ANA FILIPA PINHEIRO SEQUEIRA

Liveness Detection and Robust Recognition in Iris and Fingerprint Biometric Systems

Ph.D. Thesis

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

June 2015

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Liveness Detection and Robust Recognition in Iris and Fingerprint Biometric Systems

Ana Filipa Pinheiro Sequeira

Programa Doutoral em Engenharia Eletrotécnica e de Computadores

Supervisor: Jaime dos Santos Cardoso (PhD)

June 2015

Resumo

A Biometria constitui um regresso a uma forma mais tradicional de identificar alguém com base no que essa pessoa é em vez de ser baseada naquilo que essa pessoa sabe ou possui. Apesar de já existir uma pesquisa significativa nesta área, ainda há muito a fazer considerando que novos cenários de aplicabilidade aparecem a cada dia. Os sistemas de reconhecimento biométrico já não estão confinados à investigação forense ou ao controlo de funcionários. Estes sistemas têm vindo a ganhar uma visibilidade e aplicabilidade nos dispositivos de uso diário que reforçam a sua usabilidade em todos os aspetos da nossa rotina.

Com esta expansão de aplicações biométricas, encontradas hoje em dia nos nossos computadores pessoais, *smarthphones*, serviços de gestão bancária e de controlo nos aeroportos, há uma crescente necessidade de uma melhor segurança. A importância de proteger a nossa identidade e as nossas informações tornou-se crucial devido a os nossos dispositivos se tornarem repositórios de informação sensível a muitos níveis.

Assim, os métodos de deteção de vivacidade como medidas contra os *spoffing attacks* são mais importantes do que nunca. Novos métodos devem ser desenvolvidos que se debrucem sobre os novos cenários de aquisição de dados e que lidem com o crescente ruído nos dados biométricos recolhidos.

É então de extrema importância o desenvolvimento de métodos de deteção de vivacidade robustos. Em particular, trabalhámos com íris e impressão digital. Estes dois traços biométricos são muitas vezes escolhidos pelas suas características particulares. Entre os objetivos desta tese estavam o propósito de fazer contribuições na deteção de vivacidade relativamente a íris e impressão digital propondo novas abordagens quer do ponto de vista dos cenários de aquisição de imagem, no caso da íris, quer na abordagem usada para a classificação, no caso da impressão digital. Foram feitas contribuições relativamente a ambos os traços, que excederam o estado da arte e resultaram em publicações tanto de conferência como de revista.

Não só os *spoffing attacks* preocupam os investigadores mas também a capacidade dos métodos de lidarem com dados ruidosos. O desenvolvimento de métodos robustos que ultrapassem o facto de a qualidade dos dados estar comprometida é uma necessidade da pesquisa biométrica dos nossos dias. Assim, um outro objetivo desta tese era contribuir para o reconhecimento por impressão digital desenvolvendo métodos robustos para extração de minúcias. O trabalho desenvolvido resultou na proposta de um método para a estimação do mapa de orientação e um método de realce de uma impressão digital que melhoraram os resultados existentes.

Este trabalho pretendeu e foi bem sucedido em propôr métodos robustos e realistas tanto no problema da deteção de vivacidade em íris e impressão digital como em alguns passos do reconhecimento por impressão digital. É de ressalvar que o foco de atenção deste trabalho é a qualidade dos dados e não a eficiência computacional, e esta teria de ser tida em conta ao ambicionar uma implementanção dos métodos em sistemas aplicados ao mundo real.

Outro objetivo deste trabalho foi criar novas bases de dados e plataformas comuns de avaliação de métodos tais como competições biométricas. Assim, durante o trabalho desenvolvido, foram

construídas duas bases de dados e foram organizadas duas competições biomítricas. Ambas as bases de dados tiveram um grande impacto na comunidade científica e continuam a ser disseminadas. Publicações usando estes *datasets* são já numerosas e continuam a aparecer regularmente.

Abstract

Biometrics represents a return to a traditional way of identifying someone relying on what that person is instead of what that person knows or owns. Even though the significant amount of research that has been done in this field, there is still much to do as new emerging scenarios of application appear everyday. Biometric recognition systems are no longer restricted to forensic investigation or control management of employees. They have been gaining a visibility and applicability in daily use devices which reinforces their usability in all aspects of our day to day life.

With this spread of biometric applications, nowadays commonly found in our laptops, our smart phones, some bank management services and airport custom services, a necessity for improved security also is rising. The importance of protecting our identity and our data has become crucial as our devices are filled with sensible information of many kinds.

Therefore the *presentation attack* or *liveness detection* methods as countermeasures against spoofing attacks are more important than ever. New methods should be developed which address the new acquisition scenarios and which deal with the increased noise in the biometric data collected.

Its of utmost importance to develop robust liveness detection methods. In particular, we worked on iris and fingerprint. These two biometric traits are very often chosen against others due to its characteristics. Among the objectives of this thesis were the purpose of making contributions in iris and fingerprint liveness detection proposing novel approaches whether from the imaging scenarios perspective, in the case of iris, or from the classification approach, in the case of fingerprint. Contributions were made regarding both traits, that exceeded the state-of-the art and resulted in both conferences and journal publications.

Not only the spoofing attacks concern the biometric researchers but also the ability of the methods to deal with the noisy data. Therefore, the development of robust methods that overcome the compromised quality of data is a necessity of biometric research of nowadays. Therefore, another objective was to contribute to the fingerprint recognition problem developing robust methods to minutiae extraction. The work developed resulted in a proposed method for fingerprint orientation map estimation and a fingerprint image enhancement that over performed existing ones.

This work aimed and succeeded to propose robust and realistic methods in both the iris and fingerprint liveness detection problem as well as in some steps of fingerprint recognition. It has to be noted that the focus of attention of this work was the quality of data and not the computational efficiency, therefore this one should have to be addressed if an application of the proposed methods to a real-world scenario was aimed.

Another objective was to create new databases and promote common platforms of evaluation of methods such as biometric competitions. Therefore, along the work developed, two biometric databases were constructed and two biometric competitions were organized. Both databases had a strong impact in the research community and they continue to be disseminated. Publications using these benchmark datasets are numerous and continue to appear regularly.

Acknowledgments

I would like to express my gratitude to Professor Jaime S. Cardoso for everything that he taught me and for setting an admirable example both in his professional conduct and regarding his human qualities. I am forever grateful for all the support, encouragement and constant availability during the development of my work, that went far beyond the role of a supervisor.
I extend my gratitude to INESC TEC for all the conditions provided. I am very glad to have had the possibility of being embraced by this institution.
All the past and present members of the "Visual Computing and Machine Intelligence" (VCMI) group played a unique role in my professional and personal path during this time. I could say in which manner each one of you has helped and inspired me in a particular way and at a different moment. I will keep these memories with me wherever we go in the future. Nevertheless, a very special word of gratitude is due to Ana Rebelo, Inês Domingues, Hélder Oliveira, Ricardo Sousa, João P. Monteiro, João C. Monteiro, Eduardo Marques and Pedro Ferreira.
I shall also dedicate a few words to someone whose influence in my life spanned out the time I was his student in the degree and master. The way I reason about Mathematics, Science, Knowledge and Life itself was deeply influenced by the many things he taught me about, the many things we changed opinions about and the many things I wondered about after all that. Thank you Professor António Machiavelo.
With each one of my dear, dear friends I share all the small, medium and big steps in my life. This is one more step I took with you holding my hand, watching my back or simply bringing joy to my days. I am blessed for having you all and your presence in my life is the fuel that keeps me going. Thank you.
To my close family that have always suffered with my absence and all the consequences of my choices in life, I thank the fact that, even at a distance, you were there sharing every moment with me, good or bad, and encouraging me. Without your support, everything would have been so much heavier.

To my Mother and Father, to whom I belong and to whom I owe everything, I thank for being the certainty in my life.

The work developed was not possible without the support provided by FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology), with Ph.D grant (Grant ID - SFRH / BD / 74263 / 2010).

"Science and mathematics [are] much more compelling and exciting than the doctrines of pseudoscience [...]
But science is more intricate and subtle, reveals a much richer universe, and powerfully evokes our sense of wonder. And it has the additional and important virtue - to whatever extent the word has any meaning - of being true."

