$.

6. **YCbCr** [139]: the last skin segmentation algorithm depends on the YCbCr colour space by exploiting the spatial distribution

Algorithm 4.1 Skin-based pixels classification using HSV_2 method

```
if ($H \leqslant (-0.4V + 75)$) then
    if ($V \geqslant 40, 10 \leqslant S \leqslant (-H - 0.1V + 110),$) then
        $s_1(x, y) = 1$
    end if
end if
if ($H \geqslant 0$) then
    if ($S \leqslant (0.08(100 - V)H + 0.5V),$) then
        $s_2(x, y) = 1$
    end if
end if
if ($H < 0$) then
    if ($S \leqslant (0.5H + 35),$) then
        $s_3(x, y) = 1$
    end if
end if
if ($s_1(x, y) \,\&\, s_2(x, y) \,\&\, s_3(x, y) = 1$) then
    $s_{k4}(x, y) = 1$
end if
```

characteristics of the human skin. The chrominance components are used to classify input image pixels as

$$s_{k6}(x, y) = \begin{cases} 1 & 77 \leqslant Cb \leqslant 127, \\ & 133 \leqslant Cr \leqslant 173, \cdot \\ 0, & \text{otherwise} \end{cases} \tag{4.4}$$

Having presented the six binary skin classifiers, the next step is to fuse them into one sclera map. First, if the input eye image is in the RGB colour space, the necessary transformations to HSV, HSI and YCbCr colour spaces along with predefined scales are achieved. Next, six binary maps are extracted and the decision to classify the eye image pixels as a skin or non-skin pixel is defined as

$$\mathbf{s}_{kt}(x, y) = \begin{cases} 1, & \text{if } \sum_{i=1}^{6} s_{ki}(x, y) \geqslant th_s \\ 0, & \text{otherwise} \end{cases}, \tag{4.5}$$

where th_s is evaluated in subsection 4.7.2. To eliminate the noisy pixels, binary morphological operations involving filling the holes and

Algorithm 4.2 Skin-based pixels classification using HSI method

```
if ($I > 40$) then
    if ($S > 13, S < 110, H > 0, H < 28, H > 332, H < 360$) then
        $s_{k5}(x, y) = 1$
    end if
    if $S > 13, S < 75, H > 309, H < 331$ then
        $s_{k5}(x, y) = 1$
    end if
end if
```

removing spur pixels within s_{kt} are used. The s_{kt} mask is shown in Fig. 4.2b.

4.2.2 *Sclera Area Classifier*

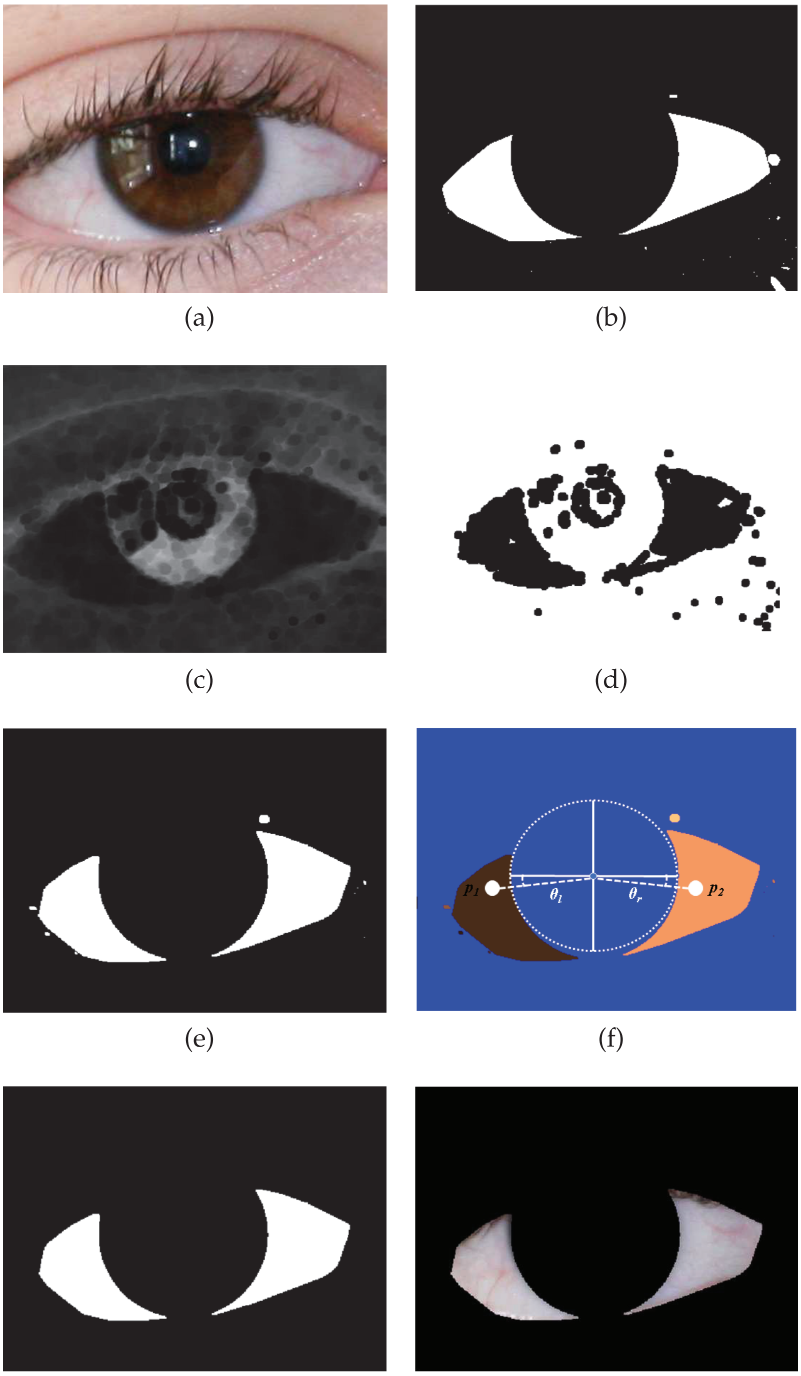
Next, to remove some misclassified skin pixels, another sclera map s_m is created based on the properties of the sclera in the saturation level S of the HSV colour space where the sclera area is presented as the darkest area as the sclera region has a low saturation value. Then, an erosion filter [140] is used where the saturation image is probed with structure element having a disk shape with size of 5×5 in order to erode the white pixels within the sclera as shown in Figs. 4.2c and 4.2d. The resulting sclera map is then converted to a binary form as

$$\mathbf{s}_m(x, y) = \begin{cases} 1, & \text{if } I > th_i \\ 0, & \text{otherwise} \end{cases}, \tag{4.6}$$

where th_i is set empirically to 0.09. Having presented the aforementioned eye skin and sclera area segmentation methods, a fusion method is next used to overcome illumination variations and noise factor problems.

4.2.3 *Fusion and Refinement Process*

A new fusion method for segmenting sclera regions is proposed which is adaptive to noise factors. This method creates two binary maps and fuses them to enhance the sclera area detection for noisy images. The



(a) (b)

(c) (d)

(e) (f)

(g) (h)

Figure 4.2: Proposed sclera segmentation masks, (a) an eye image from the UBIRIS.v2 database at a distance of 4 meters, (b) skin fusion segmentation $\mathbf{s}_{kt}$ mask, (c) an eye image in the saturation level after applying an erosion filter, (d) $\mathbf{S}_m$ mask, (e) the fused $\mathbf{S}_f$, (f) binary labelling process mask with p_1 and p_2 locations specified, (g) the final binary mask after removing all the noise pixels and (h) the segmented sclera regions.

fusion of $\mathbf{S}_{kt}$ and $\mathbf{S}_{m}$ to create the final sclera map $\mathbf{s}_{f}$ is achieved as follows

$$\mathbf{s}_{f}\left(x,y\right)=\begin{cases}1, & \text{if } s_{kt}\left(x,y\right)\bigcap s_{m}\left(x,y\right)=1\\ 0, & \text{otherwise}\end{cases}, \tag{4.7}$$

The next step is to post-process the $\mathbf{S}_{f}$ mask to mitigate noise pixels before segmenting the sclera. First, the binary mask $\mathbf{S}_{f}$ is complemented and the convex hull of sclera regions and noisy regions is applied for the final shape refinement. Then, the iris mask is applied using a logical AND operation to remove the iris region. Finally, a morphological operation is applied to remove spur pixels. Outputs of the skin clusters fusion algorithm and the erosion filter are shown in Fig. 4.2e.

4.2.4 *Sclera Map Labelling and Selection*

Although fusing two binary sclera maps and applying post-processing steps have achieved a high quality segmented sclera area, however some spur pixels will remain. Therefore, a new method for selecting valid sclera areas is proposed.

The binary labelling process involves finding the connected area in $\mathbf{S}_{f}$ and selecting the areas which overlap with predefined points referring to the location of sclera regions which are normally around the iris. For a Frontal-Gaze (F-G) eye, two points are set p_1 and p_2 as depicted in Fig. 4.2f whereas for a Left-Gaze (L-G) eye image only one point on the right is required and vice versa for a Right-Gaze (R-G) eye image. The positions of these points are defined as

$$p_1\left(x,y\right)=\left(1.15\times r\cos(\theta_r),1.15\times r\sin(\theta_r)\right), \tag{4.8}$$

and

$$p_2\left(x,y\right)=\left(1.15\times r\cos(\theta_l),1.15\times r\sin(\theta_l)\right), \tag{4.9}$$

where $(1.15\times r)$ is set to ensure that the selected right and left points are outside the iris with angles $\theta_r=-10°$ and $\theta_l=190°$. The final sclera segmentation mask is shown in Fig. 4.2g.

4.3 SCLERA TEMPLATE ROTATION ALIGNMENT AND DISTANCE SCALING

During the capturing process, the subject normally moves back and forward, rotating the head to the left and right while looking at different directions. These effects need to be considered in order to design an efficient sclera recognition system. Therefore, a new method which is named Sclera Template Rotation Alignment and Distance Scaling (STRADS) method for sclera template is proposed which is invariant to eye image rotation and distance scaling.

First, the rotation alignment process for the F-G eye image which is depicted in Fig. 4.3 is achieved as follows

i. divide each individual mask into two sections around the detected iris and extract the four internal corners c_1, c_2, c_3 and c_4 of the sclera binary mask using Harris corner detection [110];

ii. calculate the internal angles of these corners ϕ_1, ϕ_2, ϕ_3 and ϕ_4 with respect to the y-axis of the iris centre position;

iii. adjust the angles ϕ_1, ϕ_2, ϕ_3 and ϕ_4 to be equal to 45° by applying an image rotation function which uses the nearest neighbour interpolation to rotate an image with a specified angle τ.

To determine ϕ_1, ϕ_2, ϕ_3 and ϕ_4, the calculation process of ϕ_1 will be discussed only as it applies similarly to all other angles. ϕ_1 can be defined as

$$\phi_1 = \tan^{-1} \frac{(y_1 - y_0)}{(x_1 - x_0)}. \tag{4.10}$$

Then, τ_l which is the degree of rotation is calculated as

$$\tau_l = \begin{cases} (\phi_2 - \phi_1), & \text{if } \phi_1 \ > \ \phi_2 \parallel \phi_1 \ < \ \phi_2 \\ 0, & \text{otherwise,} \end{cases} \tag{4.11}$$

where τ_r is defined similarly. The value of τ determines the direction of rotation where the positive value rotates the image counter-clockwise and vice versa.

The next stage is to minimise the distance effect on the sclera template. An extraction method is designed based on RGB threshold-

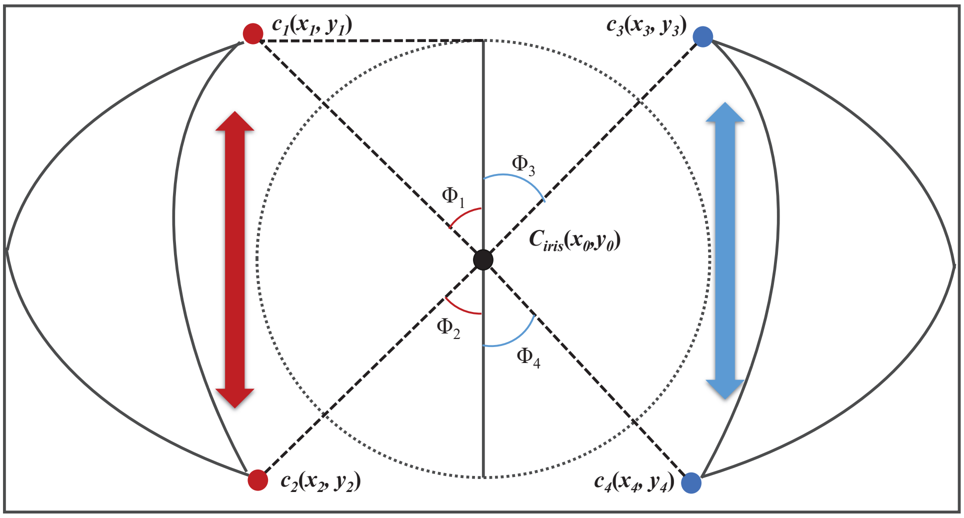


Figure 4.3: Sclera rotation alignment where internal sclera corners c_1, c_2, c_3 and c_4 are detected and ϕ_1, ϕ_2, ϕ_3 and ϕ_4 are calculated.

ing by simply extracting the pixels with the RGB values that satisfy $(R > 30, G > 25, B > 44)$. Then, the extracted areas of the sclera templates at different distances are resized into an equal size of 400×300 using bilinear interpolation. This is depicted in Fig. 4.4.

4.4 SCLERA BLOOD VESSELS ENHANCEMENT

This section will discuss the enhancement process of blood vessels within the sclera. To enhance the segregation of the sclera blood veins from their background, two steps are utilised which include extracting the green channel of the sclera image and applying CLAHE which enhances the local contrast of blood vessels. Then, an efficient method to extract the most distinctive features is suggested depending on the complex features of these blood vessels. This is achieved through the DT-CWT. In addition, there are some state-of-the-art methods used in the literature such as the Gabor Wavelet Transform (GWT), Discrete Wavelet Transform (DWT), and CDF 9/7 which will be discussed briefly and compared to the proposed sclera blood vessel enhancement method in Section 4.7.

Figure 4.4: Sclera distance scaling for sclera templates from the UBIRIS.v2 database with R-G at distances 4, 5 and 6 meters. The top row images correspond to distance 4 meters, the middle row images 5 meters and the bottom row images at 6 meters.

4.4.1 *Complex Sclera Features for Noisy Eye Image Factors*

A novel feature representation denoted as *complex-sclera* features is proposed based on the DT-CWT [141]. The advantages of the DT-CWT over the DWT [107] and GWT [142] are the shift invariance properties which better characterise blood vessel patterns; its ability to extract geometric features efficiently unlike the DWT [143]; it generates less redundant information and its relatively low computational cost unlike the GWT [144].

The 2D DT-CWT is depicted by six complex wavelets defined as

$$\psi^1(x,y) = \phi(x)\psi(y), \psi^4(x,y) = \phi(x)\overline{\psi(y)}, \tag{4.12}$$

$$\psi^2(x,y) = \psi(x)\psi(y), \psi^5(x,y) = \psi(x)\overline{\psi(y)}, \tag{4.13}$$

$$\psi^3(x,y) = \psi(x)\phi(y), \psi^6(x,y) = \psi(x)\overline{\phi(y)}, \tag{4.14}$$

where $\psi(x) = \psi_h(x) + j\psi_g(x)$ is the 1-D complex wavelet function and $\phi(x) = \phi_h(x) + j\phi_g(x)$ is the scaling function. The 1-D consists of two real DWTs in parallel to produce the real and the imaginary parts of the complex function. Each of the $\psi^{[1-6]}(x,y)$ provides six bandpass sub-bands of complex coefficients oriented at angles $-75°, -45°, -15°, 15°, 45°, 75°$. In practice, DT-CWT can be implemented efficiently by subtracting and adding two 2-D DWT in parallel which requires only fourfold computation cost of 2-D DWT and with a redundancy factor of 2:1 for the magnitude coefficients decomposition. More details related to the DT-CWT can be found in [141].

Given a sclera image of size $w_o \times h_o$, image dimensions must first be resized to a decisive size to make sure that the decomposition of the next wavelet level can be achieved for a certain wavelet level J defined as

$$w_J = w_o - w_o \bmod \left(3 \times 2^{J+1}\right), \tag{4.15}$$

$$h_J = h_o - h_o \bmod \left(3 \times 2^{J+1}\right). \tag{4.16}$$

The magnitude information of the complex coefficients is used as *complex-sclera* features as these magnitude values are slowly changing and insensitive to slight image shifts [141]. The 2D DT-CWT is applied and for each level, the magnitude of the six subbands is extracted and summed creating a total magnitude map M_j. Then, $\mathbf{m}_J$ which is the augmented feature set for a certain J levels is calculated as

$$\mathbf{m}_J^{(\rho)} = \left(M_1^{(\rho)^T}, M_2^{(\rho)^T}, \cdots, M_J^{(\rho)^T} \right), \tag{4.17}$$

where $(.)^T$ is the transpose operator, and $M_1^{(\rho)^T}$ denotes the columnised magnitude coefficients of M_j. An example of a sclera template with five complex levels extracted is shown in Fig. 4.5. The augmented feature set $\mathbf{m}_J^{(\rho)}$ resulting from J levels of DT-CWT has a size of $z = w_J \times h_J \times J$ leading to increase in dimensionality and higher processing time.

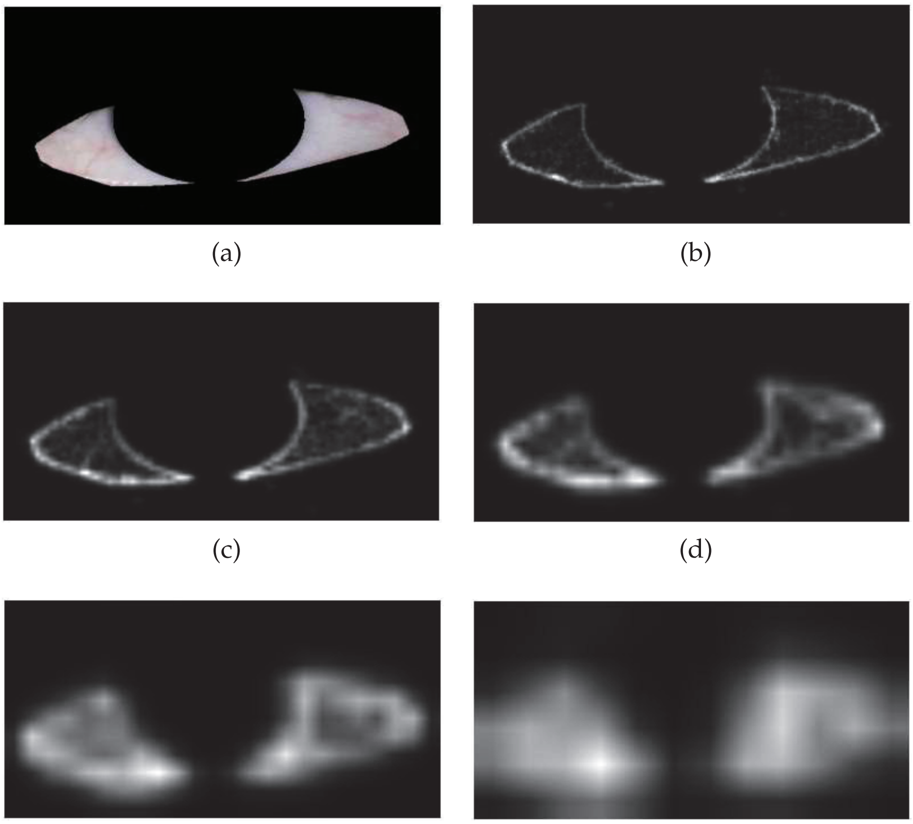


Figure 4.5: The segmented sclera map after applying DT-CWT with $J = 5$, (a) is the sclera template and (b)-(f) are the extracted complex wavelets for $J = [1 : 5]$ respectively.

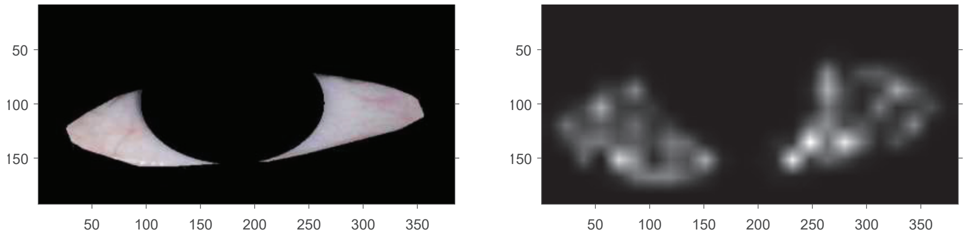
4.4.2 *2D Discrete Wavelet Transform DWT*

For a sclera image with size $x \times y$, the 2D-DWT [145] is defined as

$$W_{\varphi}\left(j_0, m, n\right) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f\left(x, y\right) \varphi_{j_0,m,n}\left(x, y\right), \tag{4.18}$$

$$W_{\psi}^{i}\left(j, m, n\right) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f\left(x, y\right) \psi_{j,m,n}^{i}\left(x, y\right), \tag{4.19}$$

where $W_{\varphi}\left(j_0, m, n\right)$ are the coefficients of the $f\left(x, y\right)$ at scale j_0, $W_{\psi}^{i}\left(j, m, n\right)$ with $i \in \{H, V, D\}$ are either the horizontal, vertical, and diagonal coefficients respectively which provide details for scales



(a) (b)

Figure 4.6: The segmented sclera map after applying DWT with $J = 5$, (a) is the sclera template and (b) is the extracted discrete wavelet for $j = 4$.