Carl Sagan, *in* Broca's Brain (1986)

Contents

List of Figures

List of Tables

Acronyms

2D	Two-dimensional
3D	Three-dimensional
APCER	Attack Presentation Classification Error Rate
ACER	Average Classification Error Rate
BSIF	Binarized Statistical Image Features
CN	Convolutional Networks
CNN	Convolutional Neural Networks
CHT	Circular Hough Transform
DoG	Difference of Gaussian
FA	False Acceptances
FAR	False Acceptance Rate
FFT	Faste Fourier Transform
FIR	False Identification Rate
FNR	False non-pupil Rate
FPR	False pupil Rate
FR	False Rejections
FRR	False Rejection Rate
FVC	Fingerprint Verification Competition
GLCM	Gray Level Co-Occurrence Matrices
GMM	Mixture of Gaussian Models
HIG	Histograms of Invariant Gradients
HOG	Histograms of Oriented Gradients
HVC	Hierarchical Visual Codebook
KNN	K Nearest Neighbors
IR	Infra-Red
LBP	Local Binary Patterns
LDA	Liner Discriminant Analysis
LPQ	Local Phase Quantization
MCC	Minutia Cylinder Code
MRF	Markov Random Field
NIR	Near Infra Red
NIST	National Institute of Standards and Technology
NPCER	Normal Presentation Classification Error Rate
OCSVM	One Class Support Vector Machine
PAD	Presentation Attack Detection
PCA	Principal Component Analysis
PIN	Personal Identification Number

RBF	Radial Basis Function
ROI	Region-of-Interest
SIFT	Scale-Invariant Feature Transform
SFFS	Sequential Float Forward Selection
SFS	Sequential Forward Selection
STFT	Short Time Fourier Transform
SVM	Support Vector Machine
TA	True Acceptances
TR	True Rejections
VL	Visible Light
VW	Visible Wavelength
WLBP	Weighted Local Binary Patterns
WLD	Weber Local Descriptor

Part I

Introduction and State of the Art

Chapter 1

Introduction

Reliable automatic recognition of persons has long been an attractive goal. In most of all daily activities, personal identification plays an important role. The most traditional techniques for personal identification can be divided in two kinds: knowledge-based and token-based. In one hand, token-based approaches take advantage of a personal item, like a passport, driver's license, ID card, credit card or a simple set of keys, on the other hand, knowledge-based approaches, are based on something the user knows that, theoretically, are not accessible to others, such as passwords or personal identification numbers. These approaches present the obvious disadvantage that tokens may be lost, stolen, forgotten or misplaced, while passwords can easily be forgotten by a valid user or guessed by an unauthorized one [112]. In fact, all of these approaches stumble upon an obvious problem: any piece of material or knowledge can be fraudulently acquired. So, biometrics represents a return to a more natural way of identification: many physiological or behavioral characteristics are unique between different persons. Testing someone by what this someone is, instead of relying on something he owns or knows seems likely to be the way forward.

As in all pattern recognition problems, the key issue is the relation between inter-class and intra-class variability: objects can be reliably classified only if the variability among different instances of a given class is less than the variability between different classes. For example, in face recognition, difficulties arise from the fact that the face is a changeable social organ displaying a variety of expressions, as well as being an active three-dimensional (3-D) object whose image varies with viewing angle, pose, illumination, accouterments, and age. Against this intra-class (same face) variability, inter-class variability is limited because different faces possess the same basic set of features, in the same canonical geometry [56]. Nevertheless, several biological traits in humans show a considerable inter-individual variability: fingerprints and palm prints, the shape of the ears, the pattern of the iris, among others. Biometrics works by recognizing patterns within these biological traits, unique to each individual, to increase the reliability of recognition.

1.1 Motivation

Iris biometrics research is an exciting, broad and rapidly expanding field. At the same time that there are successful practical applications that illustrate the power of iris biometrics, there are also many fundamental research issues to be solved on the way to larger scale and more complex applications [129]. The first iris based recognition systems were highly efficient assuming very controled imaging acquisition scenarios. A step forward was to broad the application of these systems to less controlled environments. One example is the mobile biometrics scenario. The spread of handheld devices equipped with cameras with increased quality has been a motivation for implementing iris recognition systems in this devices.

But nowadays, not only is iris playing a role in the mobile biometrics scenario, but also fingerprint based recognition systems are being implemented in handheld devices. This biometric trait was the first one to be used as a systematic way of identification in very controlled conditions and is now being applied in more unconstrained and challeging scenarios such as the identification of users in handheld devices.

Along with the challenges raised by the unconstrained conditions in the image acquisition also several security questions appear such as the vulnerability of biometric recognition systems to attacks consisting on presenting a fake or altered sample to the sensor. These kind of spoofing attacks are nowadays commonly known as “presentation attacks”.

Observing the emerging mobile biometric field, the urge for improving the security of the biometric systems implemented in mobile devices leads to an increased interest in the development of presentation attack detection methods or, in particular, liveness detection methods in these new imaging scenarios. Concerning iris liveness detection, the main problem is adjusting the methods to the new imaging scenarios for the most common and easily perpetrated attacks that are the use of iris printed images or the use of patterned contact lenses. For this purpose suitable databases are needed so that the techniques can be evaluated in images captured by handheld devices. This was a necessity we found and tried to fulfill by constructing new databases. The evaluation of iris liveness detection techniques regarding printed iris images and contact lenses iris images is one of the problems addressed in this thesis.

Regarding fingerprints, despite the interest raised by recent deployment of devices equipped with fingerprint recognition systems, there are not yet databases available with images specifically acquired in the mobile scenario. Therefore, the work developed used the available databases acquired with the traditional sensors. Nevertheless, we keep in mind that the path to pursue is to adjust the methods to this new challenging scenario. It is expected that in a near future new databases will become available with such characteristics.

The work developed in fingerprint liveness detection, was motivated by the observations of an increasing number of well documented techniques used to build fake samples of fingers. So, when we considering fingerprint spoofing attacks, it can be observed a growing variety and sophistication of spoofing materials. It seemed to us a necessity to develop methods that do not rely so much in the prior knowledge about these materials as the traditional approach usually does. Therefore,

a semi-supervised classification approach is proposed to tackle this problem.

On the fingerprint recognition problem, the work done comprises the proposal of a new method for the orientation map estimation based on Markov Random Fields. This method proved to improve the results of a recognition system for some of the tested datasets. Also an enhancement step consisting on a novel approach based on the curved gabor signal is explored.

Transversal to the use of any of the several available biometric traits, is the necessity for common platforms which allow a fair comparison between the developed methodologies. Therefore, it is of utmost importance the construction of databases, as well as the development of adequate metrics and experimental protocols. The biometric competitions are excellent opportunities to motivate the development of new methodologies and, above all, to evaluate these new methods using common benchmarks, metrics and protocols. Having this in mind, the work developed comprised the construction of two databases and the organization of two biometric competitions.

The first database constructed was a multimodal database with samples of face, iris and voice acquired with an handheld device. The construction of this database was motivated by the aim to have a benchmark which would allow the evaluation of methods adapted to this more recent context of of biometric recognition, the mobile biometrics scenario. The biometric traits were chosen due to their complementarity and also considering their adaptability to the functions available in actual mobile devices which are equipped with cameras and sound recording systems.

Nevertheless, we point out that the multimodal scenario was not the main focus of the present thesis. We may consider the construction of this database as a first approach to the mobile biometrics problem. The approach to mobile biometrics was later pursued regarding iris through the use of the iris subset. On one hand, the set of iris images by itself allowed the evaluation of methods for segmentation and recognition more adjusted to the limitations of images captured with this type of devices; and on the other hand, allowed the construction of set of printed iris fake samples captured with an handheld device which at the time appeared as a novelty.

The following direction in this work had fingerprint as the biometric trait used. This biometric trait was not included in the initial work regarding the multimodal database constructed. However, the fact that fingerprint has recently been introduced in mobile devices for authentication was a good motivation for the posterior work which includes research in liveness detection and recognition for fingerprint recognition systems.

1.2 Objectives

The objectives of this thesis were to make contributions in iris and fingerprint liveness detection using novel approaches whether from the imaging scenarios perspective, in the case of iris, or from the classification approach, in the case of fingerprint. Another objective was to contribute to the fingerprint recognition problem developing robust methods to minutiae extraction. Another objective was to create new databases and promote common platforms of evaluation of methods such as competitions. Along the work developed, two biometric databases were constructed and two biometric competitions were organized. Particularly concerning the Iris Liveness Detection

problem, the aim to impact the research community was achieved through the dissemination of the database and the contributions made by researchers from all over the world with their methods. But also the Multimodal Database continues to interest researchers in the biometrics field whether as a whole or by its subsets regarding each biometric trait.

1.3 Contributions

We summarise below the contributions of the thesis to the iris and fingerprint liveness detection field; the fingerprint recognition field; and the biometrics field in general by constructing databases and organizing competitions. In this thesis we have:

1. created a multimodal database captured with an handheld device with samples of face, iris and speech which was used as benchmark in a multimodal biometric competition we conducted;

2. created an iris database with printed fake samples captured with an handheld device which was used as benchmark in an iris liveness detection competition we conducted (embraced by a top conference in the biometric field);

3. proposed a new methodology which combined feature extraction methods for iris liveness detection and applied the methodology in mobile imaging scenarios as well as other benchmark datasets (we classified as first place in an Iris Liveness Detection Competition with the proposed methodology);

4. proposed a new approach in fingerprint liveness detection based on semi-supervised classification and compared it with the more traditional approaches based on supervised classification;

5. proposed a new method for direction map estimation of a fingerprint based on Markov Random Field and the evaluation of its impact on a fingerprint recognition system showed an improvement in the overall recognition results;

6. proposed a robust binarization method, consisting on an enhancement based on the use of a curved gaussian filter for a longitudinal enhancement combined with a transversal enhancement based on curved bandpass filter.