$j \geqslant j_0$, and $\varphi_{j_0,m,n}(x, y), \psi^i_{j,m,n}(x, y)$ are the scaled and translated basis functions defined as

$$\varphi_{j,m,n}(x, y) = 2^{j/2}\varphi\left(2^j x - m, 2^j y - n\right), \tag{4.20}$$

$$\psi^i_{j,m,n}(x, y) = 2^{j/2}\psi^i\left(2^j x - m, 2^j y - n\right), \tag{4.21}$$

where $m = n = 1, 2, \cdots, 2^j - 1$.

Similarly, the magnitude information of the DWT is utilised creating a total magnitude map W_j. Then, the augmented feature set for the J levels is defined as

$$\mathbf{w}^{(\rho)}_J = \left(W^{(\rho)^T}_1, W^{(\rho)^T}_2, \cdots, W^{(\rho)^T}_J\right). \tag{4.22}$$

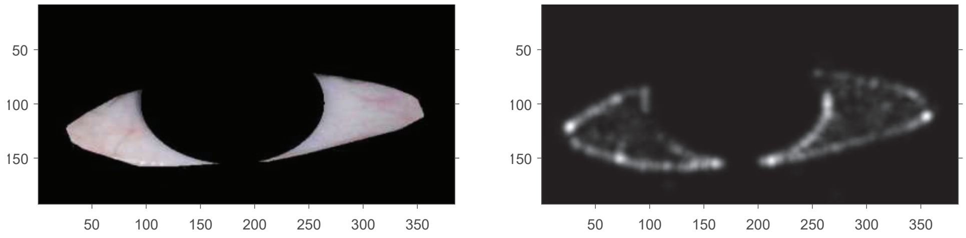
An example of sclera template with one discrete wavelet level $j = 4$ extracted is shown in Fig. 4.6.

4.4.3 *Gabor Wavelet Transform GWT*

The sclera image features are also represented by using a bank of Gabor filters [142]. The GWT can be defined as

$$\psi_{u,v}(x, y) = \frac{f_u^2}{\pi\kappa\eta} e^{-\left(\left(f_u^2/\kappa^2\right)x'^2 + \left(f_u^2/\eta^2\right)y'^2\right)} e^{j2\pi f_u x'}, \tag{4.23}$$

where $f_u = f_{max}/2^{(u/2)}$ is the centre frequency of the Gabor kernels, the parameters κ and η determine the ratio between the size of the Gaussian envelope and f_u, $x' = x\cos\theta_v + y\sin\theta_v$ and $y' = -x\sin\theta_v +$



(a) (b)

Figure 4.7: The segmented sclera map after applying GWT, (a) is the sclera template and (b) is the extracted wavelet for one scale $u = 1$ and 8 orientations $v = 8$.

$y \cos\theta_v$ with $\theta_v = v\pi/8$ representing the orientation of each kernel. The enhanced sclera green channel image $I(x, y)$ is convolved with $\psi_{u,v}(x, y)$ with scale u and orientation v as

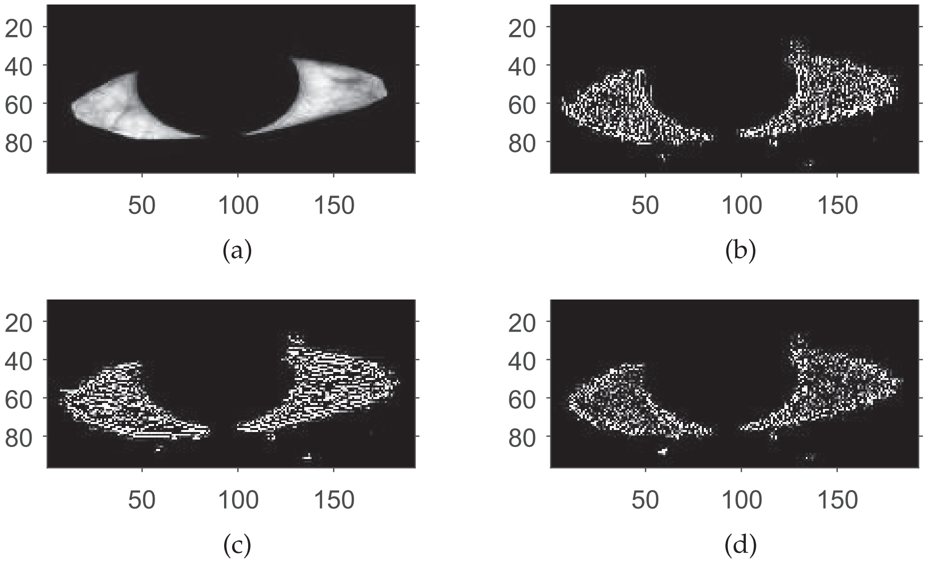
$$G_{u,v}(x, y) = I(x, y) * \psi_{u,v}(x, y), \tag{4.24}$$

where $G_{u,v}(x, y)$ is the complex filtering output decomposed into its real and imaginary parts. In this chapter, Gabor filter banks are constructed with eight orientations, $v \in \{0, \cdots, 7\}$, five scales, $u \in \{0, \cdots, 4\}$, $\kappa = \eta = \sqrt{2}$ and $f_{max} = 0.25$ as in [146, 147]. The magnitude information which varies slowly with the spatial position is utilised, while the phase is generally affected by position to overcome the non-linearity in movement of the blood veins [148, 31].

With the proposed values of the scale and orientation, a single sclera image will produce 40 Gabor magnitude representations resulting in an inflation of data dimensionality. To overcome this issue, a down-sampling technique is used. Each $G_{u,v}(x, y)$ is down-sampled by a factor $\rho = 0.5$ in order to reduce dimensionality and normalised by zero-mean and unit variance producing $G_{u,v}^{(\rho)}(x, y)$. The augmented feature set equals

$$\mathbf{g}^{(\rho)} = \left(G_{0,0}^{(\rho)^T}, G_{0,1}^{(\rho)^T}, \cdots, G_{4,7}^{(\rho)^T} \right), \tag{4.25}$$

where $G_{u,v}^{(\rho)^T}$ denotes the columnised coefficients of $G_{u,v}(x, y)$. An example of sclera template with one Gabor wavelet scale $u = 1$ and 8 orientations $v = 8$ is shown in Fig. 4.6.



(a) (b) (c) (d)

Figure 4.8: The segmented sclera map after applying CDF 9/7 with one level of decomposition, (a) is the sclera template and (b)-(d) are the extracted wavelets.

4.4.4 *Cohen-Daubeches-Feauveau 9/7 Wavelet* CDF 9/7

The 9/7 bi-orthogonal wavelet filter bank CDF 9/7 [149] has achieved a wide reputation for image coding and image compression. It has been chosen by the FBI as the best fingerprint image compression technique [150] as the FBI retains a large fingerprint database because the wavelet and scaling functions of the CDF 9/7 have similar length and are symmetrical. One major drawback of the CDF 9/7 is that, as all the filter coefficients are irrational, it therefore has a higher computational complexity as compared to the DWT using bi-orthogonal spline wavelets [92]. More information about the implementation of the CDF 9/7 is available in [149].

4.5 BLOOD VESSEL FEATURE EXTRACTION

The augmented feature set which could be denoted as $\mathbf{x}^{(\rho)}$ resulting from the DT-CWT, GWT, DWT, and CDF 9/7 filtering processes is subjected to Kernel Linear Discriminant Analysis (KLDA) [151]. The KLDA method maps the feature sets by a non-linear mapping Φ into some feature space $\mathcal{F}$ and applies Fisher's linear discriminant with a Mercer kernel strategy [152]. This strategy formulates the algorithm using a dot product $(\Phi(\mathbf{x}), \Phi(\mathbf{y}))$ to overcome the limitation of the inability to solve between and within class scatter matrices directly if $\mathcal{F}$ is very high or has infinite dimension space.

Let the $\mathbf{x}^{(\rho)}$ training feature vectors be arranged in $\mathbf{X}_N = \{\mathbf{x}_1^{c_1}, \cdots, \mathbf{x}_{n_1}^{c_1}, \cdots, \mathbf{x}_{n_N}^{c_N}\}$, $\mathbf{X}_N \in R^{d \times n_N}$ for N-classes, $c_1, c_2, \cdots, c_N$ where $n_N = \sum_{i=1}^{N} n_i$ is the total number of samples and n_i is the number of training samples per class. KLDA can be solved in $\mathcal{F}$ by finding vector $(\boldsymbol{\alpha})$ which maximises

$$J(\boldsymbol{\alpha}) = \arg\max_{\boldsymbol{\alpha}} \frac{\boldsymbol{\alpha}^T K_B \boldsymbol{\alpha}}{\boldsymbol{\alpha}^T K_W \boldsymbol{\alpha}}, \tag{4.26}$$

where K_B and K_W are the between and within class scatter matrices respectively defined as

$$K_B = \frac{1}{n_N} \sum_{i=1}^{c_N} \sum_{j=1}^{n_i} \left(\mathbf{m}_i - \mathbf{m}_j\right) \left(\mathbf{m}_i - \mathbf{m}_j\right)^T, \tag{4.27}$$

$$K_W = \frac{1}{n_N} \sum_{c_i=1}^{c_N} \frac{1}{n_i} \sum_{j=1}^{n_i} \left(\mathbf{l}_j - \mathbf{m}_i\right) \left(\mathbf{l}_j - \mathbf{m}_i\right)^T, \tag{4.28}$$

where

$$m_i = \frac{1}{n_i} \sum_{j=1}^{n_N} \sum_{k=1}^{n_i} k\left(\mathbf{x}_j, \mathbf{x}_k^i\right), \tag{4.29}$$

and

$$\mathbf{l}_j = \left(k\left(\mathbf{x}_1, \mathbf{x}_j\right), k\left(\mathbf{x}_2, \mathbf{x}_j\right), \cdots, k\left(\mathbf{x}_{n_i j}, \mathbf{x}_j\right)\right)^T. \tag{4.30}$$

The homogeneous polynomial kernel is utilised with the KLDA as in [153] which equals

$$k(\mathbf{x}, \mathbf{y}) = \left(\mathbf{x}^T \cdot \mathbf{y}\right)^d, \tag{4.31}$$

with 2^{nd} order polynomial, $(d = 2)$. Finally, the projection of the test pattern $\mathbf{y}$ onto α is defined as

$$\prod(\boldsymbol{\alpha}, \Phi(\mathbf{y})) = \sum_{i=1}^{n_N} \boldsymbol{\alpha}_i k(\mathbf{x}_i, \mathbf{y}). \tag{4.32}$$

In addition, other linear and non-linear methods such as PCA, LDA, and Kernel Principal Components Analysis (KPCA) [60, 154, 98, 151] which are also efficient for dimensionality reduction and feature classification are utilised in the results section for the sake of comparison.

4.6 MATCHING AND DECISION PROCESS

After acquiring the new feature vector for the test image and projecting it using (4.32), then the decision is made depending on the nearest neighbour classification method defined as

$$\delta(\mathbf{p}, M^c) = \min_i \delta(\mathbf{p}, M^{c_i}) \to \mathbf{p} \in M^{c_N}, \tag{4.33}$$

where $\mathbf{p}$ is the test image new feature vector, M^c is the mean of training samples for the class M^{c_i} with $i = 1, 2, \cdots, N$ is the number of classes, and δ is the similarity measure.

For the similarity distance measure, the Mahalanobis Cosine similarity distance measures is used. δ_{mahcos} of $(\mathbf{x}_I, \mathbf{x}_Q)$ is defined as

$$\delta_{mahcos}(\mathbf{x}_I, \mathbf{x}_Q) = -\frac{(\mathbf{x}_I^T C_x \mathbf{x}_Q)}{\|\mathbf{x}_I\|_{C_x} \|\mathbf{x}_Q\|_{C_x}}, \tag{4.34}$$

where $(\mathbf{x}_I, \mathbf{x}_Q)$ are the enrolled and test feature vectors respectively, $C_\mathbf{x}$ is the covariance matrix of $(\mathbf{x}_I, \mathbf{x}_Q)$ in the Mahalanobis space and $\|\|_{C_x}$ is the norm operator for $\mathbf{x}_I$ and $\mathbf{x}_Q$ defined as

$$\|\mathbf{x}_I\|_{C_x} = \sqrt{\mathbf{x}_I^T C_\mathbf{x} \mathbf{x}_I}, \tag{4.35}$$

and

$$\|\mathbf{x}_Q\|_{C_x} = \sqrt{\mathbf{x}_Q^T C_\mathbf{x} \mathbf{x}_Q}. \tag{4.36}$$

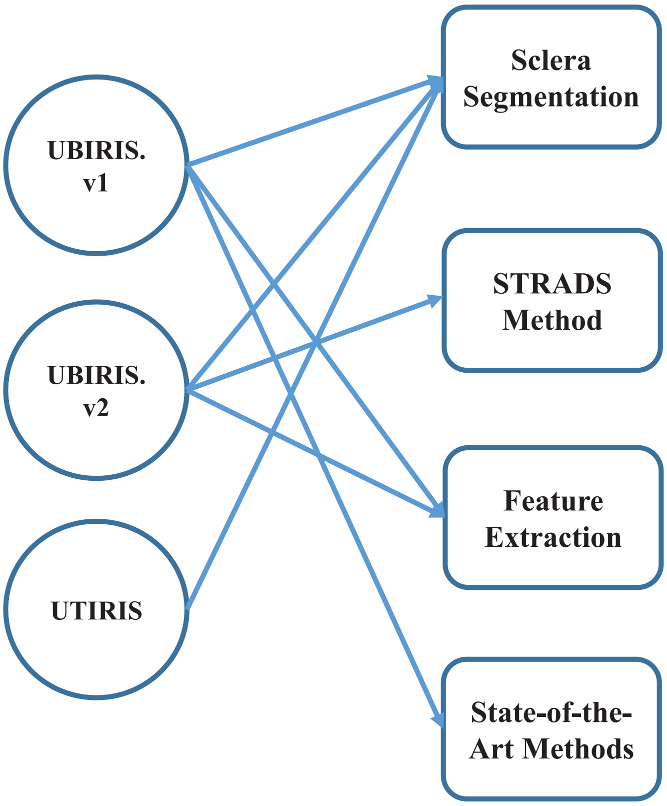


Figure 4.9: Database classification according to where it was utilised in the system.

4.7 RESULTS AND DISCUSSIONS

4.7.1 *Datasets Utilised in The Evaluation*

As mentioned earlier, there is no explicit publicly available database for the sclera, therefore most recent research on sclera recognition has used iris databases captured in the visible-wavelength illumination in order to evaluate the sclera recognition system. As this chapter proposes many resolutions for problems mentioned in Section 4.1, therefore each database should be classified in terms of what challenges do they employ and where it is suitable to be used in order to evaluate the proposed system. This is depicted in Fig. 4.9 and also summarised as

1. UBIRIS.v1 database: this is the most frequently used database in the recent research in order to evaluate a sclera recognition system. As this database has been captured in two capturing sessions in which, the first session was captured at minimum noise factors while the second session has introduced more natural

luminosity conditions adding several noise problems, therefore, it was utilised to evaluate sclera segmentation, feature extraction, and state-of-the-art methods. Because all the eye images are captured with frontal-gaze and fixed distance, the UBIRIS.v1 has not been used in STRADS method evaluation. This database has 1877 images and 72 images were discarded manually due to poor quality (blink, no sclera area).

2. UBIRIS.v2 database: this database has used an unconstrained image capturing scheme (at-a-distance and on-the-move) resulting in a highly challenging eye database. As a consequence of the capturing distances range from 4-8 meters at a rate of three images per meter and with three different iris gaze directions: F-G, R-G, and L-G, this database therefore requires more post-processing on removing noise factors, template alignment and rotation fitting design, hence it was used in the evaluation of sclera segmentation, STRADS, and feature extraction methods. No recent work on the sclera has been carried out using this database, therefore, the work in this chapter is the first to handle challenges introduced in this database. Only images captured from 4 to 6 meters and in both sessions with multiple gaze directions are used resulting in 3978 images as images captured at 7 and 8 meters have insufficient information for sclera regions.

3. UTIRIS database: this database consists of a high resolution 806 eye images captured in the visible-wavelength illumination. The issue with this iris database is that many images have been blurred in the sclera area to focus on the iris region. In addition, eye images were captured at a fixed distance and the database has not been used in recent sclera recognition research. Therefore, this database is only used in the evaluation of sclera segmentation.

Some examples of images in the above mentioned databases are shown in Fig. 4.10. Matlab (version R2014a) on a PC with Intel core i5 3.0 GHz processor and 8.0 GB RAM was used to implement and evaluate the algorithms described in this chapter.

(a) (b) (c)

Figure 4.10: Some examples of the eye images used in the proposed sclera system evaluations, (a) UBIRIS.v1, (b) UBIRIS.v2, and (c) UTIRIS sample images.

4.7.2 *Sclera Segmentation Evaluation*

The sclera segmentation accuracy is evaluated in terms of number of fused skin clusters and processing time. First, a ground truth is created for a total of 800 images which are randomly chosen and manually segmented and labelled for the UBIRIS.v1.S1, UBIRIS.v1.S2, UBIRIS.v2 and UTIRIS databases. Then, the accuracy of segmentation is measured pixel by pixel as

$$Ep = \sum_{k=1}^{n} \sum_{i=1}^{w_o} \sum_{j=1}^{h_o} bw_g\left(i,j\right) \oplus bw_t\left(i,j\right), \tag{4.37}$$

$$Acc = 1 - \left(\frac{Ep}{n \times w_o \times h_o}\right) \times 100\%, \tag{4.38}$$

where n is the number of images, Ep is the total number of incorrectly segmented pixels, bw_g and bw_t are the binary maps for the ground truth and segmented sclera image respectively.

The processing times for different image sizes are summarised in Table 4.1. The higher the resolution, the better the segmentation accuracy and with a higher processing time. Therefore, (300×400) resolution is chosen in the following evaluations for all databases as the segmentation accuracy is sufficiently high with low processing time.

In terms of skin clusters fusion threshold th_s in (4.3) and the evaluation shown in Fig. 4.11, it can be observed that for eye images captured under less constrained conditions, fusing at least any three detected skin clusters will achieve the highest segmentation accuracy

Table 4.1: Processing times for the proposed sclera segmentation using different image resolutions

No. of fused skin algorithms	Image size	Processing time	Accuracy 100%
3	600*800	3.70 s	95.37
3	300*400	0.82 s	94.22
3	150*200	0.29 s	92.97
3	75*100	0.12 s	89.70

(UBIRIS.v1.S1 Acc = 94.22 % and UTIRIS Acc = 95.99 %). On the other hand, images with severe noise factors like in UBIRIS.v1.S2 and UBIRIS.v2 have the highest segmentation accuracy if any of the skin clusters is detected. It is noteworthy that no sclera segmentation method will achieve perfect accuracy especially for pixel-wise thresholding methods as the pixel value could be altered by any noise

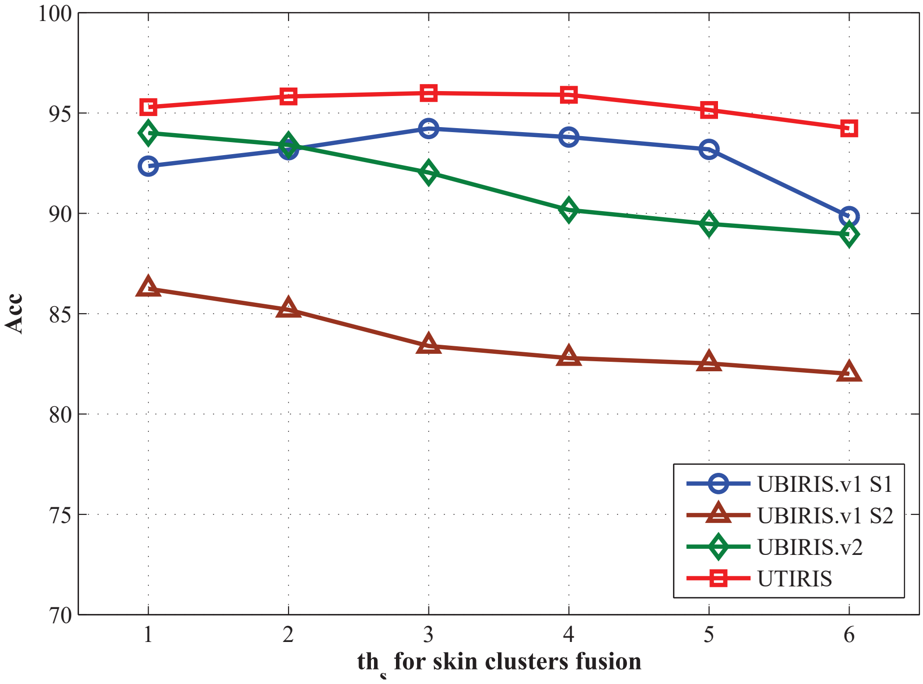
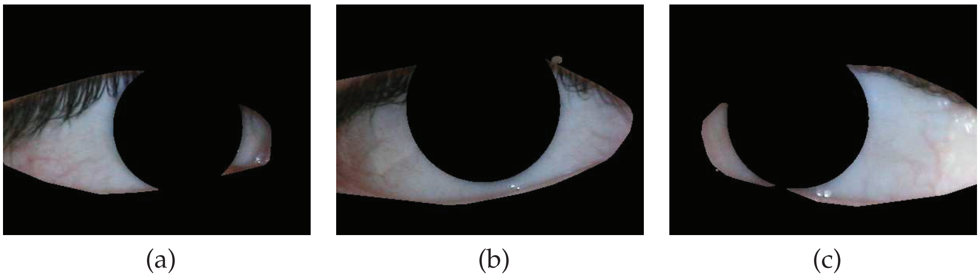


Figure 4.11: The accuracy evaluation of proposed sclera segmentation in terms of total fused skin clusters for the UBIRIS.v1, UBIRIS.v2 and UTIRIS databases.



(a) (b) (c)

Figure 4.12: Examples of the proposed sclera segmentation method using the UBIRIS.v2 database images captured at distance equals 5 meters. (a) R-G eye, (b) F-G eye, and (c) L-G eye.

factors or illumination variation. However, the proposed segmentation method has achieved promising accuracy using both constrained and unconstrained imaging conditions in the visible light. Some segmentation examples of the proposed sclera segmentation method are shown in Fig. 4.12.

4.7.3 *STRADS Method Evaluation*

The evaluation process for STRADS has been achieved in multi-session contexts with F-G, R-G, and L-G images. For each context, 2 images at a distance of 4 meters are used for training and 4 images at a distance of 5 and 6 meters are used for testing per user and gaze direction. In addition, the AdaBoost algorithm [155] is utilised to extract the ROI of the 6 meters distance image. Then, the proposed sclera segmentation along with DT-CWT and KLDA for blood vessel enhancement and feature extraction methods are used to evaluate STRADS performance.

Table 4.2: Comparison of recent state-of-the-art literature on blood vein recognition within the eye in terms of the availability of template rotation and distance scaling methods

Method	**Database (Sub./Img.)**	**Image Capturing Condition**	**Sclera Segmentation**	**Sclera Template Rotation Alignment**	**Sclera Distance Scaling**
Derakhshani et al. [17]	In-house DB(6/12)	Constrained	Manual	No	No
Derakhshani and Ross [71]	In-house DB(50/300)	Constrained	Manual	No	No
Crihalmeanu et al. [72]	In-house DB(50/100)	Constrained	Semi-automated	No	No
Thomas et al. [133]	UBIRIS.v1 (241/1205)	Constrained	Automated	No	No
Oh and Toh [78]	UBIRIS.v1 (241/1205)	Constrained	Automated	No	No
Zhou et al. [31]	UBIRIS.v1 (241/1805)	Constrained	Automated	No	No
Lin et al. [85]	UBIRIS.v1 (241/1168)	Constrained	Automated	No	No
Das et al. [86]	UBIRIS.v1 (241/1350)	Constrained	Automated	No	No
Alkassar et al. [19]	UBIRIS.v1 all UTIRIS	Constrained	Automated	No	No
Gottemukkula et al. [88]	In-house DB(226/904) UBIRIS.v1 (241/1205)	Constrained	Automated	No	Yes

Table 4.3: Comparison of EER using STRADS method for F-G, R-G, and L-G images at different distances on the UBIRIS.v2 database

Method	**With STRADS**		**Without STRADS**	
Gaze Direction	EER (%)			
	4m -5m training-testing	4m -6m training-testing	4m -5m training-testing	4m -6m training-testing
F-G	9.58	11.61	34.89	38.05
R-G	10.23	12.79	38.97	46.29
L-G	9.90	12.47	39.56	46.43

Recent work on the vascular pattern recognition within the sclera is shown in Table 4.2. Whereas the evaluation results are shown in Table 4.3 and the ROC curve is shown in Fig. 4.13. EER rates are relatively high when using noisy images with multiple gazes at different distances, as the UBIRIS.v2 database has many challenging issues resulting in high error rates, however, using the proposed sclera segmentation and STRADS methods decreased EER for F-G, R-G, and L-G images significantly. In addition, the run times of the STRADS method shown in

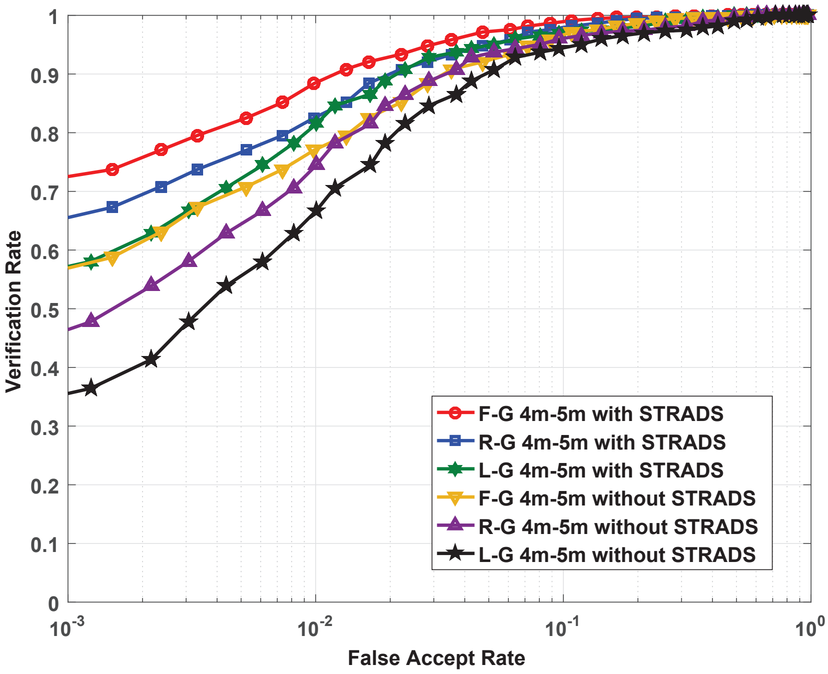


Figure 4.13: The UBIRIS.v2 ROC curve with and without the use of the STRADS method.

Table 4.4: The run times of the proposed Sclera Template Rotation Alignment and Distance Scaling STRADS technique

STRADS Steps	**Run Times**
Sclera Rotation Alignment	0.07 s
Sclera Distance Scaling	0.09 s

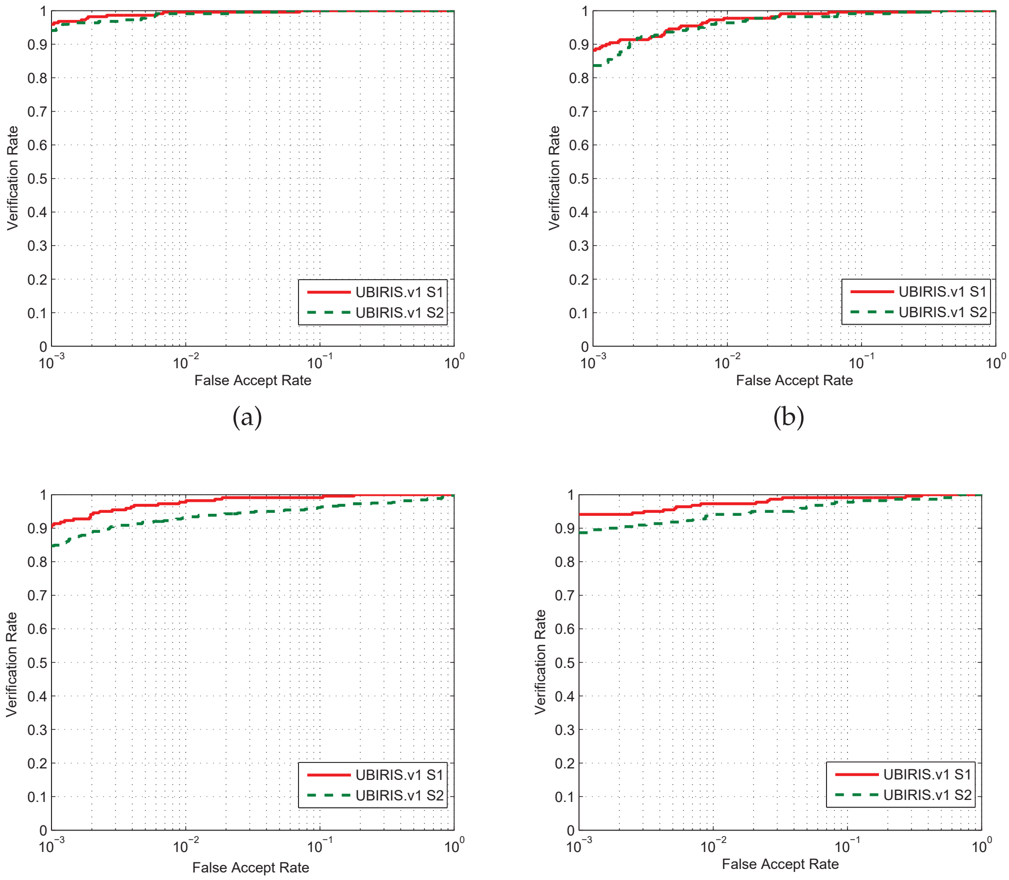
Table 4.4 reveal that the system complexity is dramatically reduced and thereby provides a practical solution for sclera recognition on the ubiquitous communication products with an integrated camera.

4.7.4 *Blood Vessel Enhancement and Feature Extraction Evaluation*

To create a ground truth of training-testing samples for evaluation, session 1 in the UBIRIS.v1 database is utilised as each user has 5 images. For all evaluation scenarios, k-fold cross validation with ($k = 5$) is used. As shown in Table 4.5, increasing the number of training samples will improve the system performance significantly while increasing the training time process. In fact, the training process is still a user-friendly process and a practical aspect as most recent real-time biometric systems such as the iPhone Touch ID uses many enrolment rounds to train fingerprint verification system and register the final

Table 4.5: Comparison of training-testing samples number for *complex-sclera* features with KLDA classification

No. of training-testing samples	**EER (%)**	**GAR (%)**	
		0.1 FAR	**0.01 FAR**
1-4	5.13	72.05	55.00
2-3	3.02	85.45	72.12
3-2	0.79	93.64	87.73
4-1	0.47	97.27	90.91



(a) (b)

(c) (d)

Figure 4.14: ROC curves for blood vessel enhancement and feature extraction methods with KLDA, (a) proposed DT-CWT method, (b) DWT method, (c) CDF 9/7 method and (d) GWT method.

fingerprint for higher recognition accuracy. Thus, it is feasible to design a sclera biometric system with a training scheme.

Next, *sclera-complex* features resulting from the proposed blood vessels enhancement are evaluated on the UBIRIS.v1 with DWT, CDF 9/7 and GWT using 3 samples for training and 2 for testing for each session. As level 4 of the DT-CWT has empirically introduced the lowest error rates, therefore one level only for the DWT, CDF 9/7, and GWT which produce the lowest EER is used for the fairness of comparisons. The ROC curves are plotted in Fig. 4.14. As shown in Table 4.6, DT-CWT has the lowest EER compared to DWT, CDF 9/7, and GWT in both sessions (EER = 0.79 % and EER = 0.91 % for session 1 and session 2 respectively). However, the processing time is slightly lower (0.083 s) for DWT which is comparable to the proposed DT-CWT.

Table 4.6: EER rates for the proposed blood vessels enhancement method compared to GWT, CDF 9/7, and DWT methods

Method	**EER (%)**		**Processing time**
	S1	**S2**	
GWT	1.5	2.27	0.847 s
CDF 9/7	1.73	5	0.095 s
DWT	2.29	2.27	0.083 s
DT-CWT	0.79	0.91	0.121 s

In addition, the classification method KLDA has been compared in sclera recognition with linear and non-linear methods such as PCA, LDA, and KPCA in Table 4.7. KLDA has superiority over PCA and LDA whereas KPCA is comparable to KLDA.

4.7.5 *A State-Of-The-Art Sclera Recognition Method Comparison*

The results of the proposed work are evaluated and analysed with recent state-of-the-art methods on sclera recognition and summarised in Table 4.8. The proposed sclera segmentation for constrained and unconstrained imaging conditions has achieved high accuracy rates and it is noteworthy that the recent methods utilised only session 1 of the UBIRIS.v1 database for evaluating sclera segmentation and no

Table 4.7: EER comparison of PCA, LDA, KPCA, and KLDA for the sclera classification process

Method	**EER (%)**	
	S1	**S2**
PCA	11.81	12.72
LDA	4.5	5.46
KPCA	2.27	2.52
KLDA	0.79	0.91

Table 4.8: A state-of-the-art comparison of recent work on the UBIRIS.v1 database

Method	**EER (%)**		**EER (%)**	
	S1	**S2**	**S1 training S2 testing**	**S2 training S1 testing**
Thomas et al. [133]	3..38	na	na	na
Oh and Toh [78]	0.47	na	na	na
Zhou etal [31]	4.09	9.98	na	na
Lin et al. [85]	3.05	na	na	na
Das et al. [86]	0.42	0.51	3.95	4.34
Alkassar et al. [19]	2.19	2.67	3.68	4.11
Proposed method	0.47	0.50	2.23	2.87

segmentation accuracy rates are reported except for the processing times. In addition, the EER performance of the designed system is significantly lower with reduced complexity compared to discrete and Gabor wavelets used by authors in Table 2.2 and thereby supporting the chapter hypothesis that the proposed *complex-sclera* features method is appropriate for real-time applications as it has low EER rate and uncomplicated hardware implementation.

4.7.6 *Empirical Parameters Tuning Analysis*

There are two parameters which control the sclera segmentation process for the skin and sclera classifiers. th_s determines how many skin classifiers will be used when creating the skin map and is evaluated in Fig. 4.11. Whereas th_I controls the sclera area binary conversion where a higher value leads to segment more inaccurate sclera regions.

4.8 CHAPTER SUMMARY

In this chapter, the aim was to assess sclera biometric system challenges when relaxed imaging conditions are introduced. The chapter first identified the challenges of segmenting the sclera area, enhancing blood vessel separation from its background, extracting distinctive features, and obtaining an efficient matching scheme. Therefore, the chapter has proposed a novel sclera segmentation method which is adaptive to the noisy images. Sclera segmentation has been achieved through an adaptive fusion method which involves combining many skin classifiers and was performed first to increase skin pixel detection and overcome various image noise factors. In addition, new methods for sclera template rotation alignment and distance scaling have been suggested to overcome the change in rotation and distance when an eye image is captured. For blood vessel enhancement and feature extraction, a new sclera features scenario was proposed denoted as a *complex-sclera* approach to achieve low error rates and decrease the complexity. The experimental results on the UBIRIS.v1, UBIRIS.v2 and UTIRIS databases showed that the proposed system outperforms the state-of-the-art sclera recognition systems. Furthermore, the results using the UBIRIS.v2 database which has many challenging aspects showed that the STRADS method has a low processing time and reduces the error rates dramatically thus introducing more robustness.

As the sclera area mainly depends on the accurate eye localisation, eye gaze direction, and also the quality of an eye image used in the sclera biometric system could affect the system performance, the design of an unsupervised system which handles the above mentioned challenges is substantial. Therefore, the next chapter will discuss the implementation of automatic eye localisation, eye corner and gaze detection methods along with an efficient image quality measure.

5

OVERCOMING CHALLENGES IN SCLERA RECOGNITION SYSTEMS UNDER VARIOUS IMAGING CONDITIONS FOR REAL-TIME APPLICATIONS

5.1 INTRODUCTION

In the previous chapters, novel methods for tackling challenges in the design of sclera recognition systems under constrained and relaxed imaging conditions were proposed which consisted of designing efficient sclera segmentation, blood vessel enchantment, feature extraction and final template production. Both sclera recognition systems in Chapters 3 and 4 as well as in the state-of-the-art systems have used supervised techniques in one or more aspects of the design. Therefore, and as the design should be applied in a real life scenario, where human intervention should be minimised, some issues emerged which consequently are addressed in this chapter.

This chapter proposes the following novel contributions to mitigate real-time limitation points (i-v) mentioned in subsection 2.3.2 on sclera recognition systems summarised as

- A practical eye localisation method which enhances the AdaBoost eye detection;
- A novel fusion method for eye corner detection which is adapted to noise factors;
- An efficient eye gaze detection method which is invariant to eye rotation;
- A sclera image quality measure and classification methods based on a sclera image focus measure.
- Successful operation with sclera images captured at-a-distance, on-the-move, using ubiquitous camera devices and under natural illumination conditions using ubiquitous camera-based devices.

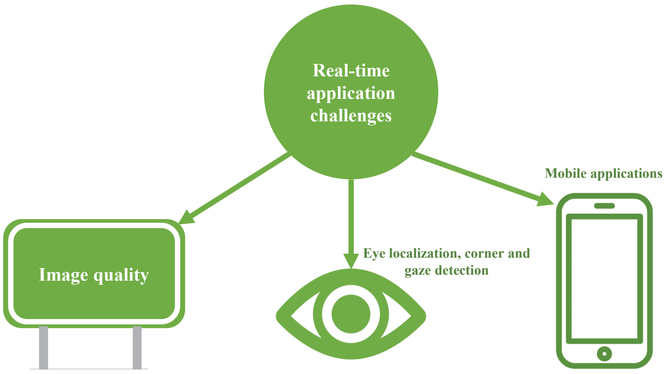


Figure 5.1: Issues and challenges for sclera recognition in a real life scenario.

This chapter will first discuss a solution for accurate eye localisation. Then, a robust method for automatic eye corner and gaze detection is proposed which involves two levels of eye corner verification to minimise eye corner point misclassification when noisy eye images are introduced. Moreover, gaze direction estimation is achieved through the pixel properties of the sclera area. Next, a new method for sclera image quality measure is achieved based on a focus measure. Then, the Excellent, Good, Bad, and Ugly (EGBU) classification technique is exploited on eye images to determine their quality in order to evaluate system performance. Finally, the evaluation and results of the proposed methods are introduced in the results and discussions section.

5.2 ACCURATE EYE LOCALISATION

The first challenge when designing a real time unsupervised sclera recognition system is locating the eye. Normally in the state-of-the art methods in sclera recognition, either the images utilised are eye cropped or the user has to follow certain rules in the image capturing session in order to capture eye regions accurately. These two cases are accepted as the design process is at early stages. However, and as the aim is to design a real-time application where the user may not cooperate perfectly, this could affect the system performance. Therefore, designing an efficient eye localisation system which is accurate and fast is necessary.

As shown in Algorithm 5.1, the process is initiated by taking the first frame from the camera video feed and apply preprocessing steps such

Algorithm 5.1 Real-time eye localisation method

```
set camera resolution to $352 \times 288$;
reset frame loop-count and feature-count;
while True && loop-count $\leqslant 300$ do
    if feature-count < 200 then
        apply eye detector on the frame to acquire the box coordinates;
        if box coordinates not empty then
            get box width and height;
            set min and max iris radius between width/10 and width/3 respectively;
            apply MinEigenFeatures method to extract tracking points;
            if points detected then
                update feature-count;
                apply integro-differential operator;
                if iris detected then
                    return frame and exit;
                else
                    move to the next frame;
                    apply tracking;
                end if
            end if
        end if
    end if
end while
```

as resizing the input frame to enhance system speed. After that, the AdaBoost eye detection method is applied to extract the anticipated eye region. As the extracted box from the AdaBoost algorithm may not contain the eye, therefore, first the minimum eigenvalue features method [157] is applied to extract tracking points within the eye area. As a result, the algorithm will be trained to detect specific features within the eye and hence, increase eye detection. Then, the integro-differential operator is applied to check if there is a valid iris area within the box. The process continues until an eye area is correctly extracted.

5.3 EYE CORNER AND GAZE DETECTION

5.3.1 *ROI Extraction*

Normally, when capturing an eye image and after applying the AdaBoost detection [155] to crop the eye region, the cropped eye image will consist of the iris, pupil, sclera and eyelid skin areas surrounding the eye. However, the variation in the capturing distance and illumination condition in a relaxed imaging scenario could lead the AdaBoost detection method to crop more area around the eye. Hence, the segmentation, rotation alignment, and distance scaling of sclera regions along with estimating eye corner positions could be affected. Therefore, it is necessary to define the region where the eye corner position is anticipated and hence, improve the system performance.

To avoid detecting redundant eye corners, a method is proposed to extract the ROI area where eye corners are predicted. Based on the skin segmentation method proposed in subsection 4.2.1, two diameter arc areas are extracted to estimate the position and the orientation of the eyelids. As depicted in Fig. 5.2, the two arc area pixels are defined as

$$\mathbf{Arc} = \mathrm{RGB}\left(x_0 + r\cos\theta, y_0 + r\sin\theta\right), \tag{5.1}$$

where $\theta \in [-5\pi/12, 5\pi/12] \cup [-7\pi/12, 7\pi/12]$ with uniform increment steps of 0.1 degrees and (r, x_0, y_0) are the iris radius and centre coordinates calculated by using the integro-differential operator [32]. Then, both arcs are divided equally at angles $(0, \pi)$ into up and down parts and a circular search for skin pixels is applied on these arcs as depicted in Fig. 5.2b. The circular search is initiated by taking a number of pixels Pn at a time in the direction $[0 \rightarrow 5\pi/12, 0 \rightarrow -5\pi/12]$ for the right arc and $[\pi \rightarrow 7\pi/12, \pi \rightarrow -7\pi/12]$ for the left arc in order to mitigate incorrect eyelid positioning, as searching the arcs pixel by pixel is not efficient. The right arrow $\rightarrow$ denotes the direction of the circular search. Each segment of the Pn pixels is checked using the skin classifier as follows

i. take the first segment in the directions depicted in Fig. 5.2b;

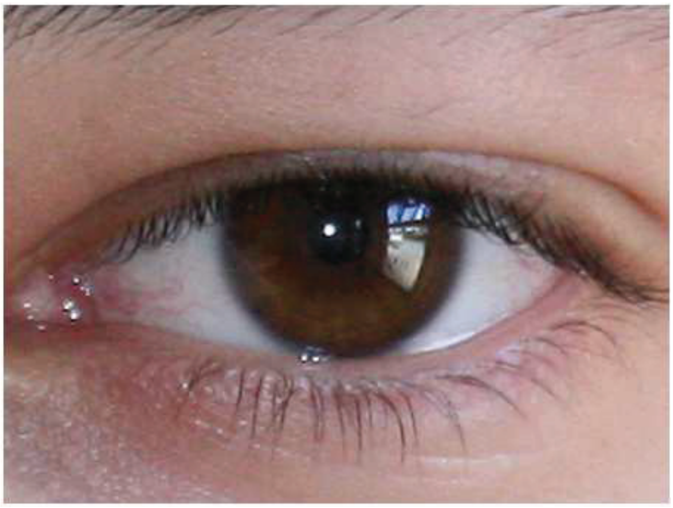
ii. for the right and left arcs, up and down partitions, check if the selected segment has full skin pixels and then, store the first pixel coordinates in that segment and stop the search. Otherwise, move to the next segment;

iii. check the distance between the detected points in the up and down partitions for the right and left arcs. If the Euclidean distance is less than th_e, for both arcs, consider the eye is partially closed and terminate;

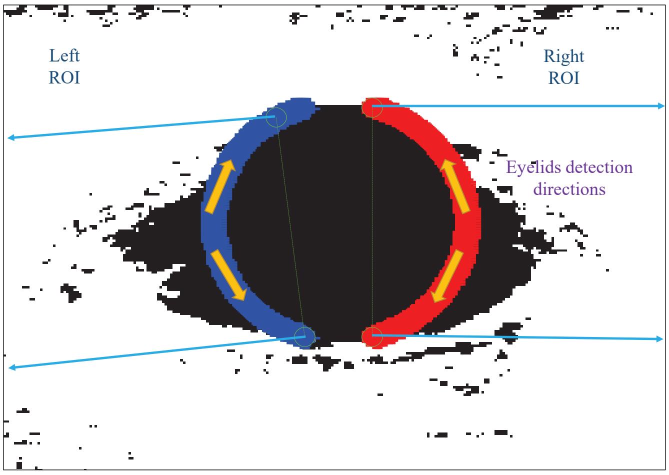
iv. otherwise, return the detected eyelid points;

where Pn and th_e are set empirically to 100 and 25 respectively.