List of Publications Related to the thesis

The work related with this thesis resulted in the followings journal paper:

- Pedro Ferreira, Ana F. Sequeira, Ana Rebelo, and Jaime S. Cardoso. Robust Orientation Map Estimation and Robust Clustering-based Segmentation Methods for Fingerprint Recognition. In International Journal on Neural Computing and Applications, (submitted).

- Ana F. Sequeira and Jaime S. Cardoso. Fingerprint Liveness Detection in the Presence of Capable Intruders. Sensors: Physical Sensors. June, 2015. DOI: 10.3390/s150614615.

- João C. Monteiro, Ana F. Sequeira, Hélder P. Oliveira, Jaime S. Cardoso. Computer Vision, Imaging and Computer Graphics: Theory and Applications, Chapter in Robust Iris Localisation in Challenging Scenarios (Sebastiano Battiano, Sabine Coquillart, Robert Laramee, Andreas Kerren, Jose Braz, eds.), Springer, 2014.

From this thesis resulted the following international conference papers:

- Ana F. Sequeira, Hélder P. Oliveira , João C. Monteiro, João P. Monteiro and Jaime S. Cardoso, MobILive 2014 - Mobile Iris Liveness Detection Competition, In Proceedings of the IEEE Int. Joint Conference on Biometrics (IJCB), pp. 1-6, Clearwater, Florida, USA, 30 September – 02 October, 2014.

- Ana F. Sequeira, Juliano Murari and Jaime S. Cardoso, Iris liveness detection methods in mobile biometrics scenario, In Proceedings of the International Joint Conference on Neural Networks (IJCNN), pp.3002-3008, Beijing, China, 6 - 11 July, 2014.

- Ana F. Sequeira, João C. Monteiro, Ana Rebelo and Hélder P. Oliveira, MobBIO: a Multimodal Database Captured with a Portable Handheld Device, In Proceedings of the 9^{th} International Conference on Computer Vision Theory and Applications (VISAPP), pp.133-139, Lisboa, Portugal, 5 - 8 January, 2014.

- Ana F. Sequeira, Juliano Murari and Jaime S. Cardoso, Iris Liveness Detection Methods in Mobile Applications, In Proceedings of the 9^{th} International Conference on Computer Vision Theory and Applications (VISAPP), pp.22-33, Lisboa, Portugal, 5 - 8 January, 2014.

- João C. Monteiro, Hélder P. Oliveira, Ana F. Sequeira and Jaime S. Cardoso, Robust Iris Segmentation under Unconstrained Settings, In Proceedings of the 8^{th} International Conference on Computer Vision Theory and Applications (VISAPP), pp. 180 - 190, Barcelona, Spain, 21 - 24 February, 2013.

We also participated in national conferences with the following papers:

- Ana F. Sequeira, Ana Rute Louro, Jaime S. Cardoso, Liveness detection methods in fingerprint recognition, In Proceedings of the 20^{th} Portuguese Conference on Pattern Recognition (RECPAD2014), Covilhã, Portugal, 31 October, 2014.

- Ana F. Sequeira, Juliano Murari, Jaime S. Cardoso, Quality measures for iris images in mobile applications, In Proceedings of the 19^{th} Portuguese Conference on Pattern Recognition (RECPAD2013), Lisbon, Portugal, 31 October, 2013.

- Juliano Murari, Ana F. Sequeira and Jaime S. Cardoso, Liveness detection in iris recognition systems, In Proceedings of the 2^{nd} PhD. Students Conference in Electrical and Computer Engineering, Porto, Portugal, 25 - 27 June, 2013.

- Ana F. Sequeira, Samaneh Khoshrou, Jaime S. Cardoso, Color feature selection for unconstrained iris recognition, In Proceedings of the 18^{th} Portuguese Conference on Pattern Recognition (RECPAD2012), Coimbra, Portugal, 26 October, 2012.

- Ana F. Sequeira, Jaime S. Cardoso, Samaneh Khoshrou, Color feature selection for an unconstrained iris recognition system, In Proceedings of the 1^{st} PhD. Students Conference in Electrical and Computer Engineering, pp. 27 - 28, Porto, Portugal, 28 - 29 June, 2012.

1.4 Document Structure

The first sections of this Chapter present the main motivations, objective and thesis contributions. In this section is presented the structure of this document. This thesis is divided into four parts comprising ten chapters in total, describing the work conducted during the last four years.

In the following figure we framed the work developed in this thesis in several areas of the biometrics scenario.

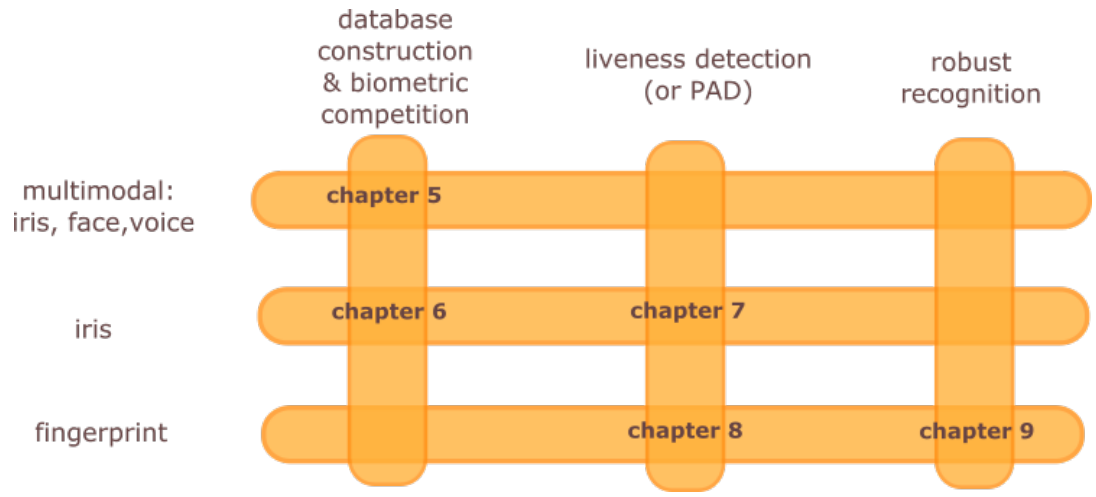


Figure 1.1: Work developed in the thesis across different areas of the biometric scenario.

The first part, **Introduction and State of the Art**, is composed by the present chapter; the chapter 2 where a brief description of the two main concepts of this work, *Biometrics* and *Liveness Detection*, are presented as well as some background concepts of *Computer vision* and *Machine learning* that were used throughout the work developed in this thesis; the chapter 3 where the literature review about iris is presented; and finally the chapter 4 where the literature review about fingerprint is presented.

The second part, **Biometric Databases and Biometric Competitions**, is dedicated to the biometric databases constructed and the biometric competitions organized in the scope of this work and is composed by chapters 5 and 6.

The third part, **Liveness Detection and Robust Recognition**, comprises the work regarding liveness detection and recognition. Chapter 7 presents the work developed regarding iris liveness detection; chapter 8 presents the work done in the fingerprint liveness detection problem; and finally chapter 9 where the work in robust fingerprint binarization is presented.

The fourth part, **Conclusions and Future work**, comprises chapter 10 with the conclusions and future work.

Annex 1 presents the work on iris segmentation which I co-authored.

Chapter 2

Background Knowledge

In this chapter are presented some basic concepts related to biometrics, in general; presentation attack detection or liveness detection techniques; and some computer vision and machine learning background knowledge.

2.1 Biometrics

In almost everyone's daily activities, personal identification plays an important role. The most traditional techniques to achieve this goal are knowledge-based and token-based automatic personal identifications. Token-based approaches take advantage of a personal item, such as a passport, driver's license, ID card, credit card or a simple set of keys to distinguish between individuals. Knowledge-based approaches, on the other hand, are based on something the user knows that, theoretically, nobody else has access to, for example passwords or personal identification numbers [203]. Both of these approaches present obvious disadvantages: tokens may be lost, stolen, forgotten or misplaced, while passwords can easily be forgotten by a valid user or guessed by an unauthorized one. In fact, all of these approaches stumble upon an obvious problem: any piece of material or knowledge can be fraudulently acquired [112].

Biometrics represents a return to a more natural way of identification as many physiological or behavioural characteristics are unique between different persons. Testing someone by what this someone is, instead of relying on something he owns or knows seems likely to be the way forward [174].

The etymology of the word "biometrics" comes from the Greek "bio-" which refers to one's life, course or way of living, and "-metros" which means measure [201]. Even though this term has been previously used in the field of statistics to refer to the analysis of biological data (which is now known as biostatistics), it is generally used to refer to an automated method of authenticating individuals based on anatomical or behavioral human characteristics [18, 249].

Architecture of a Biometric System

A biometric system can be divided in two stages: the enrollment and the identification/verification. The enrollment consists in acquiring data from specific individuals so that a database can be built. It can be said that the enrollment is the registration of individuals to the database and those will be the ones who should be recognized during the identification or verification process. The second stage of a biometric system is the identification which, no matter what feature is chosen to work with, follows the process illustrated in Figure 2.1. It can be split in five modules.

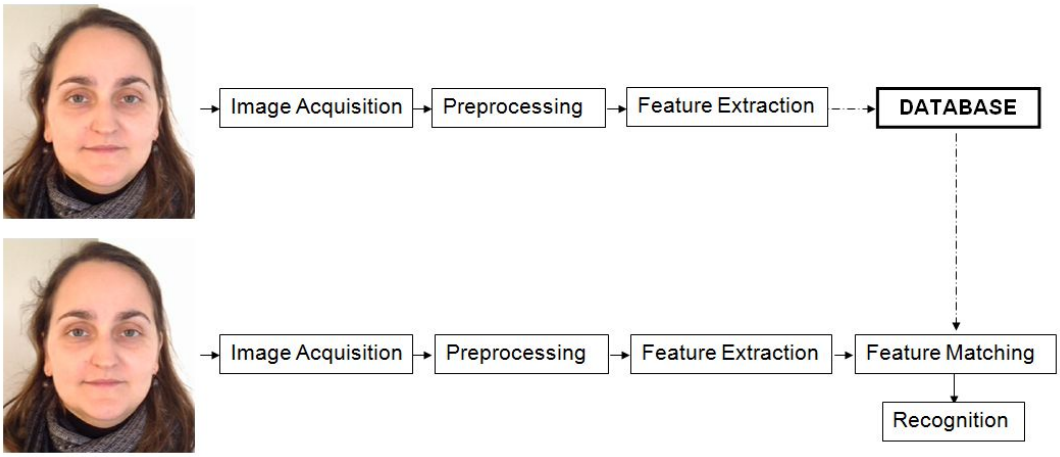


Figure 2.1: Scheme of a typical Biometric System.

The process starts with the capture of the images, acquiring a biometric sample, followed by a preprocessing module where several steps can be taken like segmentation, normalization and liveness detection. The third part of the process consists in the feature extraction, where conspicuous features are identified for classification. Lastly, the features are compared and matched with stored ones, resulting in a possible recognition [172, 216].

Operating Mode of a Biometric System

A biometric system can be used for both identification and verification. The verification system aims to confirm or deny a person's identity by comparing his/her biometric template with the stored ones, evaluating the probability of this person being who he/she claims to be. For this purpose a 1:1 matching algorithm is used. In an identification mode, the individual's biometric feature is compared to the entire database, which means that a 1:N matching algorithm is used. The objective of this process is to detect the person's ID. The success of both these operating modes depends on the presupposition that the person using the system has already been through the enrollment process. In addition, there are several ways of spoofing these systems, but this topic will be discussed hereafter.

Biometric Traits

Anatomical traits normally refer to a part of the human body like the face, the eyes, the hands, among others. On the other hand, behavioral traits relate to the conduct of a person, such as the voice or the signature. Biometric traits can also be labeled as genotypic or phenotypic. Genotypic features are genetically defined, while phenotypic features can be changed over time and depend on the surrounding environment. Several biological traits in humans show a considerable inter-individual variability: DNA, ear shape, face, facial thermogram, hand thermogram, hand veins, fingerprints, gait, hand geometry, iris, palmprints, retina, keystroke and voice, as depicted on Figure 2.2. Biometrics works by recognizing patterns within these biological traits, unique to each individual, to increase the reliability of recognition.

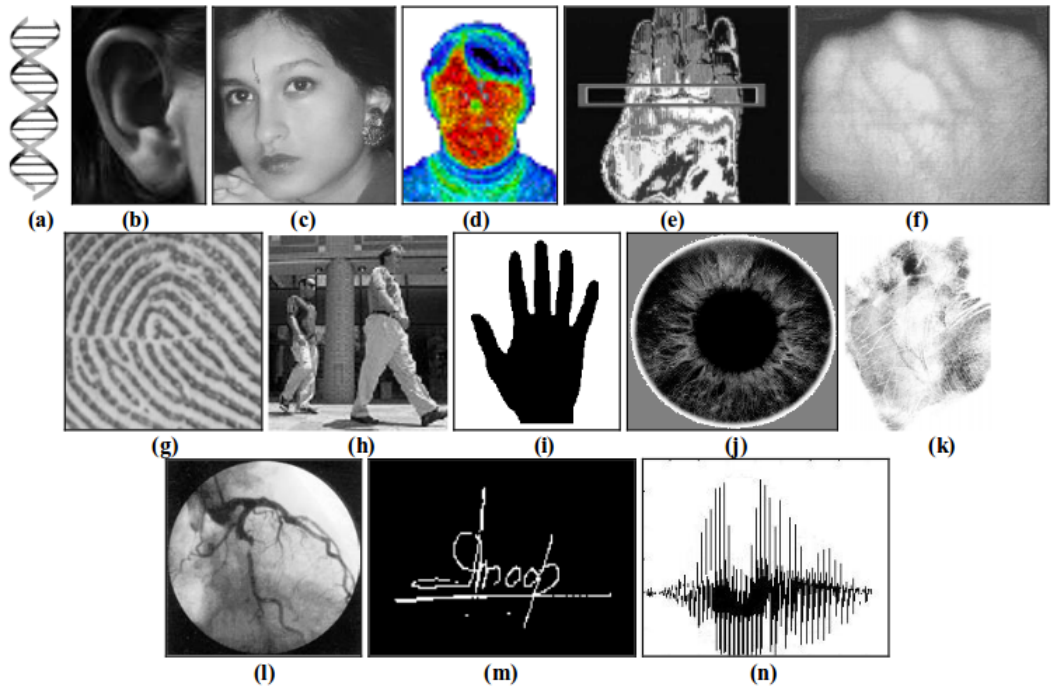


Figure 2.2: Examples of some of the most widely studied biometric traits: (a) DNA, (b) Ear shape, (c) Face, (d) Facial Thermogram, (e) Hand Thermogram, (f) Hand veins, (g) Fingerprint, (h) Gait, (i) Hand geometry, (j) Iris, (k) Palmprint, (l) Retina, (m)Signature and (n) Voice (from [110])

.

The growing need for reliability and robustness, raises some expectations and became the focal point of attention for research works on biometrics. When developing a new system based on a specific trait this is expected to meet some conditions. An ideal biometric should be *universal*, where each person possesses the characteristic; *unique*, where no two persons should share the characteristic; *collectable* where the characteristic is readily presentable to a sensor and is easily quantifiable; and *permanent*, where the characteristic should neither change nor be alterable [112]. In practice, however, a characteristic that satisfies all these requirements may not always be feasible for a useful biometric system. The designer of a practical biometric system must also consider

a number of other issues, including: *performance*, that is, a system's accuracy, speed, robustness, as well as its resource requirements, and operational or environmental factors that affect its accuracy and speed; *acceptability*, or the extent people are willing to accept for a particular biometric identifier in their daily lives; *circumvention*, as in how easy it is to fool the system through fraudulent methods [112]. In Table 2.1 it is shown a classification of the previously referred biometric traits concerning the requirements mentioned.

Table 2.1: Comparative data analysis of some common biometric traits (adapted from [112] and [208]).

	Requirements			
Traits	**Universality**	**Uniqueness**	**Collectability**	**Permanence**
DNA	High	High	Low	High
Ear	Medium	Medium	Medium	High
Face	High	Low	High	Medium
Facial Thermogram	High	High	High	Low
Hand Veins	Medium	High	High	Medium
Fingerprint	Medium	High	High	Medium
Gait	Low	Low	High	Low
Hand Geometry	Medium	Medium	High	Medium
Iris	High	High	Medium	High
Palm Print	Medium	High	Medium	High
Retina	High	High	Low	Medium
Signature	Medium	Low	High	Low
Voice	Medium	Low	Medium	Low

Applications of biometric recognition

It is possible to identify some of the main biometric applications in the forensic, civilian and commercial fields, shown in Table2.2. In forensic applications we find fingerprint as one of the most important biometric trait, nevertheless face has been gaining importance and the construction of enormous databases is a recent phenomena. Regarding the referred government/civilian applications we find fingerprint along with face and iris/periocular as the more used. These appear in commercial applications as well as voice.