The algorithm will return a decision whether the eye is partially occluded or the eyelids position is determined and the ROI is therefore cropped resulting in an efficient method for ROI extraction, eye occlusion, and eyelids position.



(a)



(b)

Figure 5.2: Eye corner ROI areas detection, (a) an eye image from the UBIRIS.v2 database captured at a distance of 4 meters and (b) the skin segmentation map where R-ROI and L-ROI areas are extracted.

5.3.2 *Eyelids Border Detection and Skin Segmentation*

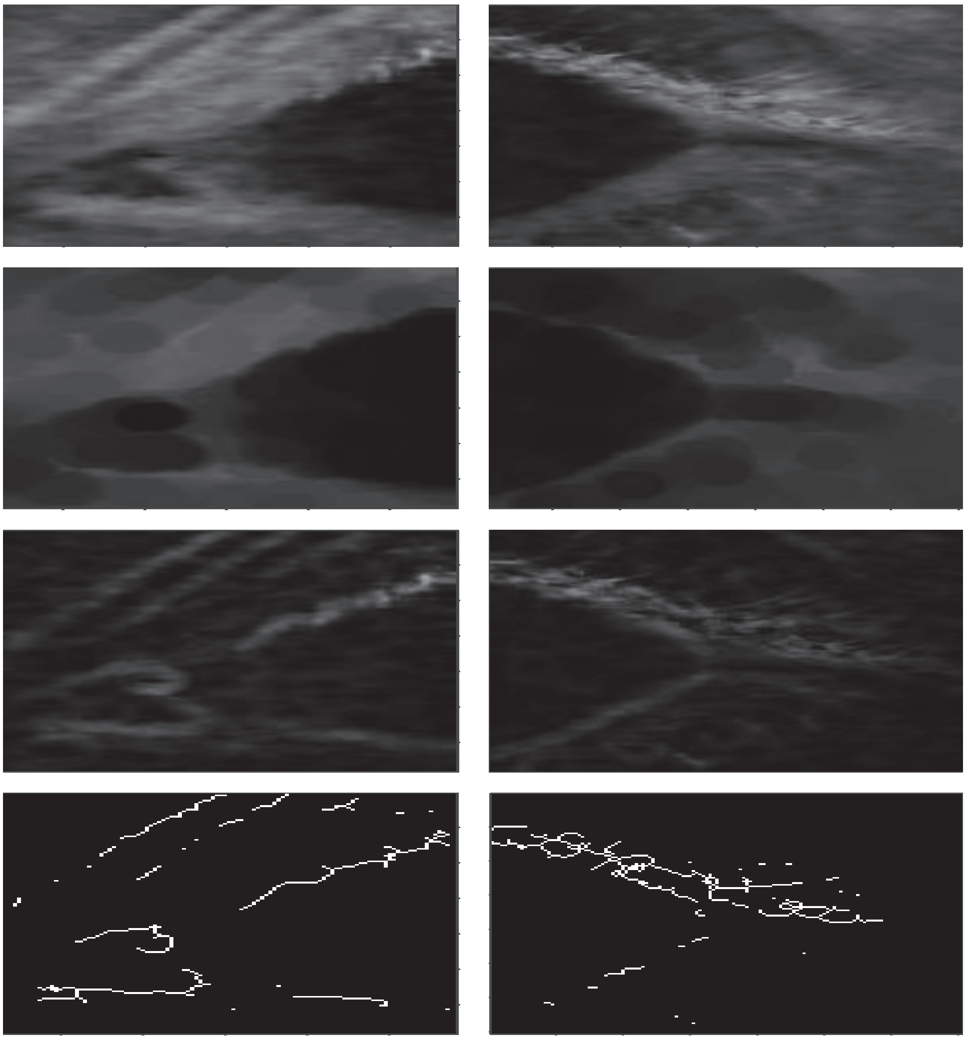


Figure 5.3: Eyelids border detection method. The top row represents $\mathbf{R}\mathtt{s}$ and $\mathbf{L}\mathtt{s}$ maps, the next row shows $\mathbf{R}_{\mathtt{er}}$ and $\mathbf{L}_{\mathtt{er}}$ maps, third row illustrate the subtracted $\mathbf{B}\mathtt{r}$ and $\mathbf{B}\mathtt{l}$ maps, whereas the bottom row shows the skeleton form of the final $\mathbf{B}\mathtt{fr}$ and $\mathbf{B}\mathtt{fl}$ maps.

The next step is to extract the skin map and eyelid border of the R-ROI and L-ROI regions. A fusion method is designed to extract corner points by using two maps. The first map represents eyelid border outlines whereas the second map consists of the eyelid skin surrounding the sclera region which is already explained Chapter 4. Therefore, only the eyelid border map extraction will be presented in this chapter.

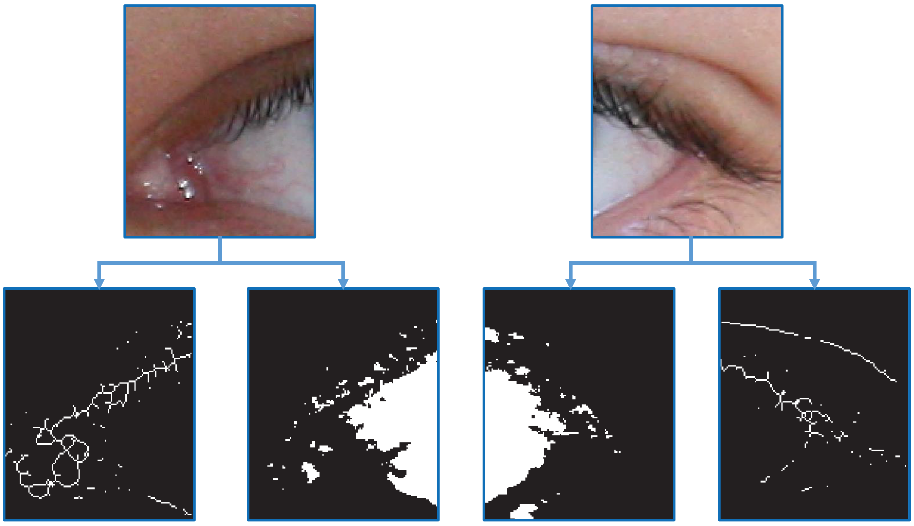


Figure 5.4: R-ROI and L-ROI regions each producing two maps for eyelid skeleton border **B**f and eyelid skin area **Sk**.

To extract the eyelid border, the saturation level S of the HSV colour space for R-ROI and L-ROI areas are extracted resulting in **R**s and **L**s maps. Next, an erosion filter with structure element having a disk shape with size 5×5 is applied on **R**s and **L**s to eliminate white pixels within the sclera as the sclera normally has the darkest pixels in the S level resulting $\mathbf{R}_{er}$ and $\mathbf{L}_{er}$. Then, $\mathbf{R}_{er}$ and $\mathbf{L}_{er}$ are subtracted from **R**s and **L**s to extract the border outlines as

$$\mathbf{B}r = \mathbf{R}s - \mathbf{R}_{er}, \tag{5.2}$$

$$\mathbf{B}l = \mathbf{L}s - \mathbf{L}_{er}. \tag{5.3}$$

Next, binary thresholding is applied on **B**r and **B**l maps as

$$\mathbf{B}f = \begin{cases} 1, & \text{if } \mathbf{B}(x,y) > th_i \\ 0, & \text{otherwise} \end{cases}, \tag{5.4}$$

where th_i is set empirically to 0.2. Then, binary morphological operations which include removing spur pixels and converting the borders into skeleton form are utilised. This is shown in Fig. 5.3.

The second map is extracted by complementing the extracted skin segmentation R-ROI and L-ROI maps producing **Sk**r and **Sk**l. The final four maps are shown in Fig. 5.4.

5.3.3 *Candidate and Final Eye corner Points Extraction and Gaze Estimation*

Next, candidate corner points in $\mathbf{B}f_r, \mathbf{B}f_l, \mathbf{Sk}r$, and $\mathbf{Sk}l$ are extracted. First, the Harris corner detector [110] is utilised to extract corners with a quality level and sensitivity factor equal to 0.3 and 0.01 respectively. To eliminate outlier detected corners, an angle threshold is set with respect to the iris centre position over $[-\pi/6, \pi/6]$ providing 60° of rotation sensitivity as depicted in Fig. 5.5.

Next, to estimate the correct corner among all the prospective corner points, the next step is calculating the Euclidean distance between all points and returning the point index which has the lowest distances as the red corner points depicted in Fig. 5.5. If the set of prospective corners is $\mathbf{c} = \{\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_n\}$, then the estimated corner point is specified as

$$\mathbf{c}_i = \underset{i}{\arg\min} \left\{ \sum_{j=1}^{n} \left\| \mathbf{c}_i - \mathbf{c}_j \right\| \right\}, \tag{5.5}$$

where $i \in \{1, \cdots, n\}$ and the symbol $\|.\|$ denotes the Euclidean distance. As the ROI extraction method ensures that the corner point will be roughly in the middle, then the estimated corner point will have the minimum Euclidean distance with respect to all prospective corner points. Prospective corner points and the final corner point are shown in Fig. 5.6.

After detecting the corner point $\mathbf{c}_i$ for eyelid skin segmentation and border maps in both right and left eye sides $\left(\mathbf{c}_{\mathbf{B}f_r}, \mathbf{c}_{\mathbf{B}f_l}, \mathbf{c}_{\mathbf{Sk}r}, \mathbf{c}_{\mathbf{Sk}l}\right)$, final corner points $(\mathbf{c}f_r, \mathbf{c}f_l)$ are calculated as the points in the middle of the straight lines which connect $\mathbf{c}_{\mathbf{B}f_r}$ with $\mathbf{c}_{\mathbf{Sk}r}$ and $\mathbf{c}_{\mathbf{B}f_l}$ with $\mathbf{c}_{\mathbf{Sk}l}$ respectively.

The final step is to estimate the eye gaze direction. The distribution of eye pixels from $\mathbf{c}f_r$ to $\mathbf{c}f_l$ of the eye blue channel in the RGB colour space is extracted, as this channel introduces more contrast between the iris and the sclera area. Values of the sclera pixel distribution are validated first to ensure that the detected pixel distribution belongs to the sclera regions. This is achieved by calculating the mean value and comparing it with the iris, and eyelids skin mean values. Then,

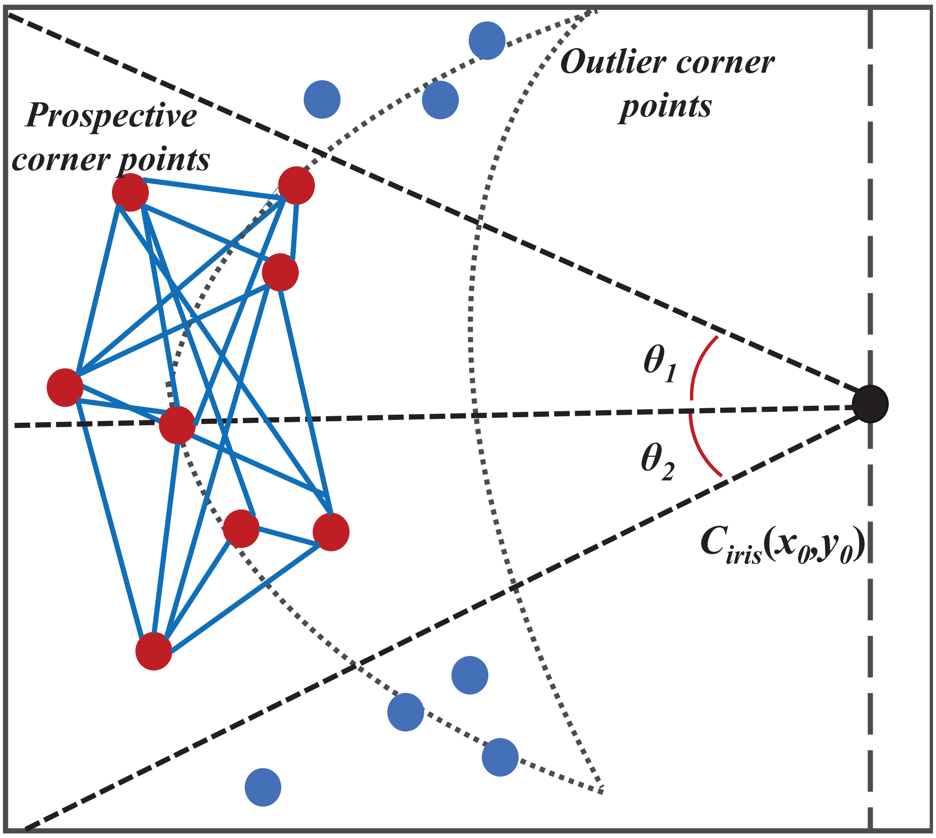


Figure 5.5: A diagram representing the proposed angle threshold with respect to the iris centre to eliminate outlier corners represented in blue dots and showing a Euclidean distance calculation scheme among prospective corners represented in red dots.

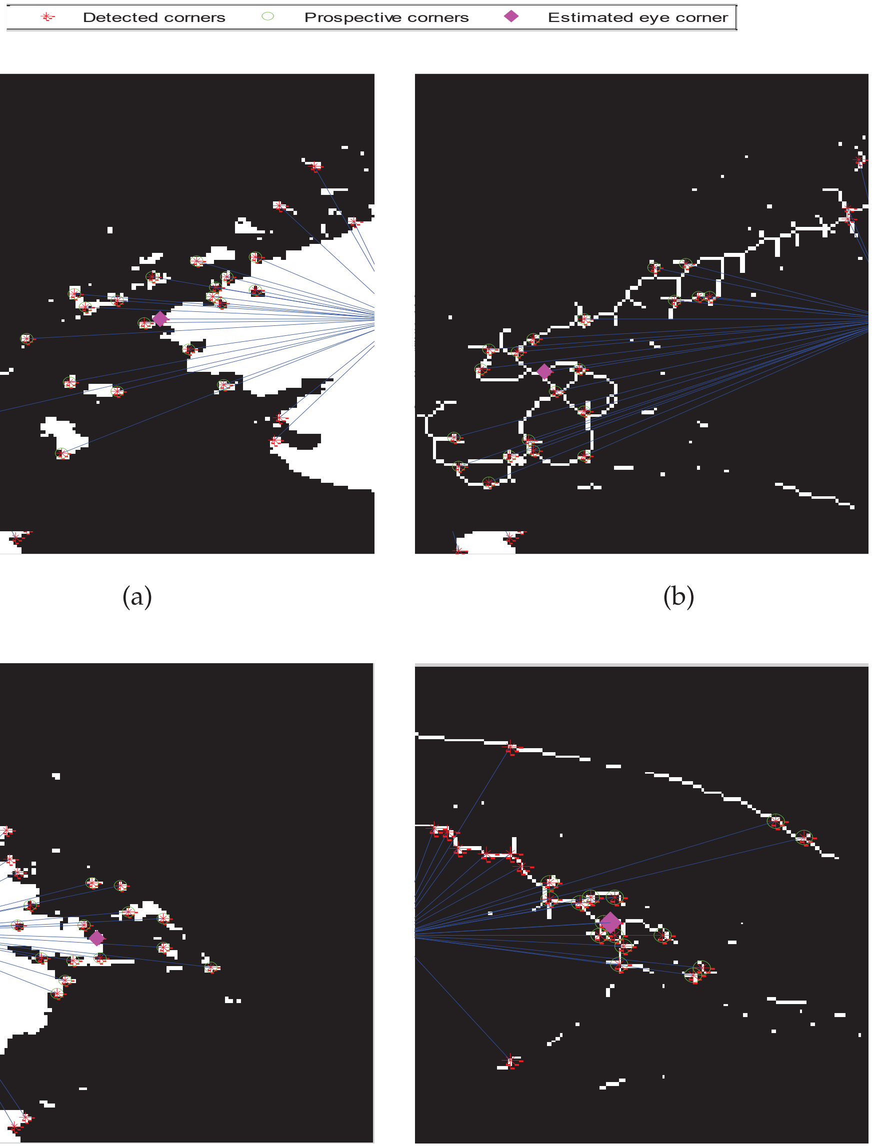
according to the width of right and left sclera sides, the decision of eye gaze is defined as

$$\text{Gaze} = \begin{cases} \text{R-G}, & \text{if } (w_l - w_r) \geqslant th_r \\ \text{L-G}, & \text{if } (w_r - w_l) \geqslant th_l \\ \text{F-G}, & \text{otherwise} \end{cases} \tag{5.6}$$

where w_r and w_l are the width of right and left sclera areas respectively calculated from the detected corners to iris borders and (th_r, th_l) are the thresholds to determine the eye gaze direction defined as

$$th_r = 2w_l/5, \; th_l = 2w_r/5, \tag{5.7}$$

which control the sensitivity of the eye gaze detection. Examples of an eye gaze detection using eyelid border, eyelid skin and fusion methods and their pixels distribution are shown in Fig. 5.7.



(a) (b)

(c) (d)

Figure 5.6: Eye corner detection for $\mathbf{B}f_r, \mathbf{B}f_l, \mathbf{Sk}r$ and $\mathbf{Sk}l$ for an eye image from the UBIRIS.v2 database at a distance of 4 meters showing the prospective and final corner points. (a) $\mathbf{Sk}l$ mask, (b) $\mathbf{B}f_l$ mask, (c) $\mathbf{Sk}r$ mask and (d) $\mathbf{B}f_r$ mask.

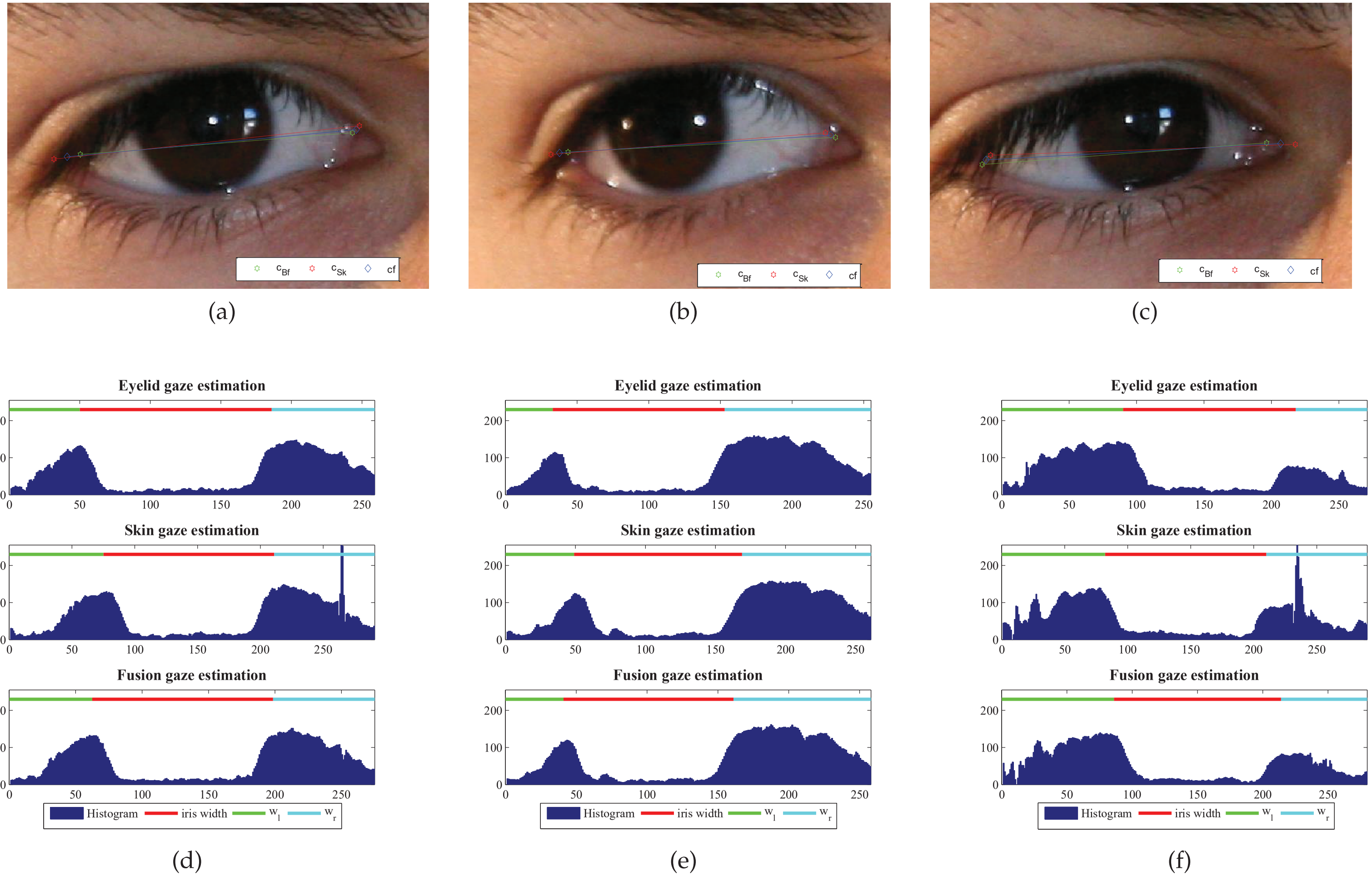


Figure 5.7: Examples of the eye gaze detection method. (a), (b) and (c) are eye images with F-G, L-G and R-G of the same individual captured at 5 meters from the UBIRIS.v2 database; (d), (e) and (f) are the pixel value distribution between detected corners utilising eyelid border, eyelid skin, and the fusion methods respectively.