Table 2.2: Biometric applications (adapted from [112]).

Forensic	**Government / Civilian**	**Commercial**
Criminal investigation	National ID	ATM
Corpse identification	Driver's licence	Cellular phone
Missing Children	Border crossing	Access control

2.2 Vulnerabilities of biometric recognition systems

Biometric systems can offer several advantages over classical security methods as they rather rely on intrinsic characteristics of an individual instead of relying on some of the individual's knowledge or possessions. However, in spite of their advantages, biometric systems have some drawbacks, including: i) the lack of secrecy (e.g. everybody knows our face or could get our fingerprints), and ii) the fact that a biometric trait cannot be replaced (for example, no new iris can be generated if an impostor "steals" it). Furthermore, biometric systems are vulnerable to external attacks which could decrease their level of security. Concerning these vulnerabilities we find in the literature [82] an analysis of the eight different points of attack on biometric recognition systems previously identified by Ratha et al. [218]. These points are illustrated in Fig. 2.3.

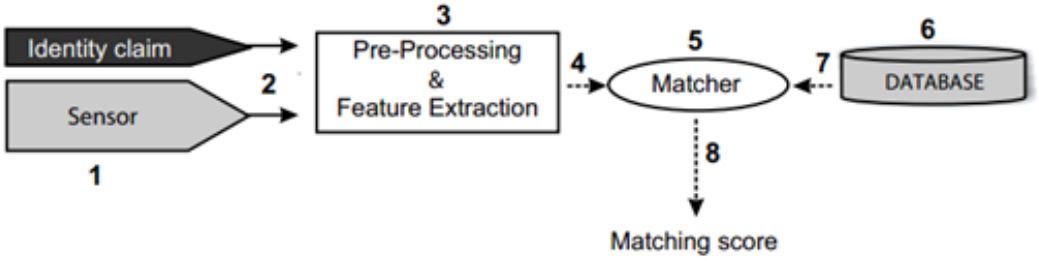


Figure 2.3: Architecture of an automated biometric verification system. Possible attack points are numbered from 1 to 8, from [82]

The attacks to biometric recognition systems can be divided into two main groups: direct and indirect attacks.

- *Direct attacks:* the first vulnerability point in a biometric security system is the possibility to generate synthetic biometric samples (for instance, speech, fingerprints or face images) in order to fraudulently access a system. These attacks at the sensor level are referred to as direct attacks. It is worth noting that in this type of attacks no specific knowledge about the system operation is needed (matching algorithm used, feature extraction, feature vector format, etc). Furthermore, the attack is carried out in the analogue domain, outside the digital limits of the system, so the digital protection mechanisms (digital signature, watermarking, etc.) cannot be used.

- *Indirect attacks:* this group includes all the remaining seven points of attack. Attacks 3 and 5 might be carried out using a Trojan Horse that bypasses the feature extractor, and the matcher respectively. In attack 6 the system database is manipulated (a template is changed, added or deleted) in order to gain access to the application. The remaining points of attack (2, 4, 7 and 8) are thought to exploit possible weak points in the communication channels of the system, extracting, adding or changing information from them. In opposition to direct attacks, in this case the intruder needs to have some information about the inner working of the recognition system and, in most cases, physical access to some of the application components (feature extractor, matcher, database, etc.) is required.

The direct attacks or spoofing attacks have been recently started to be denominated as “presentation attacks” . These can be perpetrated by presenting an altered or fake biometric sample of an authorized user to the sensor in order to obtain access.

2.3 Presentation Attack Detection

Methods to determine if a biometric sample is altered or fake constitute the denominated “presentation attack detection methods”. In particular, the liveness detection methods are applied in order to determine whether a sample is fake or real.

The problem of liveness detection of a biometric trait can be seen as a two class classification problem where an input trait sample has to be assigned to one of two classes: real or fake. The key point of the process is to find a set of discriminant features which permits to build an appropriate classifier to predict the probability of the sample vitality given the extracted set of features [83].

In the literature we found that the methods of liveness detection may be classified into four categories based on the physical features of biometric and liveness data and the timing of measurement [263]. In this framework, “biometric data” is the one used in the recognition and the “liveness data” is the one used in the liveness detection. We can itemize the four categories:

- Perfect matching model: Both biometric and liveness data are simultaneously obtained from the same physical feature.
- Simultaneous measuring model: Biometric and liveness data are simultaneously obtained from different physical features.
- Same biometric measuring model: Biometric and liveness data are obtained from the same physical feature with different timings.
- Independent measuring model: Biometric and liveness data are obtained from different features with different timings.

The ideal configuration of liveness detection for biometrics recognition is represented by the perfect matching model with the highest ability to distinguish between live and fake irises [127].

Liveness detection methods can be categorized as hardware or software-based whether the detection is performed through additional hardware or by processing the obtained image [45]. Hardware-based solutions work by measuring some physical characteristics (such as blood pressure, temperature, pulse, or pupil dilation, voluntary eye blink, among others) and have the disadvantage of being expensive and requiring intervention at the device level. Memon et al. [169] make a detailed classification of the several methods of hardware based fingerprint liveness detection methods dividing them in three categories: *biological or involuntary signals*, *skin physiology characteristics* and *stimulation response*. Software-based liveness detection methods can be divided in two categories based on whether they work with a single static 2D scan (static methods),

or need 2D scans at different time points during the acquisition process that support the observation of the dynamic properties (dynamic methods) [251]. Different sub-categorization can be mada accordingly to the biometric trait which we are studying.

2.4 Accuracy and Evaluation Measurements

2.4.1 Performance Evaluation of a Biometric System

Even though it is not possible to give a single value that reflects the accuracy of a recognition system, there are some measures of accuracy that can be used under the same data and following the same protocol that can be helpful when trying to evaluate the effectiveness of a biometric system.

One way of comparing one system to another, both using the same biometric trait, is by analyzing its False Acceptance Rate (FAR) [208]. The FAR is the proportion between the number of False Acceptances (FA), i.e. the number of impostors that were able to enter the system, and the total number of impostors that try to access the system: the falsely accepted and the Truly Rejections (TR):

$$FAR = \frac{FA}{FA + TR} \tag{2.1}$$

The FAR measures, then, the probability of confusing two identities, but it is only meaningful when presented simultaneously with the False Rejection Rate (FRR).

The FRR is the probability of the identity of a valid user being denied and it can be calculated as the proportion between the False Rejections (FR) and the total number of users that try to access the system, the False Rejections and the Truly Acceptances (TA).

$$FRR = \frac{FR}{FR + TA} \tag{2.2}$$

Figure 2.4 shows a graphical representation of FAR and FRR values for distinct similarity threshold values. The point where the two lines intersect represents the Equal Error Rate (EER) and is a very common measure of the biometric systems accuracy. It gives the average error rate when the FAR and FRR are approximately the same.

Another metric used is the False Identification Rate (FIR) corresponding to the ratio between the number of wrong recognition attempts and the total number of recognition attempts, is given by:

$$FIR = \frac{FA + FR}{(FA + TA + FR + TR)}. \tag{2.3}$$

2.4.2 Performance Evaluation of Presentation Attack Techniques

The “Normal Presentation Classification Error Rate” (NPCER) is given by the proportion of normal (live) presentations incorrectly classified as attack (non-live) presentations and the “Attack

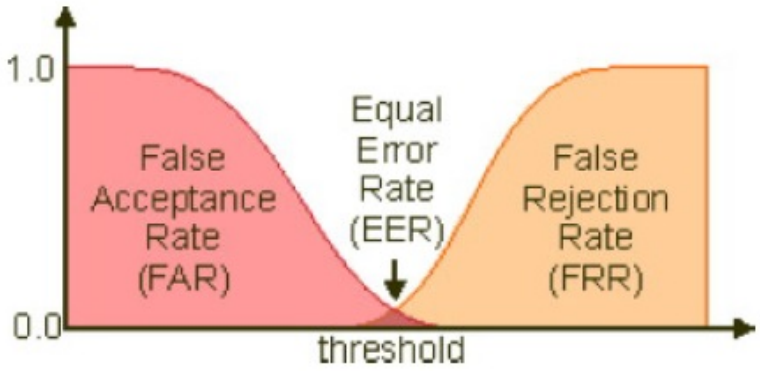


Figure 2.4: Representation of FAR and FRR [172].