5.3.4 *Improving Sclera Segmentation and STRADS Methods using Extracted Corner Points*

As mentioned in the previous section, determining eye gaze direction has a crucial role in sclera template segmentation and the design of alignment algorithms. Although two methods for sclera segmentation have been proposed for various imaging conditions in the previous chapters, both methods however need human intervention regarding the eye gaze direction. Therefore, applying automatic eye corner and gaze detection methods will improve the segmentation and alignment processes significantly and introduce a reliable system for unsupervised applications.

For the sclera segmentation using the proposed method in subsection 3.2.3, the eye gaze detection method will specify how many initial seeds of the active contours are required. While the sclera segmentation method proposed in Section 4.2, and although the segmentation method itself does not need an eye gaze decision before the segmentation, it does however use the direction of the eye gaze. The gaze direction is utilised to determine the labelling points required for the map refinement post-processing.

On the other hand, the STRADS method which is invariant to pose rotation and capturing distance uses the eye gaze decision to rotate and align the right, left, or both sides of the sclera region. In addition, eye corner positions are exploited in the alignment process as depicted in Fig. 5.8 where each part of the sclera region is moved to the corner position first before the rotation alignment process is initiated providing more robustness to the STRADS method.

5.4 SCLERA IMAGE QUALITY MEASURE

In any biometric system, failing to check the quality measure of an input image could affect the entire processing flow thus resulting in high error rates. As images involved in the sclera recognition are not necessarily captured under constrained imaging conditions, therefore the quality of the utilised images must be examined in order to have more insight about the suitability of these images for the sclera recognition process.

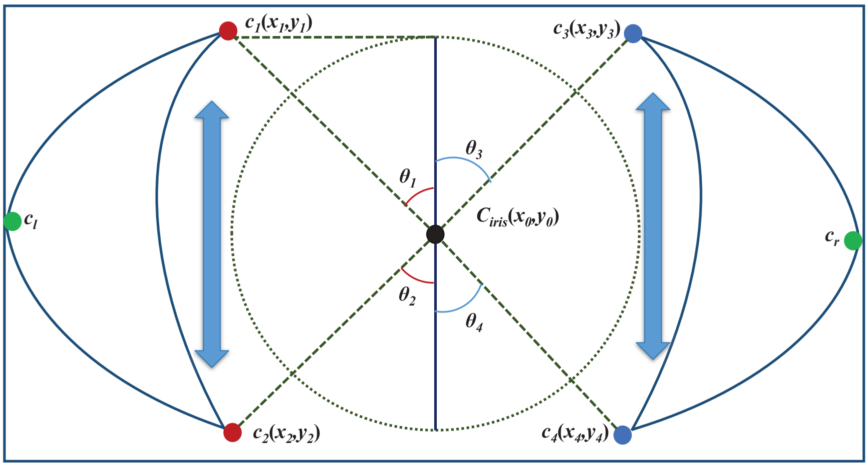


Figure 5.8: Using eye corner points shown as green dots to align the sclera regions to a fixed position before the rotation alignment process.

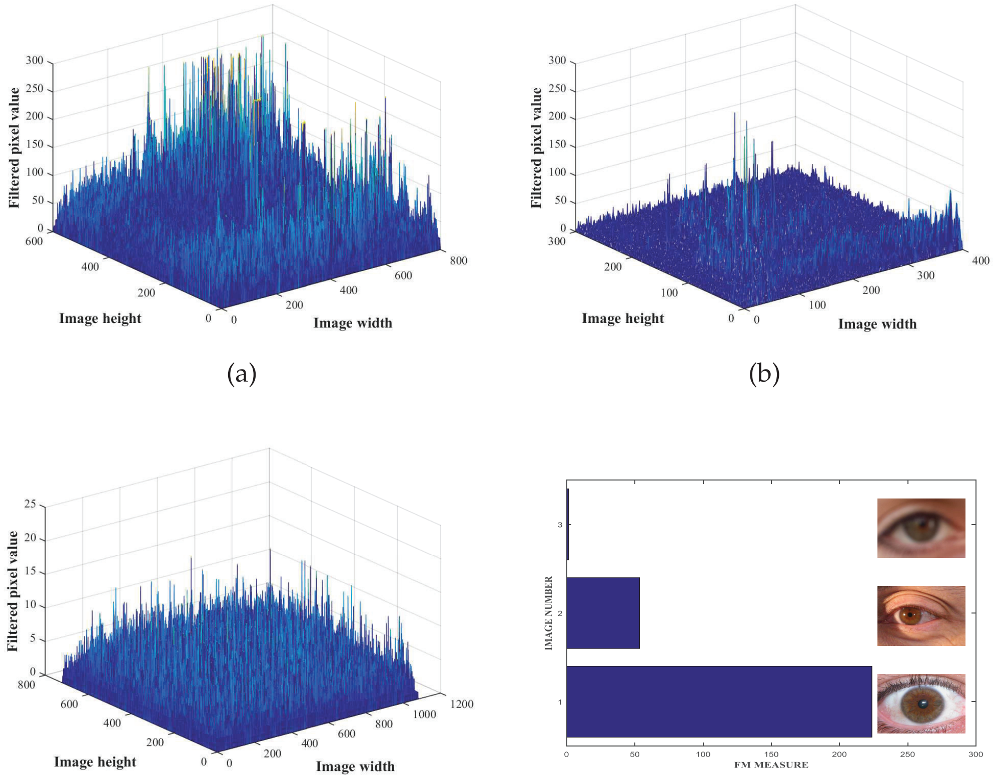
The quality of the utilised image is first classified into four categories: EGBU images inspired by the research by Phillips et al. [158] where face images are classified based on their matching difficulties into three categories only which are the Good, Bad, and Ugly (GBU) whereas in this chapter, another category where the Excellent quality category is added to investigate the challenges of different image qualities.

In order to classify each image to EGBU categories, a focus measure (fm) is proposed to determine sharp and smooth edges of an input image and hence, define its quality accordingly. For a real-time application, the employed quality measure must be efficient in terms of processing time and computational load to avoid adding extra expensive processing time and introducing an impractical sclera recognition system. To obtain fm, the second derivative method proposed in [159] is used first to detect the sharp and smooth edges within the input eye image as it passes the high spatial frequencies associated with sharp edges.

Let the intensity of an input image be $I(x,y)$, then fm is defined as

$$fm(I) = \sum_{i=1}^{x} \sum_{j=1}^{y} \left| L(x,y) - \overline{L} \right|^2, \tag{5.8}$$

where $L(x,y)$ is the convolution of $I(x,y)$ with the Laplacian operator L to detect I edges which is defined as



(a) (b)

(c) (d)

Figure 5.9: Examples of three different quality images where the convolution with the Laplacian operator and the final fm value are shown. (a), (b), and (c) are L outputs of images from the UBIRIS.v1, UBIRIS.v2, and MICHE databases respectively, and (d) shows their fm values.

$$L = \frac{1}{6}\begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix}, \tag{5.9}$$

and $\overline{L}$ is the mean of the absolute value calculated as

$$\overline{L} = \frac{1}{xy}\sum_{i=1}^{x}\sum_{j=1}^{y}\left|L\left(x,y\right)\right|. \tag{5.10}$$

High and low frequency edges of three different quality images are shown in Figs. 5.9a, 5.9b, 5.9c, whereas examples of the fm for the same images are shown in Fig. 5.9d. For high resolution images where the edges are strong, the variance between image pixels is high in which the image texture is clear whereas for low resolution images, most of the image pixels present low frequency edges where only a few pixels have a high edge value. While for the blurry image, the majority of pixels have the same low edge value resulting in smooth edges as each pixel value is subtracted by $\overline{L}$. The quality based on EGBU classification will be assigned to each image according to some empirical results of the utilised datasets which will be discussed in subsection 5.5.5.

5.5 RESULTS AND DISCUSSIONS

5.5.1 *Datasets Utilised in The Evaluation*

As the previous chapters have used the UBIRIS.v1 and UBIRIS.v2 datasets to evaluate the proposed sclera recognition methods, these datasets will therefore be utilised in this chapter. In addition, and to observe sclera recognition in a real-time scenario, images from two mobile devices and one webcam are used in the evaluation. Some properties of these databases and the camera devices used to capture their eye images are summarised in Table 5.1. In addition, examples of collected images for the aforementioned databases are shown in Fig. 5.10. The following are a brief description of databases used to evaluate the proposed system:

i. The UBIRIS.v1 database [96]: This database is used as a quality reference only for the Excellent classification. It is noteworthy that the UBIRIS.v1 database has another version of image resolution which is 200 ×150. However, the images with the aforementioned resolution have not used and hence, this resolution is ignored in Table 5.1.

ii. The UBIRIS.v2 database [112]: As this database has image conditions include blurring, motion, different capturing distance, multiple eye gaze directions, eye-glasses images, and different illumination scenarios, it has been utilised in the evaluation of image quality measure, eye gaze and corner detection. 100 images per distance and gaze are selected randomly resulting in 1500 images. Then, these images are manually annotated for eye corners for each image in order to create ground-truth comparison samples.

Figure 5.10: Examples of various eye databases under visible wavelength used to evaluate the proposed sclera recognition methods. The top row represents UBIRIS.v1 database samples, the middle row represents the UBIRIS.v2 database, the third row represents the MICHE database,while the bottom row represents the In-house Webcam database.

iii. The Mobile Iris CHallenge Evaluation (MICHE) database [114]: Due to the low resolution of the front cameras of iPhone 5, Samsung Galaxy S4, and Samsung Galaxy Tab 2 (as Samsung Galaxy Tab 2 does not have a rear camera), therefore, images captured with these cameras have been discarded resulting in processing only 1279 images.

iv. In-house HD Webcam database: In order to check the eye localisation method in a real-time scenario, and as the above mentioned datasets have static images, an eye image acquisition system is suggested. Therefore, an HD webcam and PC are chosen and hence, most image acquisition systems are covered in this chapter which are shown in Table 5.1. The webcam is a Logitech HD webcam c920 fixed on tripod as shown in Fig. 5.11.

5.5.2 *Eye Localisation Evaluation*

The eye localisation method is tested in a real-time scenario at the PhD study room during the daylight with all the natural reflections and noise factors. As shown in Fig. 5.12a, the user should sit in front of

Table 5.1: Brief description for various eye databases captured under visible wavelength in terms of capturing devices and image properties

Dataset	**Camera Type**	**Image size and Resolution**	**Capturing Distance**
UBIRIS.v1 [96]	Nikon E5700 5 MP	600×800 300 dpi	20cm
UBIRIS.v2 [112]	Canon EOS 5D 13.3 MP	300×400 72 dpi	5m-8m
MICHE iPhone 5 [114]	iSight with 8 MP	1536×2048 72 dpi	13cm
MICHE Samsung Galaxy S4 [114]	CMOS with 13 MP	2322×4128 72 dpi	13cm
In-house Webcam	Logitech HD Webcam c920 15 MP	1920×1080 30 fps[a]	30cm

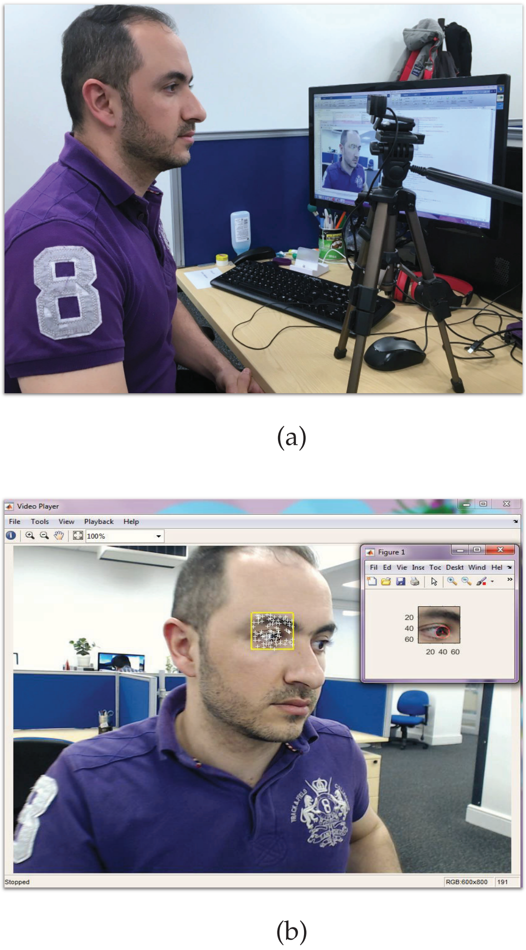
[a] fps: frame per second

Figure 5.11: The Logitech HD c920 webcam setup which is fixed on a tripod for real-time image acquisition.

the webcam while the eye localisation is initialised. Then, the process continues until an eye image which consists of a valid eye is captured as shown in Fig. 5.12b or the threshold of the frame count is reached.

5.5.3 *Eye Corner Detection Evaluation*

Simulations were performed using Matlab (version R2016a) on a PC with Intel core i5 3.0 GHz processor and 8.0 GB RAM. In Fig. 5.13, ROC curves of the eye corner detection accuracy using the proposed fusion method are plotted with images captured at 5 meters for multiple gazes



(a)

(b)

Figure 5.12: Eye localisation method. (a) Acquiring eye frames from the user and (b) detection of a valid frame which consist of the iris and the sclera.

(F-G, R-G, and L-G). The horizontal axis represents the error threshold where the Euclidean distances are calculated for the detected eye corners with the ground-truth eye corners which manually annotated. It is evident that the fusion method performs best in the front eye gaze direction whereas it achieves lower corner detection accuracy for a deviated eye gaze.

To observe the capturing distance effect on the proposed eye corner detection method, ROC curves and boxplots for F-G images captured at distances ranging from 4-8 meters are presented in Fig. 5.14. For the boxplot, the red cross represents outliers, the top and bottom whiskers are the extreme values, the horizontal solid red line represents the median value, while the bottom and the top of boxes are the first and the third quartile. The highest detection rate is observed when an eye

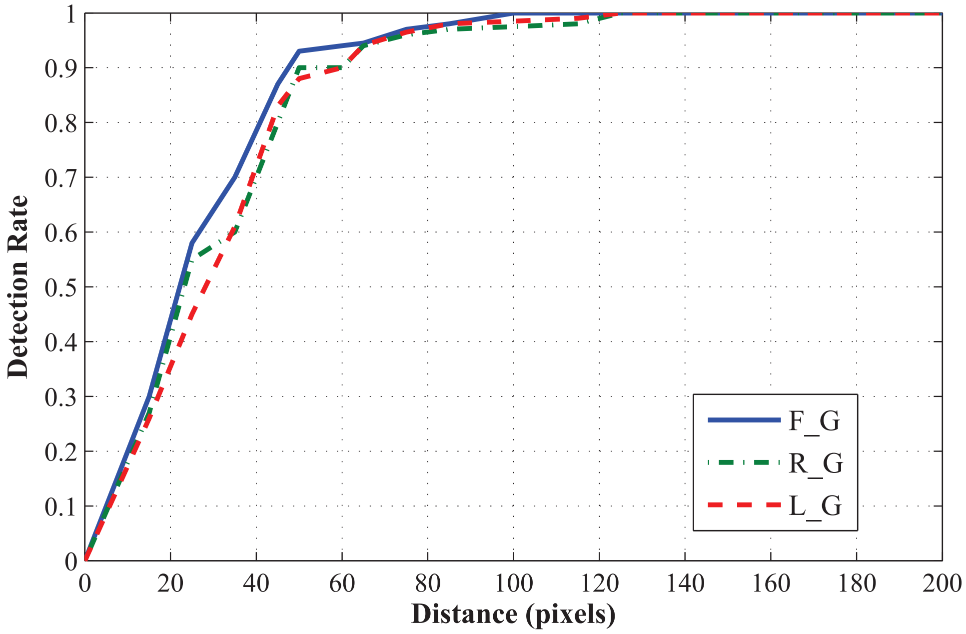
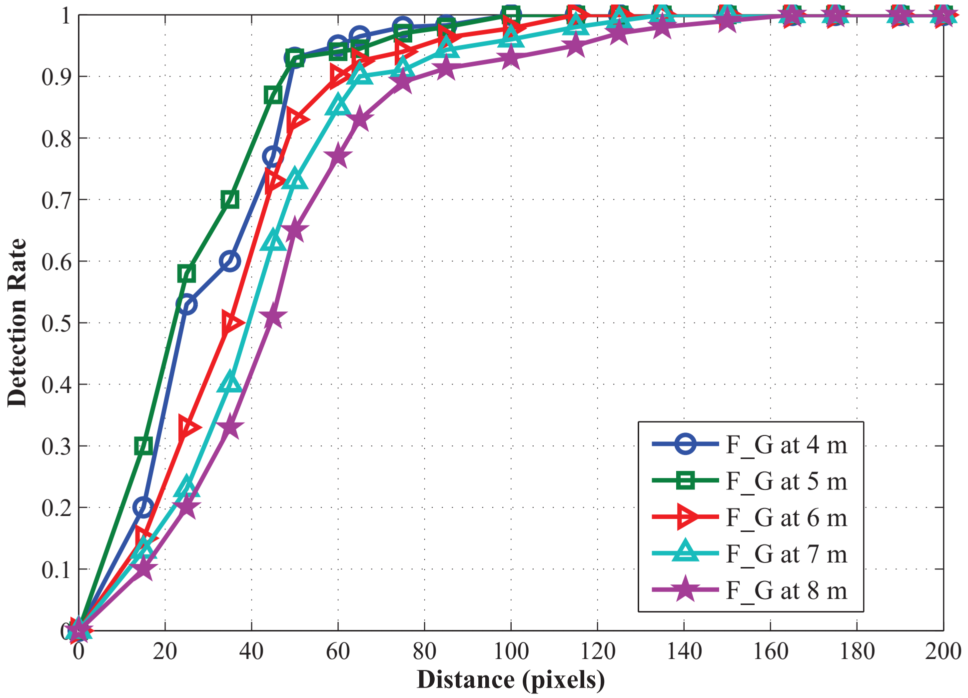


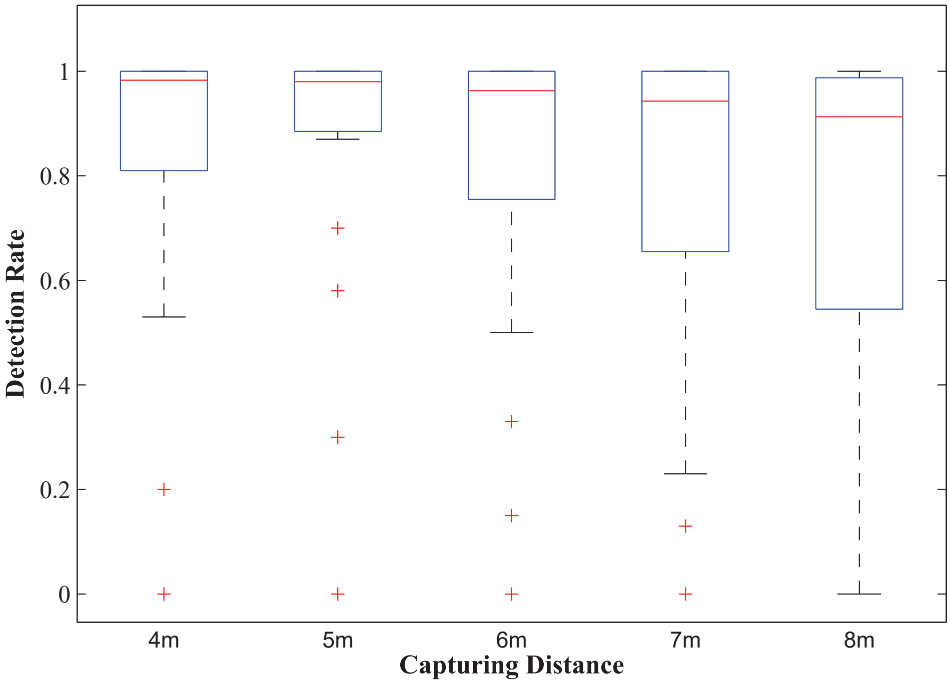
Figure 5.13: The UBIRIS.v2 ROC curves for the eye corner detection accuracy with multiple gaze directions (F-G, R-G, and L-G) using images captured at 5 meters.

image is captured at 5 meters as shown in Fig. 5.14a. This is due to some images in the UBIRIS.v2 database captured at 4 meters distance being cropped removing some parts of the eye corners. However, the detection rate is inversely proportional to the capturing distance. Nevertheless, the median values (horizontal solid lines) and the third quartiles values observed in Fig. 5.14b, which compares boxplots of the eye corner detection rate, show that the proposed method achieves high performance with images captured at-a-distance ranging from 4-8 meters. However, the detection rate in the first quartiles values degrades as the distance increases and hence, the quality of images degrades severely.

In terms of comparing the eye corner detection method with state-of-the-art methods on the UBIRIS.v2 database, only one work proposed by Santos and Proenca [160] is found. However, their method did not involve detecting eye gaze direction. The eye corner detection rate for F-G images is compared using 300 images captured at 5 meters. As shown in Fig. 5.15, the proposed detection method is better than the Santos and Proenca method when the error threshold is less than 40 pixels whereas both methods achieve higher rates after 50 pixels distance threshold.



(a)



(b)

Figure 5.14: Comparison of eye corner detection rates using the proposed fusion method for F-G images captured at distances ranging from 4 to 8 meters. (a) ROC curves and (b) boxplot of the detection rate.

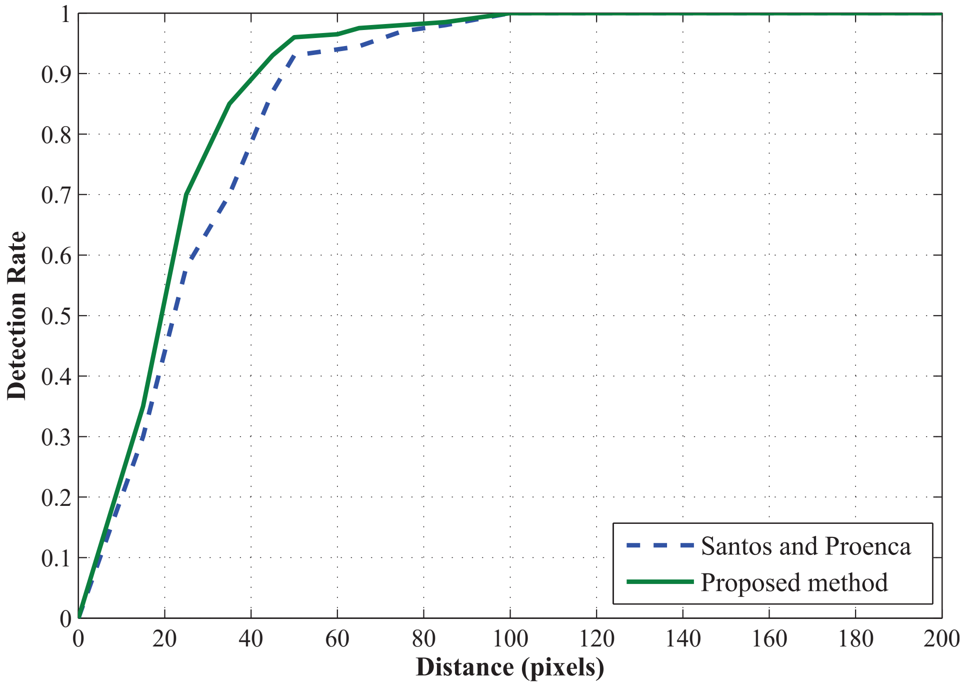


Figure 5.15: The UBIRIS.v2 ROC curve comparison of proposed eye corner detection with state-of-the-art method due to Santos and Proenca.

5.5.4 *Eye Gaze Estimation Evaluation*

Finally, the proposed eye gaze detection methods involving eyelid border detection, eyelid skin detection and fusion detection methods are assessed as shown in Table 5.2. As mentioned above, due to the incorrect image cropping for 4 meters captured images which led to removal of eye corners in the UBIRIS.v2 database, higher gaze detection rate (95%) was achieved with images captured at 5 meters using the proposed gaze detection method. It is also evident that the larger the image capturing distance, the lower eye corner and gaze detection rates become. In addition, it is noteworthy that the gaze detection rate is higher than eye corner detection rate as it is challenging to detect the exact location of both eye corner points in a noisy imaging environment.

5.5.5 *Fm Measure for the Employed Datasets*

For the sclera image quality measure, the fm measure is used to classify each image according to the sharpness and smoothness of

Table 5.2: Eye gaze detection rates for the UBIRIS.v2 database images captured at distances range from 4 to 8 meters

Gaze Direction	**Image Capturing Distance**				
	4 m	**5 m**	**6 m**	**7 m**	**8 m**
F-G	89.6 %	**91.4 %**	85.8 %	81.5 %	76.4 %
R-G	87.4 %	88.7 %	83.2 %	79.3 %	74.6 %
L-G	86.9 %	89.3 %	82.7 %	78.7 %	72.3 %

(a) Eyelid border method

Gaze Direction	**Image Capturing Distance**				
	4 m	**5 m**	**6 m**	**7 m**	**8 m**
F-G	88.3 %	**90.5 %**	83.6 %	79.9 %	72.6 %
R-G	86.6 %	89.2 %	80.1 %	76.6 %	71.6 %
L-G	87.4 %	88.1 %	81.2 %	77.1 %	70.2 %

(b) Eyelid skin method

Gaze Direction	**Image Capturing Distance**				
	4 m	**5 m**	**6 m**	**7 m**	**8 m**
F-G	93.3 %	**95.2 %**	92.7 %	89.3 %	83.1 %
R-G	90.1 %	94.2 %	92.5 %	85.7 %	81.7 %
L-G	91.0 %	91.9 %	93.1 %	84.4 %	80.4 %

(c) Fusion method

the edge components within the image. Most of the recent work on sclera recognition which achieved high recognition rates has used the UBIRIS.v1 database. Therefore, it is selected as an Excellent reference

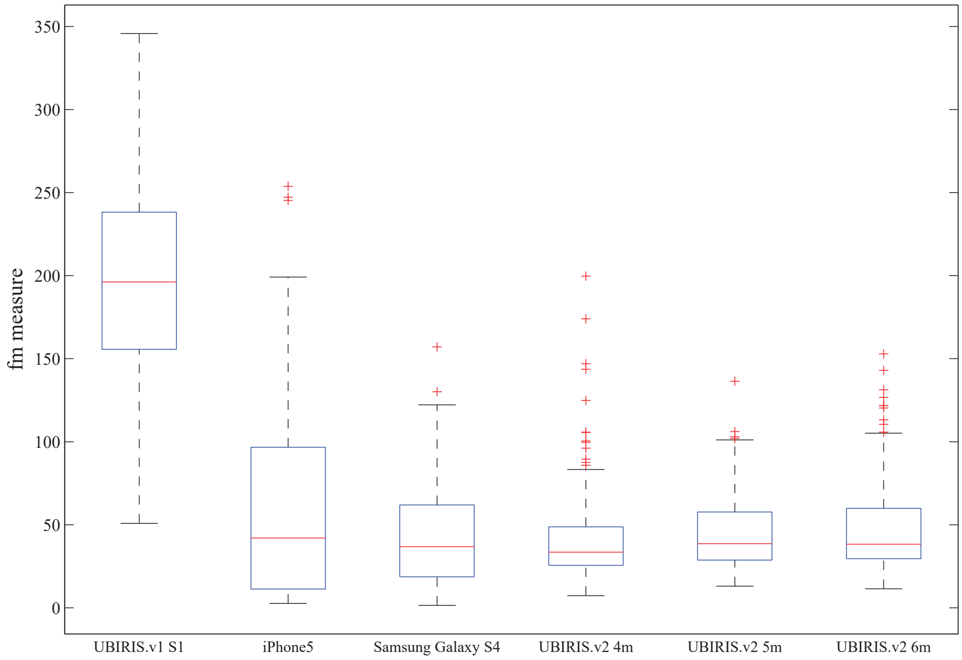


Figure 5.16: Boxplots of fm measure for the UBIRIS.v1.S1, UBIRIS.v2 at 4-5-6 meters, and MICHE iPhone 5 and Samsung Galaxy S4 databases.

to other exploited databases in this chapter. Boxplots of the fm measure for the exploited databases are shown in Fig 5.16. The horizontal solid red line represents the median value, the bottom and the top of boxes are the first and the third quartile of fm values, the top and bottom whiskers are the extreme fm values, while the red cross represents outliers.

It is observed that the median value for the constrained UBIRIS.v1 database is higher than the top extreme values and some top outlier values for the unconstrained imaging databases. This confirms how challenging it is to employ relaxed captured images in sclera recognition. However, trying to overcome these challenges will pave the way for practical sclera recognition system for a real-time application. In addition, Fig. 5.17 shows the histogram distributions of the utilised datasets. For the constrained images, the majority of fm values are greater than 200 whereas for unconstrained images, most of the fm values fall below 100.

Table 5.3: Sclera segmentation accuracy of the proposed sclera segmentation with respect to GBU classification

Database	**G**			**B**			**U**		
	Correct	**Fair**	**Poor**	**Correct**	**Fair**	**Poor**	**Correct**	**Fair**	**Poor**
MICHE iPhone 5	**43%**	35%	22%	29%	**40%**	31%	18%	21%	**51%**
MICHE Samsung Galaxy S4	**40%**	36%	24%	25%	**39%**	36%	15%	28%	**57%**
UBIRIS.v2 at 4m	**66%**	22%	17%	**42%**	32%	26%	32%	**40%**	28%
UBIRIS.v2 at 5m	**53%**	18%	29%	37%	**45%**	18%	25%	35%	**40%**
UBIRIS.v2 at 6m	33%	**43%**	24%	30%	**41%**	29%	17%	37%	**46%**

To classify each image according to the EGBU categories and for relaxed imaging capturing, the following classifier is proposed as

$$Image\ condition = \begin{cases} E & \text{if } fm > 150 \\ G & \text{if } 100 \leqslant fm \leqslant 150 \\ B, & \text{if } 50 \leqslant fm < 100 \\ U, & \text{otherwise} \end{cases}, \tag{5.11}$$

where for each image condition, fm thresholds have been set empirically.

5.5.6 *Sclera Segmentation Under EGBU Hypothesis*

For the proposed sclera segmentation, first the segmentation accuracy is compared for each image according to their condition in terms of G, B, and U. The E category which mostly apply on the UBIRIS.v1 database is discarded as the sclera segmentation on this database is already evaluated in Chapters 3 and 4. To create a ground truth, 250 images for each database are manually segmented and labelled resulting in a total of 1000 images divided approximately equally among GBU image conditions. For the authentication protocol for sclera segmentation, a pixel binary classification method is used to

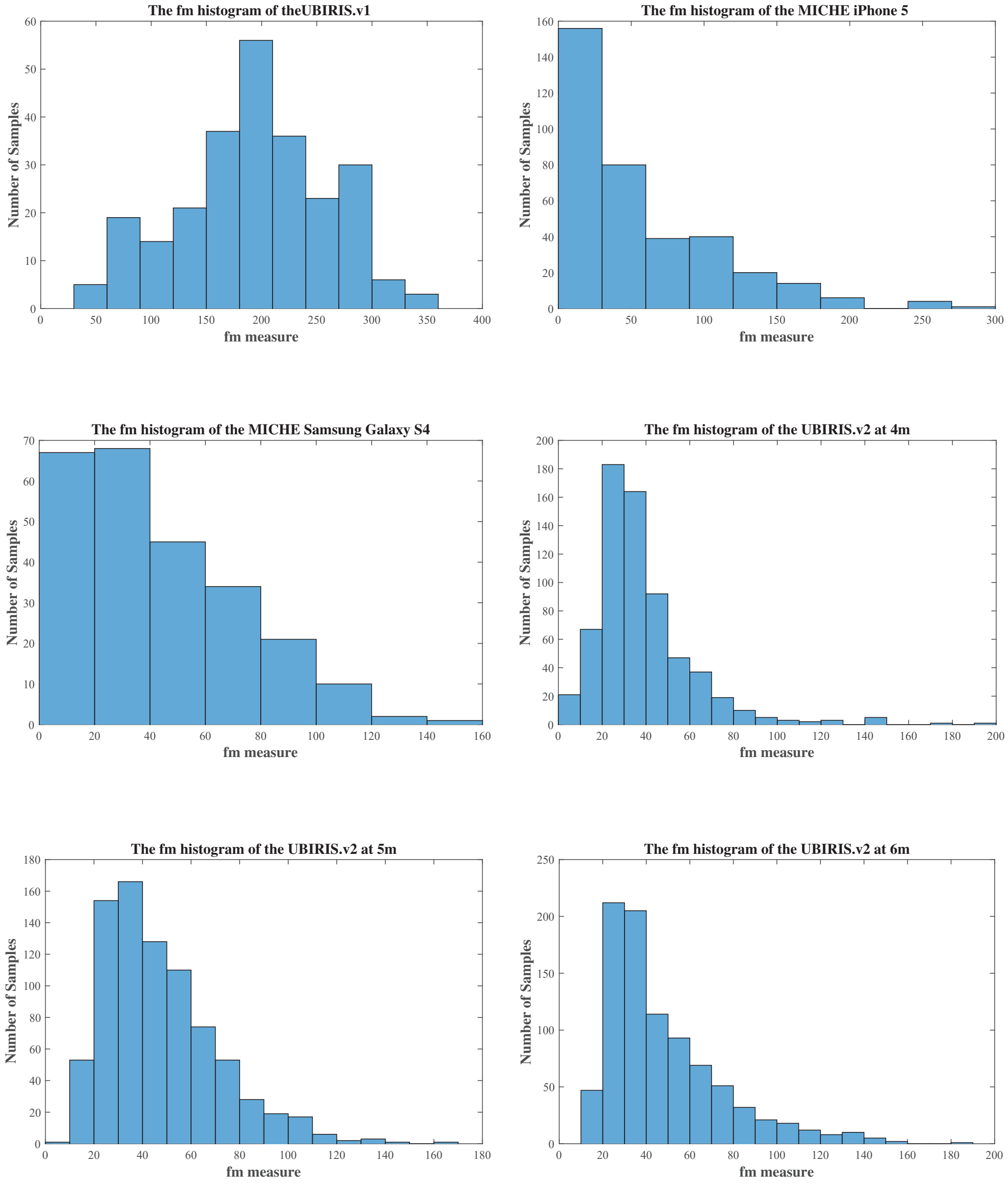


Figure 5.17: Histogram distributions of fm values for constrained and unconstrained eye image datasets.

calculate the total error pixels and segmentation accuracy using 4.37 and 4.38.

As shown in Table 5.3, the sclera Ep offset is classified for each GBU category into correct where $Ep < 10$, fair where $10 < Ep < 30$, and poor where $Ep > 30$ pixels. It is expected that the G images have the highest correct segmentation rate where the UBIRIS.v2 at 4m has 66% and the rate decreased as the capturing distance increased.

Whereas for the mobile images of the iPhone 5 and the Galaxy S4 achieved less segmentation rate where these images share most of the UBIRIS.v2 noise factors, but the lack of precise localisation and fixed capturing distance introduced more challenge to sclera segmentation. While for the B and U classification, it is typical that the fair and poor segmentation are higher except for B images of the UBIRIS.v2 captured 4 meters where correct segmentation has the highest rate (42%).

In addition, and for the unconstrained mobile datasets in terms of a number of skin classifiers th_s in (4.5) which are required to achieve overall highest segmentation accuracy. As shown in Fig. 5.18a, for the UBIRIS.v2 database at 4-5-6 meters, using at least any correct segmented pixel from the descriptor will achieve the highest segmentation accuracy while for the MICHE database for both devices, it requires all skin pixel descriptors to achieve the maximum segmentation accuracy. This confirms that using a mobile device for the sclera recognition has challenging issues which need to be considered in order not to affect the system performance. Some examples of the sclera segmentation with $th_s = 1$ and $th_s = 6$ for the skin descriptors are shown in Fig. 5.18b.

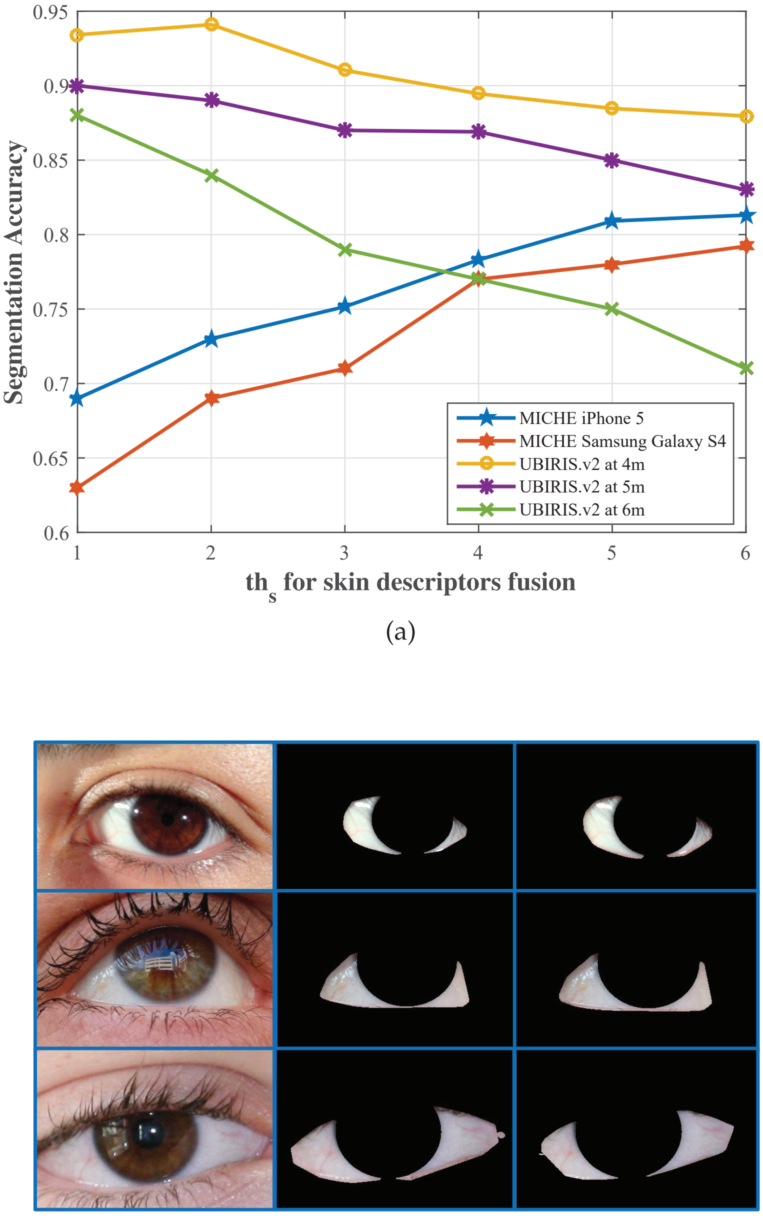
5.5.7 *Sclera Template Rotation Alignment and Distance Scaling Evaluation*

For the final part of the proposed system, the STRADS method is evaluated in terms of calculating the EER, plotting ROC curves, and evaluating the processing times for the entire proposed system under various

Table 5.4: The EER of the proposed sclera recognition with and without STRADS using the UBIRIS.v2 and MICHE databases

Method	**With STRADS**	**Without STRADS**
Database	**EER (%)**	
UBIRIS.v2 [112]	7.32	32.31
MICHE iPhone 5 [114]	7.69	36.54
MICHE Samsung Galaxy S4 [114]	9.06	34.07

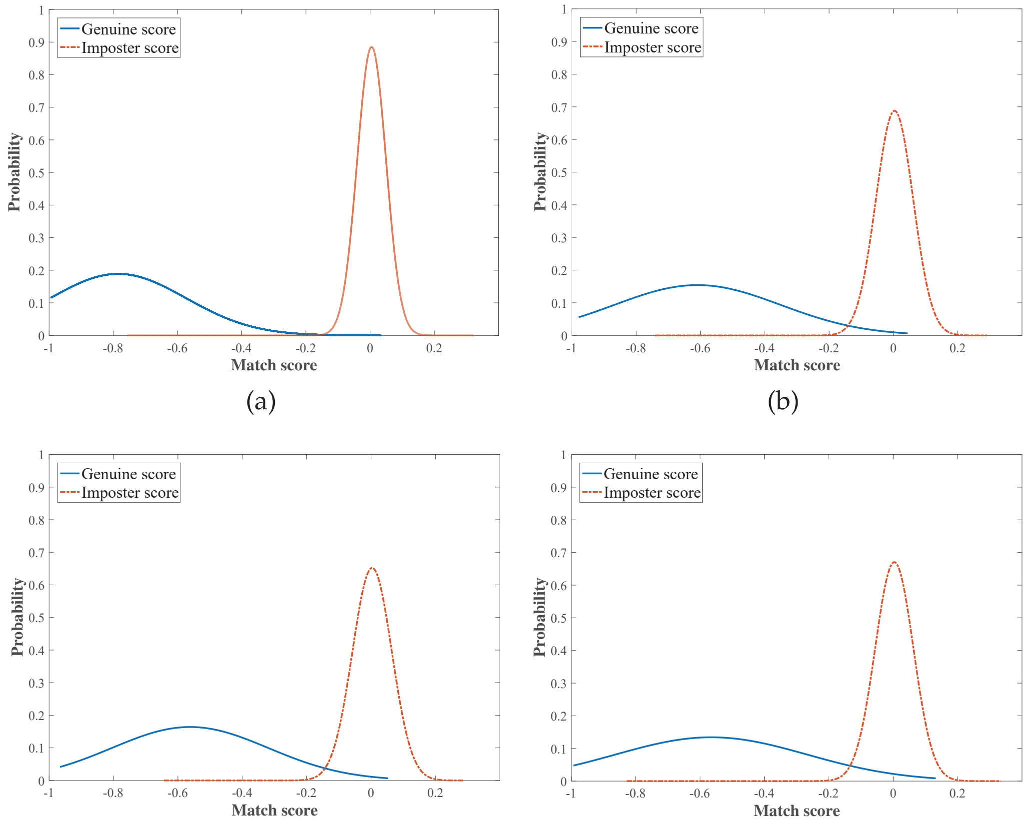
image qualities. For image enhancement, feature extraction and matching process, the complex-sclera method along with KLDA and cosine



(a)

(b)

Figure 5.18: Sclera segmentation accuracy evaluation for different skin descriptor threshold th_s. Examples of using different skin descriptors for sclera segmentation. The left column represents from the top to bottom images from the MICHE iPhone 5, MICHE Samsung Galaxy S4 and UBIRIS.v2 at 4 meters databases. The middle column is the resulting sclera map when $\mathrm{th}_s = 1$ while the right column is when $\mathrm{th}_s = 6$.