Presentation Classification Error Rate" (APCER) is given by the proportion of attack (non-live) presentations incorrectly classified as normal (live) presentations. The *Average Classification Error Rate (ACER)* is given by the mean of the NPCER and the APCER error rates. The ACER is the metric used to evaluate the results and corresponds to the misclassification rate of the classifier.

$$ACER = \frac{(NPCER + APCER)}{2} \tag{2.4}$$

These metrics are suggested by the standardization project, *ISO/IEC 30107 Presentation Attack Detection* [108], which focuses on techniques for the automated detection of presentation attacks undertaken by data capture subjects at the point of presentation and collection of the relevant biometric characteristics.

We can assess the performance of a liveness detection method in terms of its error rate or its accuracy [160]. As observed before, two types of error may occur while performing a liveness detection action: a fake image can be labeled as a real one or a real image can be considered as fake.

The implications of these errors in terms of security are weighted differently. Usually, if a real image is classified as a fake one and the access of an authorized person is denied, that can be inconvenient but does not introduce insecurity to the process, whereas if a fake sample is considered as a real one, a potential impostor could have access to whatever was protected by the system, compromising its security. However, there could be also systems where the compromise between the two different types of errors is balanced in such way that the denial of access to real users is more inconvenient than the authorization of non-registered users.
In a general way, the error rate is calculated by finding the ratio between the number of misclassified images, that is, the sum of real images considered as fakes and fake images accepted as real, and the total number of images in the testing set, given by Equation 2.5. The accuracy is the percentage of correctly classified images, therefore given by Equation 2.6. Using this formula, the type of error is not distinguished, which means that the number of fake accepted images or real

rejected images is not being taken into account, as it only uses the total number of classification errors.

$$\text{Error Rate} = \frac{\text{misclassified images}}{\text{total number of images}} \times 100(\%) \tag{2.5}$$

$$\text{Accuracy} = 100\% - \text{Error Rate} \tag{2.6}$$

2.5 Computer Vision Background Knowledge

2.5.1 Gray Level Co-Occurrence Matrices

The Gray Level Co-Occurrence Matrices (GLCM) [97] analysis is one of the most prominent methodologies in textural analysis. Co-occurrence matrices characterize the relationship between neighboring pixels (Figure 2.5). Each element $p(i, j)$ of a GLCM matrix represents the relative frequency with which two neighboring pixels separated by a certain distance occur, one with a gray scale i and another with a gray scale j.

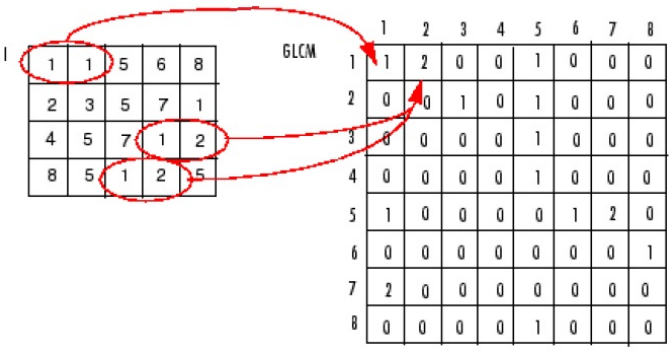


Figure 2.5: Example of the creation of a GLCM matrix [176].

Haralick et al. [97] defined 14 features that can be extracted from a GLCM matrix:
1) Angular Second Moment (ASM):

$$f_1 = \sum_i \sum_j \{p(i, j)\}^2 \tag{2.7}$$

where $p(i, j)$ is the (i, j)th entry in a normalized gray-tone spacial-dependence matrix.
2) Contrast:

$$f_2 = \sum_{n=0}^{N_{g-1}} n^2 \{ \sum_{\substack{i=1 \\ |i-j|=n}}^{N_g} \sum_{j=1}^{N_g} p(i, j) \} \tag{2.8}$$

where N_g is the number of distinct gray levels in the quantized image.

3) Correlation:

$$f_3 = \frac{\sum_i \sum_j (i,j) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{2.9}$$

4) Sum of Squares (Variance):

$$f_4 = \sum_i \sum_j (i-\mu)^2 p(i,j) \tag{2.10}$$

5) Inverse Difference Moment:

$$f_5 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j) \tag{2.11}$$

6) Sum Average:

$$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i) \tag{2.12}$$

7) Sum Variance:

$$f_7 = \sum_{i=2}^{2N_g} (i-f_8)^2 p_{x+y}(i) \tag{2.13}$$

8) Sum Entropy:

$$f_8 = -\sum_{i=2}^{2N_g} p_{x+y}(i) log\{p_{x+y}(i)\} \tag{2.14}$$

9) Entropy:

$$f_9 = -\sum_i \sum_j p(i,j) log(p(i,j)) \tag{2.15}$$

10) Difference Variance:

$$f_{10} = \text{variance of } p_{x-y} \tag{2.16}$$

11) Difference Entropy:

$$f_{11} = -\sum_{i=0}^{N_{g-1}} p_{x-y}(i) log\{p_{x-y}(i)\} \tag{2.17}$$

12) & 13) Information Measures of Correlation:

$$f_{12} = \frac{HXY - HXY1}{max(HX, HY)} \tag{2.18}$$

$$f_{13} = (1 - exp[-2.0(HXY2 - HXY)])^{\frac{1}{2}} \tag{2.19}$$

$$HXY = -\sum_i \sum_j p(i,j) log(p(i,j)) \tag{2.20}$$

where HX and HY are entropies of p_x and p_y, and

$$HXY1 = -\sum_i \sum_j p(i,j) log\{p_x(i)p_y(j)\} \tag{2.21}$$

where $p_x(i)$ and $p_y(j)$ are the ith entry in the marginal-probability matrix, obtained by summing the rows/columns of $p(i,j)$.

$$HXY2 = -\sum_i \sum_j p_x(i)p_y(j) log\{p_x(i)p_y(j)\} \tag{2.22}$$

14) Maximal Correlation Coefficient:

$$f_{14} = (\text{second largest eigenvalue of } Q)^{\frac{1}{2}} \tag{2.23}$$

where

$$Q(i,j) = \sum_k \frac{p(i,k)p(j,k)}{p_x(i)p_y(k)} \tag{2.24}$$

These features are orientation dependent so four values can be obtained for each feature based on the four orientations $\{0°, 45°, 90°, 135°\}$. The mean and standard deviation of the four values (one for each of the four orientations) of each of the 14 measures, compose a set of 28 features.

2.5.2 Local Binary Pattern

Local Binary Pattern (LBP) [199] encodes the information between the target pixel and the neighboring pixel intensities: for each target pixel, x, it takes P neighbors sampled uniformly on a circle of radius R centered on x. These pixels are then compared with x, taking only the sign of the difference, and forming thus a vector of P values, which are then converted in a decimal number. The formula to determine the value for pixel x is given by:

$$LBP = \sum_{i=0}^{P-1} u(x_i - x)2^i \tag{2.25}$$

where x_i is the i-th neighbor of pixel x, $u(x) = 1$ when $x \geq 0$ and 0 otherwise.

2.5.3 Scale-Invariant Feature Transform

The Scale Invariant Feature Transformation (SIFT) was introduced by Lowe [149] and is currently used in many computer vision applications. SIFT encoding is a procedure that enables the extrac-

tion of local information from digital images. This technique consists in normalizing local patches around robust scale covariant image key points. This is a methodology which transforms an image into a large collection of local feature vectors, each one is invariant to image translation, scaling, and rotation, partially invariant to illumination changes, noise, minor changes in viewpoint and is robust to local geometric distortion (scale and rotation). These features are defined as maxima and minima of the result of the aforementioned DoG applied in scale space to a series of smoothed and resampled images (see Figure 2.6). With Lowe's method, it is possible to extract features which are highly distinctive and relatively easy to extract, and to identify objects correctly with low probability of mismatch. They are relatively easy to match against a database of local features - however, the high dimensionality can be an issue, and generally probabilistic algorithms are used.

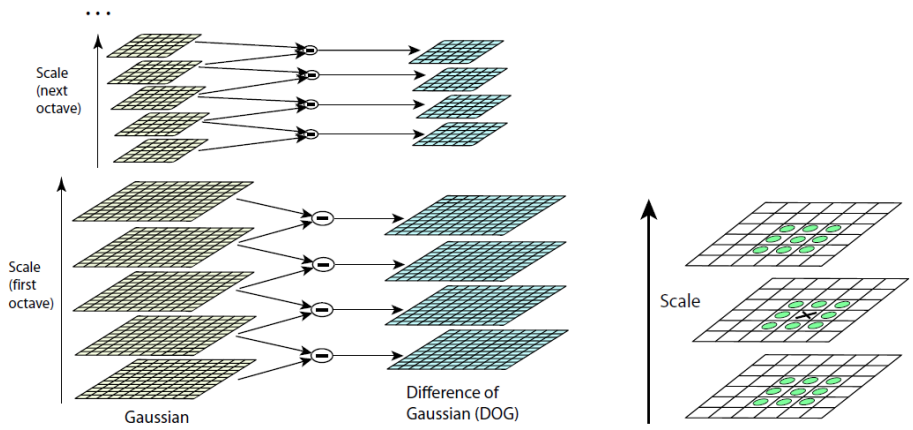


Figure 2.6: Difference of Gaussian space-scale pyramid visual description (left) with extremal pixel finding by comparison to the neighbours at each level (right) (from [149]).

The SIFT algorithm includes a robust descriptor. Considering a window of 8×8 around the feature location, using the level of the Gaussian filter in which the feature was detected, then the descriptor is based on computing the gradient in each of the pixels contained in that window, downgraded by a Gaussian fall-off function (to reduce the effect of those pixels far from the feature center). Then, in each quadrant of that window, it develops an orientation histogram by adding the value of each pixel to one of the eight orientation bins using tri-linear interpolation in a $2 \times 2 \times 2$ histogram (4 eight-bin histogram for each quadrant). That forms a 128-dimension feature vector. An example is depicted in Figure 2.7.

2.6 Machine Learning Background Knowledge

Different approaches were evaluated in this work for the classification problem: linear discriminant analysis (LDA), k-nearest neighbor (kNN), support vector machine (SVM), one class SVM (OCSVM) and gaussian mixture models (GMM). For optimizing the parameters we used nested

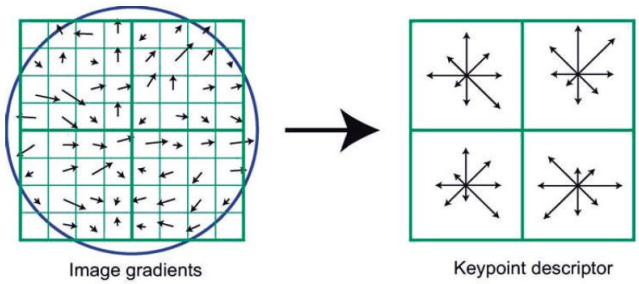


Figure 2.7: Example of a 2×2 descriptor array of orientation histograms (right) computed from an 8×8 set of samples (left). The orientation histograms are quantized into 8 directions and the length of each arrow corresponds to the magnitude of the histogram entry (from [149]).

cross validation. Also the Principal Component Analysis (PCA) feature dimensionality reduction technique was applied. These concepts are detailed in this section.

2.6.1 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) [75, 79] is a technique used when the classes are known *a priori*. Therefore, given π_k observations, $k = 1, \ldots, K$, and one observation to allocate to one of the existing classes the discriminant rule will separate the space from the sample in sets R_k and classify the sample x as belonging to π_k if the sample belongs to R_k. The main purpose is to find the optimal regions R_j which minimize the classification error.

In a more general way, given a data matrix $A \in \mathbb{R}^{N \times n}$ LDA aims to find a transformation $G \in \mathbb{R}^{N \times l}$ that maps each column a_i of A for $1 \leq i \leq n$, in the $N-$dimensional space to a vector b_i in the $l-$dimensional space. That is $G : a_i \in \mathbb{R}^N \rightarrow b_i = G^T a_i \in \mathbb{R}^l \ (l < N)$. Equivalently, LDA aims to find a vector space $\mathscr{G}$ spanned by $\{g_i\}_{i=1}^{l}$, where $G = [g_1, ..., g_l]$, such that each a_i is projected onto $\mathscr{G}$ by $(g_1^T \cdot a_i, ..., g_l^T \cdot a_i)^T \in \mathbb{R}^l$.

Assume that the original data in A is partitioned into k classes as $A = \{\Pi_i, ..., \Pi_k\}$ where Π contains n_i data points from the ith class, and $\sum_{i=1}^{k} n_i = n$. LDA aims to find the optimal transformation G such that the class structure of the original high-dimensional space is preserved in the low-dimensional space.

2.6.2 K Nearest Neighbours

K-nearest neighbour (kNN) [79] is a supervised learning method for classifying objects based on training examples in the feature space. This algorithm belongs to a set of techniques called Instance-based Learning. The K-nearest neighbor algorithm is very simple and basically it does not have train. It starts by extending the local region around a data point x until the k^{th} nearest neighbor is found. The most represented class in the k-closest samples defines the predicted class. Training lies only in the estimation of the best k. In this work the Euclidean distance was used :

$$d_{euclidean}(a,b) = \sqrt{\sum_{i=1}^{d} (a_i - b_i)^2}$$

2.6.3 Support Vector Machines

Support Vector Machines (SVMs), pioneered by Vapnik [269], deal with the problem of classification as a problem of quadratic optimization. In the base of this technique is the idea of the construction of a hyperplane as the decision surface in such a way that the margin of separation between positive and negative examples is maximized. Support vector machines classify the data using support vectors [98].

SVMs work by maximizing the margin of the optimal hyperplane which separates the data. This is typically done in a much higher dimension than that of the original feature space. Formally, given the training set $\{x_i, y_i\}_{i=1}^{N}$ with input data $x_i \in \mathbf{R}^P$ and the corresponding binary class labels $d_i \in \{-1, 1\}$, the linear separable optimal hyperplane is defined by $g(\mathbf{x}) = \mathbf{w}^t \varphi(\mathbf{x}) + b$ where $\varphi(\mathbf{x})$ denotes a fixed-feature space transformation and b a bias parameter. An observation $\mathbf{x}$ is assigned to class 1 if $g(\mathbf{x}) > 0$ or to -1 if $g(\mathbf{x}) < 0$. This is equivalent to have $d_i(\mathbf{w}^t \varphi(\mathbf{x}) + b) \geq 1$, $i = 1, \ldots, N$. In fact, maximizing the margin is equivalent to solving

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2}\mathbf{w}^T\mathbf{w} \\ s.t \quad & d_i(\mathbf{w}^T \varphi(\mathbf{x}) + b) \geq 1, i = 1, \ldots, N \end{aligned} \tag{2.26}$$

However, if the training classes are not linearly separable, the above conditions and problem formulation can not be sustained. For this reason, slack variables ξ_i, $i = 1, \ldots, N$ are added. These allow to have a penalty for the data points wrongly classified. Finally, the objective is to minimize the error. That is,

$$\begin{aligned} \min_{w,b,C,\xi_i} \quad & \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{N}\xi_i \\ s.t \quad & d_i(\mathbf{w}^T \varphi(\mathbf{x}) + b) \geq 1 - \xi_i, i = 1, \ldots, N \\ & \xi_i \geq 0 \end{aligned} \tag{2.27}$$

where the parameter $C > 0$ controls the trade-off between the training error and the margin.

In the feature space it is easier to solve the *dual problem*, and sometimes it is the only way to train the support vector machines. It is possible to formulate the dual problem for a sample of training $\{x_i, y_i\}_{i=1}^{N}$ not separable as follows:

$$\begin{aligned} \max_{\alpha} \quad & \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j d_i d_j k(x_i, x_j) \\ s.t \quad & \textstyle\sum_{i=1}^{N}\alpha_i d_i = 0 \\ & 0 \leq \alpha_i \leq C \quad i = 1, 2, ..., N \end{aligned} \tag{2.28}$$

where $k(x_i, x_j) = \varphi^T(x_i)\varphi(x_j) = \sum_{l=0}^{m_1}\varphi_l(x_i)\varphi_l(x_j)$, $i = 1, 2, ..., N$ and $j = 1, 2, ..., N$. $\varphi_l(x_i)$ is the l component in the application $\varphi(x_i)$ of x_i; m_1 is the dimension of the feature space.

The three most common types of inner-product kernels for SVMs are: polynomial learning machine, radial-basis function network and tangent hyperbolic. In this work a radial-basis function network was used, given by:

$$k(\mathbf{x},\mathbf{x}_i) = exp(-\gamma||\mathbf{x}-\mathbf{x}_i||^2), \gamma \geq 0 \tag{2.29}$$

2.6.4 One Class SVM

One class support vector machine (OCSVM) [230, 261] was proposed by Scholkopf *et al.* for estimating the support of a high dimensional distribution. Given training vectors $x_i \in \mathbb{R}^n, i = 1, ..., l$ without any class information, the primal problem of one-class SVM is

$$\begin{aligned} \min_{w,\xi,\rho} \quad & \frac{1}{2}\mathbf{w}^T\mathbf{w} - \rho + \frac{1}{vl}\sum_{i=0}^{l}\xi_i \\ s.t \quad & d_i[\mathbf{w}^T\varphi(\mathbf{x_i})] \geq \rho - \xi_i, \xi_i = 1, \ldots, l \end{aligned} \tag{2.30}$$

The dual problem is then

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2}\alpha^T Q_\alpha \\ s.t \quad & 0 \leq \alpha_i \leq (1/vl), \quad i = 1, ..., l \\ & e^T\alpha = 1, \end{aligned} \tag{2.31}$$

where $Q_{ij} = K(x_i, x_j) = \varphi^T(x_i)\varphi(x_j)$. The decision function is

$sgn\left(\sum_{i=1}^{l}\alpha_i K(x_i, x) - \rho\right)$.