(a) (b) (c) (d)

Figure 5.19: Normalised genuine and impostor scores for proposed STRADS using (a) the UBIRIS.v1, (b) the UBIRIS.v2, (c) MICHE iPhone 5, and (d) Samsung Galaxy S4 images.

similarity are utilised. For the UBIRIS.v2, the evaluation scenario has been achieved in multi-session contexts with F-G, R-G and L-G eye gazes. For each context, 2 images at a distance of 4 meters for training and 4 images at a distance of 5 and 6 meters for testing per user and gaze direction are exploited. While for the MICHE, 4 images for training and 4 for testing for both the iPhone 5 and Samsung Galaxy S4 are used. K-fold cross-validation with the number of trials equal to five has been utilised in the evaluation.

As shown in Table 5.4, although the resulting EER after using the STRADS is not necessarily low, however, the average EER for the UBIRIS.v2 for 4,5,6 meters and the MICHE database have been improved significantly when applying the STRADS method. The improvement rate is more than 25% for all exploited databases. This confirms the need for a rotation alignment and distance scaling method when dealing with the sclera under relaxed imaging conditions.

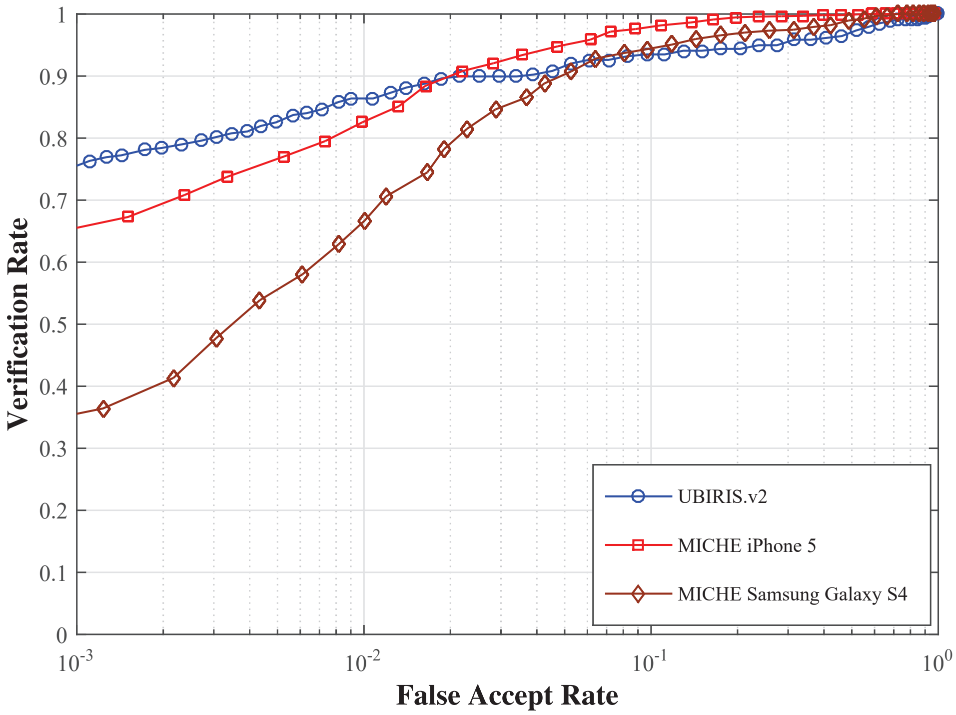


Figure 5.20: ROC curves of the proposed STRADS method for the UBIRIS.v2, MICHE iPhone 5, and MICHE Samsung Galaxy S4 datasets.

The normalised distribution of genuine and imposter scores using the UBIRIS.v1, UBIRIS.v2, and both MICHE iPhone 5 and Samsung Galaxy S4 images are plotted in Fig. 5.19. For the E images within the UBIRIS.v1 database, the distribution of genuine and impostor scores shows significant separation and thus low error rates as most of the noise factors have been minimised. While for the GBU images, it is notable that as the UBIRIS.v2 images are less constrained in terms of the capturing process which has not achieved by the participated users themselves, therefore the matching score is better separated than the MICHE mobile devices images where users are responsible for capturing their images. In addition, from the ROC curves for the recognition rate as shown in Fig. 5.20, it is evident that sclera recognition using relaxed imaging conditions and without user co-operation is a major challenge which needs more investigation.

Finally, for the processing time evaluation, as shown in Table 5.5, the proposed eye image quality measure and STRADS methods achieved significantly lower processing time (0.24s and 0.16s respectively). While the sclera segmentation process and eye gaze estimation consume higher times as the operation is pixel-wise and thus the greater the image resolution, the higher the processing time and the better sclera segmentation accuracy.

Table 5.5: The process times of the proposed sclera system under relaxed imaging conditions

Sclera Segmentation Steps	**Run Times**
Eye image fm measure	0.24 s
Sclera Segmentation	1.58 s
Eye gaze estimation	0.93 s
STRADS	0.16 s

5.5.8 *Empirical Parameters Tuning Analysis*

The eye gaze direction detection mainly depends on th_r and th_l. A higher value of th_r and th_l means the right and left eye gaze directions can only be detected if the iris position is extremely next to eye corners.

5.6 CHAPTER SUMMARY

In this chapter, challenges and issues in sclera recognition under various imaging conditions and for real-time applications were considered. The chapter has identified three major issues related to eye localisation, eye corner detection, eye gaze estimations, and eye image quality measure. Therefore, an efficient eye localisation method is proposed and tested in a real-time scenario using live video feed captured with a webcam. In addition, novel eye corner and gaze detection methods have been proposed to overcome limitations in using relaxed imaging constraints in sclera recognition. Eye corner detection has been achieved through an adaptive fusion method depending on the eyelid borders and eyelid skin properties. Furthermore, an efficient method for eye gaze detection has been suggested to overcome the change in rotation and distance of sclera areas when an eye image is captured. The eye image quality measure is achieved by estimating the focus through high and low frequencies within the image using a practical method for real time application.

The evaluation process is achieved on the In-house database, UBIRIS.v1, UBIRIS.v2 and MICHE iPhone 5 and Samsung Galaxy S4 databases and the results are promising in terms of how using unconstrained images could affect the sclera recognition and ways to overcome challenges introduced. The proposed eye corner and gaze detection methods have achieved high detection rates for different distances and gazes thus introducing more robustness. In addition, this chapter extends the knowledge of challenges facing the sclera recognition design in real time applications and will serve as a base for future studies in the field. The next chapter will draw the conclusion of the entire work on designing an efficient unsupervised sclera recondition system under various imaging conditions and propose future works which will push the development of sclera recognition system design even further.

6

CONCLUSIONS AND FUTURE WORK

6.1 SUMMARY AND CONCLUSION

To conclude, this thesis has provided a significant development in sclera recognition technology. In particular, a comprehensive survey on the improvement of sclera recognition has been presented in Chapter 2. The recent sclera recognition methods have been listed and analysed in terms of the properties of utilised database, segmentation, vessel enhancement, feature extraction, template registration, and matching techniques along with their best-reported results. In addition, an overview of recent publicly available datasets of eye images in the visible spectrum has been listed with their merits and disadvantages.

On the other hand, Chapter 3 proposed a new sclera biometric method for indoor applications where the noise factors are minimised. First of all, efficient sclera segmentation was designed which is based on the sclera shape using active contours without edges method. Initialisation of the contour seeds was achieved through an enhanced geometric method. Then, a new method to enhance and extract the binary network of the vascular pattern in the sclera was proposed based on the IUWT. After that, template registration based on selecting discriminant features from the training samples was proposed which only stores permanent blood vessel features over different samples and updates the final user template. The proposed algorithm was tested on the UBIRIS.v1 and UTIRIS dataset which have minimised noise factors simulating indoor constrained image capturing. Each part of the algorithm was first evaluated individually and the entire system was then tested. In terms of sclera segmentation, the proposed method achieved high segmentation rates (for instance 95.38% for the UBIRIS.v1 session 1 images). In addition, the active contours without edges method was compared with the state-of-the-art active contour methods and the proposed method performed better both in terms of segmentation accuracy and processing times. Then, the sclera recognition algorithm was evaluated with the state-of-the-art methods for sclera recognition and the reported EER was 3.68% for the UBIRIS.v1 multi-session context and 6.67% for the UTIRIS database. Furthermore, in order to provide a comparison between sclera recognition and iris recognition under visible wavelength spectrum, the proposed sclera recognition system was compared with the iris recognition system using the UBIRIS.v1 database and the results were promising. Finally, the processing times for the proposed recognition system were reported.

In Chapter 4, as most recent methods on the sclera biometric have been designed under constrained imaging conditions, a new sclera recognition method under unconstrained imaging conditions was proposed. Firstly, an efficient sclera segmentation method was designed based on the properties of the sclera region and the skin around the eye. This involved fusing multiple sclera and skin classifiers to achieve high segmentation accuracy despite the noise included in the exploited eye images. Then, post-processing operations were applied to refine the final sclera mask. Next, a novel method was proposed to overcome challenges related to eye rotation due to head pose movement and different sclera image capturing distance. The proposed method first measured the rotation degree of sclera regions (left, right, or both sclera sides depending on the gaze direction). Then, each sclera side is rotated according to the degree of rotation which is calculated efficiently using the novel approach. After that, sclera templates captured at different distances were scaled to improve the recognition rate. In terms of vessel enhancement and feature extraction methods, a holistic approach was suggested because of the difficulty of extracting the binary vascular network under relaxed imaging conditions. As a result, a novel method to extract complex features of the sclera blood vessels based on the DT-CWT method was proposed. For feature classification, the KLDA was used for dimensionality reduction along with the nearest neighbour method with the Mahalanobis Cosine similarity measure. To evaluate the proposed method, the UBIRIS.v2 database which has challenging imaging conditions such as capturing at-a-distance and on-the-move is utilised in addition to the UBIRIS.v1 and UTIRIS databases. The proposed segmentation method was evaluated first by creating ground truth samples for the above-mentioned datasets and a pixel-based comparison was performed. The segmentation accuracy results were significantly high for both constrained and unconstrained images (best-reported segmentation accuracy results for the UBIRIS.v1 session 1, UBIRIS.v2, and UTIRIS databases were all above 90%). Next, the STRADS method was tested on the UBIRIS.v2 database with only images with capturing distance ranges from 4 to 6 meters. Although EER results were high, the proposed STRADS method has however improved these rates dramatically showing evidence of the effectiveness of the proposed rotation alignment and distance scaling method. Then, the proposed sclera complex features were tested with features extracted using state-of-the-art feature extraction methods and the results were better in terms of EER and processing times (0.79% of EER and 0.121 sec of processing time for the UBIRIS.v1 session 1 images). Finally, the system was assessed and compared to recent sclera recognition methods and the achieved EER result was 0.47%.

Furthermore, several issues and challenges regarding a sclera biometric system particularly when applied in real-time applications have been discussed and resolutions have been introduced in Chapter 5. These issues include eye localisation, eye corner and gaze detection, image quality, and non-cooperative relaxed imaging capturing using mobile devices. First of all, eye localisation was designed and examined using a live-feed video obtained via an HD webcam where the algorithm monitors the live video frame-by-frame until an eye region is spotted and hence, bypassing the sclera biometric failure-to-enrol processing. Next, an eye gaze detection method was proposed based on novel eye corner points detection. A fusion method was proposed to enhance eye corner points detection and improve eye gaze estimation thus developing an automated unsupervised sclera recognition system. In addition, a sclera image quality measure was used based on the focus measure. The sclera image was checked first by examining high and low frequencies to assess its quality. Then based on the focus measure degree, four categories defining sclera image quality were suggested based on the EGBU classification. Moreover, sclera recognition system design was evaluated under each quality category. Furthermore, four different databases have been employed to evaluate the aforementioned algorithm which are the UBIRIS.v1, UBIRIS.v2, MICHE, and in-house webcam database simulating the constrained, relaxed, non-cooperative, and real-time scenarios. As a consequence, the proposed eye localisation method has performed accurately as it does not pass any video frame to the next stage until an eye is detected. The eye corner detection algorithm achieved high-performance rates with images captured at a distance (best mean corner detection accuracy was achieved with images captured at 5 meters which equals 97%). While for the eye gaze estimation under relaxed imaging condition, the proposed method achieved high gaze estimation rates with all capturing distances of the UBIRIS.v2 database (the highest gaze estimation was 95.2% with images captured at 5 meters and with frontal-gaze direction). For the EGBU classification, images from the four datasets were evaluated in terms of segmentation accuracy, STRADS method accuracy and recognition rate providing a full review of how the quality of sclera images affects the sclera biometric and the challenges which emerge when using mobile devices.

6.2 FUTURE WORK

The aspirations of this thesis involve presenting new research horizons in the future of sclera recognition. However, there are some limitations which need to be considered as future work in order to improve the eye biometric which includes the sclera biometric implicitly. Key issues are

i. As all the utilised public datasets in recent sclera recognition systems evaluation are intended to assess the iris biometric under visible wavelength illumination, it is therefore necessary to collect a database specifically for evaluating the sclera biometric using various imaging scenarios and devices. This may include exploiting modern mobile devices and experimenting with several capturing scenes on a wider population to examine time-invariance, uniqueness, and universality of sclera recognition.

ii. Sclera biometric depends largely on locating the iris position correctly which improves the segmentation and eye gaze detection methods. As Daugman's iris segmentation method has only been adopted in this thesis, it is essential to develop a new method for iris positioning under visible wavelength illumination fit for purpose. On the other hand, designing sclera segmentation and eye gaze detection methods without the need to locate the iris is the ultimate goal.

iii. Although two methods in this thesis have been designed for sclera segmentation which were shown to perform efficiently, the processing time however requires more improvement especially as the resolution of modern cameras is increasing with time.

iv. Anti-spoofing methods are required to be considered for the sclera biometric as this is an emerging technology and more research in this area is important to avoid any weaknesses which could be exploited by imposters.

v. Designing efficacious fusion method for multi-angle sclera regions where sclera regions change with respect to gaze direction.

vi. Fusing multiple modalities with the sclera such as the iris and periocular regions using different fusion techniques (for instance score fusion, feature fusion, or matching fusion) is suggested to improve the recognition rate.

vii. The glossy surface of the sclera which could reflect natural lights and eyelashes which could block sclera regions are considered as the next step for sclera biometric under relaxed imaging conditions.

viii. Employ recent mobile devices in the study of sclera biometric as the front-facing camera of these modern devices such as the iPhone X and Samsung Galaxy S8 has improved dramatically in terms of resolution and colour depths.

ix. Pathologies and medical conditions affecting the eye, and in particular the sclera, such as conjunctivitis, eye allergy, dry eye syndrome, blepharitis, Melanosis, Scleral Coloboma, Ectasia, and subconjunctival haemorrhage are required to be studied and their effect on the sclera biometric performance.

Part II

Thesis References

REFERENCES

[1] Z. Boren, "There are officially more mobile devices than people in the world," *The Independent*, 2014. [Online]. Available: http://www.independent.co.uk/life-style/gadgets-and-tech/news/there-are-officially-more-mobile-devices-than-people-in-the/-world-9780518.html [Accessed: 24/07/2017]

[2] "Number of mobile phone users worldwide 2013-2019 | statista," *Statista*, 2017. [Online]. Available: https://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/ [Accessed: 24/07/2017]

[3] S. K. Vashist, O. Mudanyali, E. M. Schneider, R. Zengerle, and A. Ozcan, "Cellphone-based devices for bioanalytical sciences," *Analytical and bioanalytical chemistry*, vol. 406, no. 14, pp. 3263–3277, 2014.

[4] "Number of mobile phone users in India from 2013 to 2019 (in millions)," *Statista*, 2016. [Online]. Available: https://www.statista.com/statistics/274658/forecast-of-mobile-phone-users-in-india/ [Accessed: 27/07/2017]

[5] D. Casciani, "cybercrime and fraud scale revealed in annual figures," *BBC News*, 2017. [Online]. Available: http://www.bbc.co.uk/news/uk-38675683 [Accessed: 21/07/2017]

[6] B. Laberis, "20 eye-opening cybercrime statistics," *Security Intelligence*, 2016. [Online]. Available: https://securityintelligence.com/20-eye-opening-cybercrime-statistics/ [Accessed: 24/07/2017]

[7] "Trust in storing biometric data for payments 2016 | Great Britain," *Statista*, 2016. [Online]. Available: https://www.statista.com/statistics/679346/

trust-in-storing-biometric-data-for-payments-by-company-or/-institution-great-britain/ [Accessed: 25/07/2017]

[8] D. Maltoni, D. Maio, A. Jain, and S. Prabhakar, *Handbook of fingerprint recognition*. Springer Science & Business Media, 2009.

[9] C. Rathgeb, A. Uhl, and P. Wild, *Iris biometrics: from segmentation to template security*. Springer Science & Business Media, 2012, vol. 59.

[10] K. W. Bowyer and M. J. Burge, *Handbook of iris recognition*. Springer, 2016.

[11] J. G. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 11, pp. 1148–1161, 1993.

[12] Y. E. Du, *Biometrics: from fiction to practice*. CRC Press, 2013.

[13] H. Proenca, "On the feasibility of the visible wavelength, at-a-distance and on-the-move iris recognition," in *IEEE Workshop on Computational Intell. in Biometrics: Theory, Algorithms, and Applicat.*, 2009, pp. 9–15.

[14] H. Proenca and L. A. Alexandre, "The NICE.I: Noisy iris challenge evaluation - part i," in *First IEEE Int. Conf. on Biometrics: Theory, Applicat., and Syst.*, 2007, pp. 1–4.

[15] K. W. Bowyer, "The results of the NICE.II iris biometrics competition," *Pattern Recognition Lett.*, vol. 33, no. 8, pp. 965–969, 2012.

[16] K. B. Raja, R. Raghavendra, V. K. Vemuri, and C. Busch, "Smartphone based visible iris recognition using deep sparse filtering," *Pattern Recognition Lett.*, vol. 57, pp. 33–42, 2015.

[17] R. Derakhshani, A. Ross, and S. Crihalmeanu, "A new biometric modality based on conjunctival vasculature," *Proc. of Artificial Neural Networks in Eng. (ANNIE)*, pp. 1–8, 2006.

[18] U. Park, A. Ross, and A. K. Jain, "Periocular biometrics in the visible spectrum: A feasibility study," in *IEEE 3rd Int. Conf. on Biometrics: Theory, Applicat., and Syst. (BTAS)*, 2009, pp. 1–6.

[19] S. Alkassar, W. L. Woo, S. S. Dlay, and J. A. Chambers, "Robust sclera recognition system with novel sclera segmentation and validation techniques," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 3, pp. 474–486, 2017.

[20] S. Alkassar, W. L. Woo, S. S. Dlay, and J. A. Chambers, "A novel method for sclera recognition with images captured on-the-move and at-a-distance," in *4th Int. Workshop on Biometrics and Forensics (IWBF)*, 2016, pp. 1–6.

[21] S. Alkassar, W. L. Woo, S. S. Dlay, and J. A. Chambers, "Enhanced segmentation and complex-sclera features for human recognition with unconstrained visible-wavelength imaging," in *Int. Conf. on Biometrics (ICB)*, 2016, pp. 1–8.

[22] S. Alkassar, W. L. Woo, S. S. Dlay, and J. A. Chambers, "Efficient eye corner and gaze detection for sclera recognition under relaxed imaging constraints," in *24th European Signal Processing Conf. (EUSIPCO)*, 2016, pp. 1–5.

[23] S. Alkassar, W. L. Woo, S. Dlay, and J. A. Chambers, "Sclera recognition: on the quality measure and segmentation of degraded images captured under relaxed imaging conditions," *IET Biometrics*, vol. 6, no. 4, pp. 266–275, 2017.

[24] T. Dunstone and N. Yager, *Biometric system and data analysis: Design, evaluation, and data mining*. Springer Science & Business Media, 2008.

[25] A. Jain, L. Hong, and S. Pankanti, "Biometric identification," *Commun. of the ACM*, vol. 43, no. 2, pp. 90–98, 2000.

[26] A. Jain, P. Flynn, and A. A. Ross, *Handbook of biometrics*. Springer Science & Business Media, 2007.

[27] Y. B. Reeba and R. Shanmugalakshmi, "Spoofing face recognition," in *Int. Conf. on Advanced Computing and Commun. Syst.*, 2015, pp. 1–5.

[28] R. B. Hill, "Retina identification," *Biometrics: Personal Identification in Networked Soc.*, pp. 123–141, 2002.

[29] C.-L. Lin and K.-C. Fan, “Biometric verification using thermal images of palm-dorsa vein patterns,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 2, pp. 199–213, 2004.

[30] N. Miura, A. Nagasaka, and T. Miyatake, “Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification,” *Mach. Vision and Applicat.*, vol. 15, no. 4, pp. 194–203, 2004.

[31] Z. Zhou, E. Y. Du, N. L. Thomas, and E. J. Delp, “A new human identification method: sclera recognition,” *IEEE Trans. Syst., Man, Cybern. A*, vol. 42, no. 3, pp. 571–583, 2012.

[32] J. Daugman, “How iris recognition works,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 1, pp. 21–30, 2004.

[33] Y. Eliza Du, “Review of iris recognition: cameras, systems, and their applications,” *Sensor review*, vol. 26, no. 1, pp. 66–69, 2006.

[34] M. A. Turk and A. P. Pentland, “Face recognition using eigenfaces,” in *Proc. IEEE Comput. Soc. Conf. on Comput. Vision and Pattern Recognition*, 1991, pp. 586–591.

[35] M. J. Moscou and A. J. Bogdanove, “A simple cipher governs DNA recognition by TAL effectors,” *Science*, vol. 326, no. 5959, pp. 1501–1501, 2009.

[36] L. Zhang, L. Zhang, D. Zhang, and H. Zhu, “Online finger-knuckle-print verification for personal authentication,” *Pattern recognition*, vol. 43, no. 7, pp. 2560–2571, 2010.

[37] R. R. O. Al-Nima, S. S. Dlay, S. A. Al-Sumaidaee, W. L. Woo, and J. A. Chambers, “Robust feature extraction and salvage schemes for finger texture based biometrics,” *IET Biometrics*, vol. 6, no. 2, pp. 43–52, 2016.

[38] A. Kumar and K. V. Prathyusha, “Personal authentication using hand vein triangulation and knuckle shape,” *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 2127–2136, 2009.

[39] D. Hosseinzadeh and S. Krishnan, “Gaussian mixture modeling of keystroke patterns for biometric applications,” *IEEE Trans. Syst., Man, Cybern. C*, vol. 38, no. 6, pp. 816–826, 2008.

[40] D. Srinivasan, W. Ng, and A. Liew, "Neural-network-based signature recognition for harmonic source identification," *IEEE Trans. Power Del.*, vol. 21, no. 1, pp. 398–405, 2006.

[41] J. Man and B. Bhanu, "Individual recognition using gait energy image," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 2, pp. 316–322, 2006.

[42] R. M. Bolle, J. Connell, S. Pankanti, N. K. Ratha, and A. W. Senior, *Guide to biometrics*. Springer Science & Business Media, 2013.

[43] S. Z. Li and A. Jain, *Encyclopedia of biometrics*. Springer Publishing Company, Incorporated, 2015.

[44] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 1, pp. 4–20, 2004.

[45] S. Prabhakar, S. Pankanti, and A. K. Jain, "Biometric recognition: Security and privacy concerns," *IEEE Security & Privacy*, vol. 99, no. 2, pp. 33–42, 2003.

[46] J. L. Wayman, "Fundamentals of biometric authentication technologies," *Int. J. of Image and Graphics*, vol. 1, no. 01, pp. 93–113, 2001.

[47] T. Dunstone, *Biometric system and data analysis: Design, evaluation, and data mining*. Springer, 2008.

[48] A. Martin, G. Doddington, T. Kamm, M. Ordowski, and M. Przybocki, "The DET curve in assessment of detection task performance," DTIC Document, Tech. Rep., 1997.

[49] C. E. Metz, "Basic principles of ROC analysis," in *Seminars in nuclear medicine*, vol. 8, no. 4. Elsevier, 1978, pp. 283–298.

[50] T. Fawcett, "An introduction to ROC analysis," *Pattern recognition lett.*, vol. 27, no. 8, pp. 861–874, 2006.

[51] S. Ryan, A. Schachat, C. Wilkinson, D. Hinton, S. Sadda, and P. Wiedemann, *Retina*. Elsevier Health Sciences, 2012. [Online]. Available: https://books.google.co.uk/books?id=PdAsuzFRv50C [Accessed: 07/08/2017]

[52] A. Jain, R. Bolle, and S. Pankanti, *Biometrics: personal identification in networked society*. Springer Science & Business Media, 2006, vol. 479.

[53] H. Farzin, H. Abrishami-Moghaddam, and M.-S. Moin, "A novel retinal identification system," *EURASIP J. on Advances in Signal Process.*, vol. 2008, no. 1, p. 280635, 2008.

[54] N. A. Rahman, A. S. Mohamed, and M. E. Rasmy, "Retinal identification," in *Cairo Int. Conf. on Biomedical Eng.*, 2008, pp. 1–4.

[55] M. B. Patwari, R. R. Manza, Y. M. Rajput, M. Saswade, and N. Deshpande, "Personal identification algorithm based on retinal blood vessels bifurcation," in *Int. Conf. on Intell. Computing Applicat.*, 2014, pp. 203–207.

[56] M. Watanabe, T. Endoh, M. Shiohara, and S. Sasaki, "Palm vein authentication technology and its applications," in *Proc. of the Biometric Consortium Conf.* Citeseer, 2005, pp. 19–21.

[57] C. Wilson, "Vein pattern recognition," *A Privacy-Enhancing Biometric. Boca Raton: CRC*, 2010.

[58] C.-L. Lin and K.-C. Fan, "Biometric verification using thermal images of palm-dorsa vein patterns," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 2, pp. 199–213, 2004.

[59] Y. Zhou and A. Kumar, "Human identification using palm-vein images," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 4, pp. 1259–1274, 2011.

[60] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and intell. laboratory syst.*, vol. 2, no. 1-3, pp. 37–52, 1987.

[61] J. Hashimoto, "Finger vein authentication technology and its future," in *Symp. on VLSI Circuits*, 2006, pp. 5–8.

[62] N. Miura, A. Nagasaka, and T. Miyatake, "Extraction of finger-vein patterns using maximum curvature points in image profiles," *IEICE Trans. on Inform. and Syst.*, vol. 90, no. 8, pp. 1185–1194, 2007.

[63] J. Yang, Y. Shi, and G. Jia, "Finger-vein image matching based on adaptive curve transformation," *Pattern Recognition*, vol. 66, pp. 34 – 43, 2017.

[64] P. Lü and S. Lai, "Image enhancement and feature matching of palm-dorsa vein," *Comput. Eng. and Design*, vol. 28, no. 15, pp. 3623–3625, 2007.

[65] C.-l. Mei, X. Xiao, G.-h. Liu, Y. Chen, and Q.-a. Li, "Feature extraction of finger-vein image based on morphologic algorithm," in *Sixth Int. Conf. on Fuzzy Syst. and Knowledge Discovery (FSKD)*, vol. 3. IEEE, 2009, pp. 407–411.

[66] N. Miura, A. Nagasaka, and T. Miyatake, "Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification," *Mach. Vision and Applicat.*, vol. 15, no. 4, pp. 194–203, 2004.

[67] M. Tistarelli and M. S. Nixon, *Advances in Biometrics: Third International Conferences, ICB 2009, Alghero, Italy, June 2-5, 2009, Proceedings*. Springer, 2009, vol. 5558.

[68] C. Oyster, *The Human Eye: Structure And Function*. Sinauer Associates Incorporated, 1999.

[69] R. Broekhuyse, "The lipid composition of aging sclera and cornea," *Ophthalmologica*, vol. 171, no. 1, pp. 82–85, 1975.

[70] C. G. Owen, T. J. Ellis, A. R. Rudnicka, and E. G. Woodward, "Optimal green (red-free) digital imaging of conjunctival vasculature," *Ophthalmic and Physiological Optics*, vol. 22, no. 3, pp. 234–243, 2002.

[71] R. Derakhshani and A. Ross, "A texture-based neural network classifier for biometric identification using ocular surface vasculature," in *Int. Joint Conf. on Neural Networks*, 2007, pp. 2982–2987.

[72] S. Crihalmeanu and R. Derakhshani, "Enhancement and registration schemes for matching conjunctival vasculature," in *Proc. of the 3rd IAPR/IEEE Int. Conf. on Biometrics (ICB)*, 2009, pp. 1240–1249.

[73] N. L. Thomas, “A new approach for human identification using the eye,” Master’s thesis, Purdue University, Indianapolis, Indiana, US, 2010.

[74] Z. Zhou, E. Y. Du, N. L. Thomas, and E. J. Delp, “A comprehensive approach for sclera image quality measure,” *Int. J. of Biometrics*, vol. 5, no. 2, pp. 181–198, 2013.

[75] S. P. Tankasala, P. Doynov, R. R. Derakhshani, A. Ross, and S. Crihalmeanu, “Biometric recognition of conjunctival vasculature using GLCM features,” in *Int. Conf. on Image Inform. Process.*, 2011, pp. 1–6.

[76] S. Crihalmeanu and A. Ross, “On the use of multispectral conjunctival vasculature as a soft biometric,” in *IEEE Workshop on Applicat. of Computer Vision (WACV)*, 2011, pp. 204–211.

[77] Z. Zhou, E. Y. Du, N. L. Thomas, and E. J. Delp, “Multi-angle sclera recognition system,” in *IEEE Workshop on Computational Intell. in Biometrics and Identity Manag. (CIBIM)*, 2011, pp. 103–108.

[78] K. Oh and K. A. Toh, “Extracting sclera features for cancelable identity verification,” in *5th IAPR Int. Conf. on Biometrics (ICB)*, 2012, pp. 245–250.

[79] V. Gottemukkula, S. K. Saripalle, S. P. Tankasala, R. Derakhshani, R. Pasula, and A. Ross, “Fusing iris and conjunctival vasculature: Ocular biometrics in the visible spectrum,” in *IEEE Conf. on Technologies for Homeland Security (HST)*, 2012, pp. 150–155.

[80] S. Crihalmeanu and A. Ross, “Multispectral scleral patterns for ocular biometric recognition,” *Pattern Recognition Lett.*, vol. 33, no. 14, pp. 1860–1869, 2012.

[81] S. P. Tankasala, P. Doynov, and R. Derakhshani, “Visible spectrum, bi-modal ocular biometrics,” *Procedia Technology*, vol. 6, pp. 564–573, 2012.

[82] A. Das, U. Pal, M. A. F. Ballester, and M. Blumenstein, “Sclera recognition using dense-SIFT,” in *13th Int. Conf. on Intell. Syst. Design and Applicat.*, 2013, pp. 74–79.

[83] S. P. Tankasala, P. Doynov, and R. Derakhshani, "Application of pyramidal directional filters for biometric identification using conjunctival vasculature patterns," in *IEEE Int. Conf. on Technologies for Homeland Security (HST)*, 2013, pp. 639–644.

[84] A. Das, U. Pal, M. A. F. Ballester, and M. Blumenstein, "A new method for sclera vessel recognition using OLBP," in *Biometric Recognition*. Springer, 2013, pp. 370–377.

[85] Y. Lin, E. Y. Du, Z. Zhou, and N. L. Thomas, "An efficient parallel approach for sclera vein recognition," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 2, pp. 147–157, 2014.

[86] A. Das, U. Pal, M. A. F. Ballester, and M. Blumenstein, "A new efficient and adaptive sclera recognition system," in *IEEE Symp. on Computational Intell. in Biometrics and Identity Management (CIBIM)*, 2014, pp. 1–8.

[87] S. P. Tankasala and P. Doynov, "Multi scale multi directional shear operator for personal recognition using conjunctival vasculature," in *IEEE Int. Symp. on Technologies for Homeland Security (HST)*, 2015, pp. 1–6.

[88] V. Gottemukkula, S. Saripalle, S. P. Tankasala, and R. Derakhshani, "Method for using visible ocular vasculature for mobile biometrics," *IET Biometrics*, vol. 5, no. 1, pp. 3–12, 2016.

[89] Q. Li, S. Sone *et al.*, "Selective enhancement filters for nodules, vessels, and airway walls in two-and three-dimensional CT scans," *Medical physics*, vol. 30, no. 8, pp. 2040–2051, 2003.

[90] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, 1979.

[91] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. on Inform. Theory*, vol. 8, no. 2, pp. 179–187, 1962.

[92] D. Wei, J. Tian, R. Wells, and C. S. Burrus, "A new class of biorthogonal wavelet systems for image transform coding," *IEEE Trans. Image Process.*, vol. 7, no. 7, pp. 1000–1013, 1998.

[93] A. Likas, N. Vlassis, and J. J. Verbeek, "The global K-means clustering algorithm," *Pattern recognition*, vol. 36, no. 2, pp. 451–461, 2003.

[94] S. Periaswamy and H. Farid, "Elastic registration in the presence of intensity variations," *IEEE Trans. Med. Imag.*, vol. 22, no. 7, pp. 865–874, 2003.

[95] Z. Zhou, E. Y. Du, C. Belcher, N. L. Thomas, and E. J. Delp, "Quality fusion based multimodal eye recognition," in *IEEE Int. Conf. on Syst., Man, and Cybern. (SMC)*, 2012, pp. 1297–1302.

[96] H. Proença and L. A. Alexandre, "UBIRIS: A noisy iris image database," in *Int. Conf. on Image Anal. and Process. (ICIAP)*. Springer, 2005, pp. 970–977.

[97] R. M. Haralick, K. Shanmugam *et al.*, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, no. 6, pp. 610–621, 1973.

[98] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Human Genetics*, vol. 7, no. 2, pp. 179–188, 1936.

[99] R. S. Hunter, "Photoelectric color difference meter," *J. of Optical Soc. of America*, vol. 48, no. 12, pp. 985–995, 1958.

[100] K. Oh, B.-S. Oh, K.-A. Toh, W.-Y. Yau, and H.-L. Eng, "Combining sclera and periocular features for multi-modal identity verification," *Neurocomputing*, vol. 128, pp. 185–198, 2014.

[101] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, 2006.

[102] I. W. Selesnick, "A new complex-directional wavelet transform and its application to image denoising," in *Proc. Int. Conf. on Image Process.*, vol. 3, 2002, pp. 573–576.

[103] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[104] W. Bu, X. Wu, and E. Gao, "Hand vein recognition based on orientation of LBP," *SPIE Defense, Security, and Sensing*, 2012.

[105] B. N. Li, C. K. Chui, S. Chang, and S. H. Ong, "Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation," *Comput. in Biology and Medicine*, vol. 41, no. 1, pp. 1–10, 2011.

[106] A. Das, U. Pal, M. A. F. Ballester, and M. Blumenstein, "Fuzzy logic based selera recognition," in *IEEE Int. Conf. on Fuzzy Syst. (FUZZ-IEEE)*, 2014, pp. 561–568.

[107] I. Daubechies, *Ten lectures on Wavelets* . SIAM, 1992, vol. 61.

[108] M. H. Kabir, T. Jabid, and O. Chae, "A local directional pattern variance (LDPV) based face descriptor for human facial expression recognition," in *7th IEEE Int. Conf. on Advanced Video and Signal Based Surveillance*, 2010, pp. 526–532.

[109] P. Radu, J. Ferryman, and P. Wild, "A robust sclera segmentation algorithm," in *IEEE 7th Int. Conf. on Biometrics Theory, Applicat. and Syst. (BTAS)*, 2015, pp. 1–6.

[110] C. Harris and M. Stephens, "A combined corner and edge detector." in *Alvey Vision Conf.*, vol. 15, Manchester, UK, 1988, p. 50.

[111] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," *Compu. Vision–ECCV*, pp. 430–443, 2006.

[112] H. Proenca, S. Filipe, R. Santos, J. Oliveira, and L. A. Alexandre, "The UBIRIS.v2: A database of visible wavelength iris images captured on-the-move and at-a-distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 8, pp. 1529–1535, 2010.

[113] M. Hosseini, B. Araabi, and H. Soltanian-Zadeh, "Pigment melanin: Pattern for iris recognition," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 4, pp. 792–804, 2010.

[114] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler, "Mobile iris challenge evaluation (MICHE)-I, biometric iris dataset and protocols," *Pattern Recognition Lett.*, vol. 57, pp. 17–23, 2015.

[115] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, 2001.

[116] J. Daugman, "New methods in iris recognition," *IEEE Trans. Syst., Man, Cybern. B*, vol. 37, no. 5, pp. 1167–1175, 2007.

[117] W. W. Boles and B. Boashash, "A human identification technique using images of the iris and wavelet transform," *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 1185–1188, 1998.

[118] L. Ma, T. Tan, Y. Wang, and D. Zhang, "Efficient iris recognition by characterizing key local variations," *IEEE Trans. Image Process.*, vol. 13, no. 6, pp. 739–750, 2004.

[119] R. P. Wildes, "Iris recognition: an emerging biometric technology," *Proc. of the IEEE*, vol. 85, no. 9, pp. 1348–1363, 1997.

[120] J. Kovac, P. Peer, and F. Solina, "2D versus 3D colour space face detection," in *Proc. EC-VIP-MC 2003. 4th EURASIP Conf. focused on Video/Image Process. and Multimedia Commun. (IEEE Cat. No.03EX667)*, vol. 2, 2003, pp. 449–454 vol.2.

[121] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. of Comput. Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[122] V. Caselles, F. Catté, T. Coll, and F. Dibos, "A geometric model for active contours in image processing," *Numerische Mathematik*, vol. 66, no. 1, pp. 1–31, 1993.

[123] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Topology-independent shape modeling scheme," in *Int. Symp. on Optics, Imaging, and Instrumentation, SPIE*. Int. Soc. for Optics and Photonics, 1993, pp. 246–258.

[124] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *Int. J. of Comput. Vision*, vol. 22, no. 1, pp. 61–79, 1997.

[125] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Commun. on Pure and Applied Mathematics*, vol. 42, no. 5, pp. 577–685, 1989.

[126] J.-L. Starck, J. Fadili, and F. Murtagh, "The undecimated wavelet decomposition and its reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 297–309, 2007.

[127] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3136–3144, 2001.

[128] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun. of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[129] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.

[130] L. D. Cohen, "On active contour models and balloons," *CVGIP: Image Understanding*, vol. 53, no. 2, pp. 211–218, 1991.

[131] N. Paragios, O. Mellina-Gottardo, and V. Ramesh, "Gradient vector flow fast geodesic active contours," in *Proc. Eighth IEEE Int. Conf. on Comput. Vision. ICCV*, vol. 1, 2001, pp. 67–73.

[132] H. Proenca, "Iris recognition: On the segmentation of degraded images acquired in the visible wavelength," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 8, pp. 1502–1516, 2010.

[133] N. L. Thomas, Y. Du, and Z. Zhou, "A new approach for sclera vein recognition," in *SPIE Defense, Security, and Sensing*. Int. Soc. for Optics and Photonics, 2010.

[134] H. Proenca and L. Alexandre, "Toward noncooperative iris recognition: A classification approach using multiple signatures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 607–612, 2007.

[135] G. Gomez and E. Morales, "Automatic feature construction and a simple rule induction algorithm for skin detection," in *Proc. of the ICML Workshop on Mach. Learning in Comput. Vision*, 2002, pp. 31–38.

[136] S. Tsekeridou and I. Pitas, "Facial feature extraction in frontal views using biometric analogies," in *9th European Signal Proc. Conf. (EUSIPCO)*, 1998, pp. 1–4.

[137] C. Garcia and G. Tziritas, "Face detection using quantized skin color regions merging and wavelet packet analysis," *IEEE Trans. Multimedia*, vol. 1, no. 3, pp. 264–277, 1999.

[138] S. Hsieh, K.-C. Fan, and C. Lin, "A statistic approach to the detection of human faces in color nature scene," *Pattern Recognition*, vol. 35, no. 7, pp. 1583–1596, 2002.

[139] D. Chai and K. N. Ngan, "Face segmentation using skin-color map in videophone applications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 4, pp. 551–564, 1999.

[140] P. Soille, *Morphological Image Analysis: Principles and Applications*. Springer Science & Business Media, 2013.

[141] I. W. Selesnick, R. G. Baraniuk, and N. G. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Proc. Mag.*, vol. 22, no. 6, pp. 123–151, 2005.

[142] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Wurtz, and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," *IEEE Trans. Comput.*, vol. 42, no. 3, pp. 300–311, 1993.

[143] S. Mallat, "Wavelets for a vision," *Proc. of the IEEE*, vol. 84, no. 4, pp. 604–614, 1996.

[144] L. Shen and L. Bai, "A review on Gabor wavelets for face recognition," *Pattern Anal. and Applicat.*, vol. 9, no. 2-3, pp. 273–292, 2006.

[145] R. Gonzalez and R. Woods, *Digital Image Processing*, ser. Pearson education. Pearson/Prentice Hall, 2008. [Online]. Available: https://books.google.co.uk/books?id=EtLdPQAACAAJ [Accessed: 04/08/2017]

[146] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 467–476, 2002.

[147] C. Liu, "Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 5, pp. 725–737, 2006.

[148] B. Zhang, S. Shan, X. Chen, and W. Gao, "Histogram of Gabor phase patterns (HGPP): A novel object representation approach for face recognition," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 57–68, 2007.

[149] A. Cohen, I. Daubechies, and J.-C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Commun. on pure and applied mathematics*, vol. 45, no. 5, pp. 485–560, 1992.

[150] J. N. Bradley, C. M. Brislawn, and T. Hopper, "The FBI wavelet/scalar quantization fingerprint image compression standard," in *Proc. Conf. Solid-State Memory Tech.,(Pasadena, CA) A*, vol. 11, 1994.

[151] B. Scholkopft and K.-R. Mullert, "Fisher discriminant analysis with kernels," in *Proc. of the 1999 IEEE Signal Process. Soc. Workshop Neural Networks for Signal Process. IX, Madison, WI, USA*, 1999, pp. 23–25.

[152] S. Saitoh, *Reproducing Kernels and Their Applications*. Springer Science & Business Media, 1999, vol. 3.

[153] V. Štruc and N. Pavešić, "Gabor-based kernel partial-least-squares discrimination features for face recognition," *Informatica*, vol. 20, no. 1, pp. 115–138, 2009.

[154] B. Scholkopf, A. Smola, and K.-R. Muller, "Kernel principal component analysis," in *Artificial Neural Networks, ICANN'97*. Springer, 1997, pp. 583–588.

[155] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. of Comput. Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[156] Z. Zhou, E. Y. Du, and N. L. Thomas, "A comprehensive sclera image quality measure," in *11th Int. Conf. on Control Automation Robotics Vision*, 2010, pp. 638–643.

[157] J. Shi, "Good features to track," in *Proc. IEEE Comput. Soc. Conf. on Comput. Vision and Pattern Recognition (CVPR)*, 1994, pp. 593–600.

[158] P. J. Phillips, J. R. Beveridge, B. A. Draper, G. Givens, A. J. O'Toole, D. S. Bolme, J. Dunlop, Y. M. Lui, H. Sahibzada, and S. Weimer, "An introduction to the good, the bad, & the ugly face recognition challenge problem," in *Face and Gesture*, 2011, pp. 346–353.

[159] J. L. Pech-Pacheco, G. Cristobal, J. Chamorro-Martinez, and J. Fernandez-Valdivia, "Diatom autofocusing in brightfield microscopy: a comparative study," in *Proc. 15th Int. Conf. on Pattern Recognition ICPR*, vol. 3, 2000, pp. 314–317.

[160] G. Santos and H. ProenÃ§a, "A robust eye-corner detection method for real-world data," in *Int. Joint Conf. on Biometrics (IJCB)*, 2011, pp. 1–7.