In LIBSVM, which is the library used in this work, it is solved a scaled version of the dual problem:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2}\alpha^T Q_\alpha \\ s.t \quad & 0 \leq \alpha_i \leq 1, \quad i = 1, ..., l \\ & e^T\alpha = vl, \end{aligned} \tag{2.32}$$

2.6.5 Gaussian Mixture Models

A Gaussian Mixture Model (GMM) [109, 220] is a parametric probability density function represented as a weighted sum of Gaussian component densities. GMMs are commonly used as a parametric model of the probability distribution of continuous measurements or features in a biometric system.

A Gaussian mixture model is a weighted sum of M component Gaussian densities as given by the equation:

$$p(x|\lambda) = \sum_{i=1}^{M} w_i g(x|\mu_i, \Sigma_i), \tag{2.33}$$

where x is a $D-$dimensional continuous-valued data vector (i.e. measurement or features), $w_i, i = 1,...,M$, are the mixture weights, and $g(x|\mu_i, \Sigma_i)$, $i = 1,...,M$, are the component Gaussian densities. Each component density is a $D-$variate Gaussian function of the form:

$$g(x|\mu_i, \Sigma_i) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} exp\left\{-\frac{1}{2}(x-\mu_i)'\Sigma_i^{-1}(x-\mu_i)\right\}, \tag{2.34}$$

with mean vector μ_i and covariance matrix Σ_i. The mixture weights satisfy the constraint that $\sum_{i=1}^{M} w_i = 1$. The complete Gaussian mixture model is parameterized by the mean vectors, covariance matrices and mixture weight from all component densities. These parameters are collectively represented by the notation

$$\lambda = \{w_i, \mu_i, \Sigma_i)\}, i = 1,...,M. \tag{2.35}$$

Typically, GMM parameters are estimated from training data using the iterative Expectation-Maximization (EM) algorithm.

2.6.6 Principal Component Analysis

Principal component analysis (PCA) [125] is a well-established technique for dimension reduction. The most common definition of PCA is a standardized linear projection which maximizes the variance of the data in the projected space. For a set of observed $d-$dimensional data vectors t_n, $n \in 1,...,N$ the *q principal axes* w_j, $j \in 1,...,q$, are those orthonormal axes onto which the retained variance under projection is maximal. It can be shown that the vectors w_j are given by the q dominant eigenvectors (i.e., those with the largest associated eigenvalues λ_j) of the sample covariance matrix $S = E\left[(t-\nu)(t-\nu)^T\right]$ such that $S_{w_j} = \lambda_j w_j$. The q principal components of the observed vector t_n are given by the vector $x_n = W^T(t_n - \nu)$, where $W^T = (w_1, w_2, ..., w_q)^T$. The variables x_j are then correlated such that the covariance matrix $E\left[xx^T\right]$ is diagonal with elements λ_j. A complementary property of PCA, is that of all orthogonal linear projections $x_n = W^T(t_n - \nu)$, the principal component projection minimizes the squared reconstruction error $\sum \|t_n - \overset{t}{}_n\|^2$, where the optimal linear reconstruction of t_n is given by $\overset{t}{}_n = W_{x_n} + \nu$.

2.6.7 Nested Validation Procedure

Often to optimize the parameters to be used in classifiers, such as the number of neighbors k in kNN or the several parameters used in SVMs, can be performed a nested validation procedure. Exemplifying with a nested hold out, the initial data is divided in two sets *Train* and *Test*, and then the *Train* set is divided in two new sets, *Train1* and *Test1*, as depicted in Figure2.8 (according to

a defined proportion). The optimization of parameters is then performed by a grid-search while training in *Train1* and evaluating in *Test1*. Then the best parameters found, evaluated by the misclassification rate, are used to train the model in *Train* set and then the model is evaluated in *Test* set. This process can be repeated a defined number of times and the final error rate will be the averaged misclassification rate obtained on the evaluation in the *Test* set.

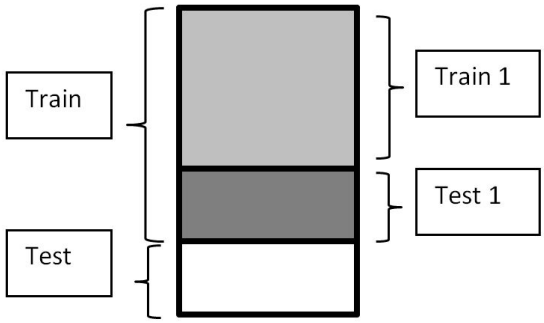


Figure 2.8: Division of the data in training and test sets for the nested cross-validation.

Replacing the hold out by a cross validation scheme, one obtains a nested cross validation, which was selected in this work (since it uses more efficiently the data).

Chapter 3

Iris Liveness Detection and Recognition Literature Review

In this chapter we contextualize the iris recognition problem and present its origins and the most relevant works in the unconstrained scenario. We discuss some questions posed by the new scenarios in the practical use of this biometric trait. One of the aspects to be taken in account is its security and its vulnerability to attacks. In particular we analyze the presentation attack problem and present some literature review in Iris "Presentation Attack" or "Liveness" Detection Methods. Also, the most important databases for iris recognition and iris liveness detection are described.

3.1 Eye Anatomy

The eye is a globular and hollow part of the human body composed by three layers: the internal, external and middle layer [172]. The external layer can also be called fibrous tunic and is constituted by the sclera and the cornea. The middle layer, or uvea/vascular tunic, has the iris and the cilliar body in it; and the internal layer, called nervous tunic, is composed by the retina. Figure 3.1 illustrates the composition of an human eye.

When observing a human eye, as depicted in Figure 3.2, in a non-invasive way, three of its features can be seen: the sclera, the pupil and the iris. The sclera, commonly known as the white area of the eye, is a tough and fibrous tissue that surrounds the eye and aims to protect it, maintain the 3D shape of it and connect it with some of the movement muscles. The pupil is a black hole located in the center of the eye that allows light to enter the retina. The iris is the colored ring between the sclera and the pupil, it is made of an elastic tissue and it aims to control the amount of light entering through the pupil [231]. The iris begins to form during the third month of gestation and its structure is complete by the eighth month, although pigmentation continues through the first year after birth. These biological characteristics and the chaotic appearance of the iris patterns make it one of the most suitable traits for biometric purposes [208].

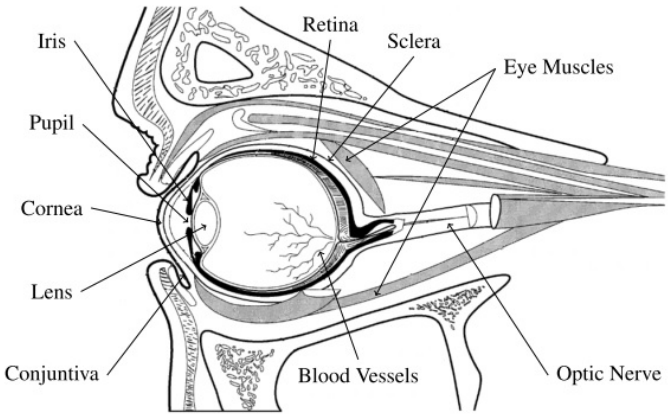


Figure 3.1: Human Eye Anatomy [208].

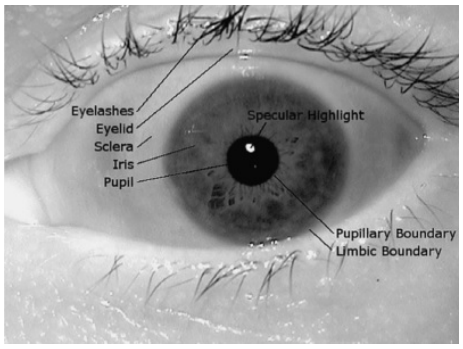


Figure 3.2: Photograph of a Human Eye [27].

3.2 Iris Recognition Methods

The two seminal methods in the iris recognition field were proposed by Daugman [52] and Wildes [289], and were considered as pioneers in this area.

Daugman's method

In his early work, Daugman established the main principles of a biometric system based on iris. His method for iris recognition [52] can be decomposed in four main steps as depicted in Figure 3.3. These main steps are: iris segmentation; iris normalization; feature extraction and feature comparison or matching.

Wildes' method

In 1977 Wildes presented a method [289] that differs substantially from Daugaman's method. Wildes' method for iris recognition can be divided in three parts: image acquisition; image segmentation and pattern matching, as depicted in Figure 3.4.

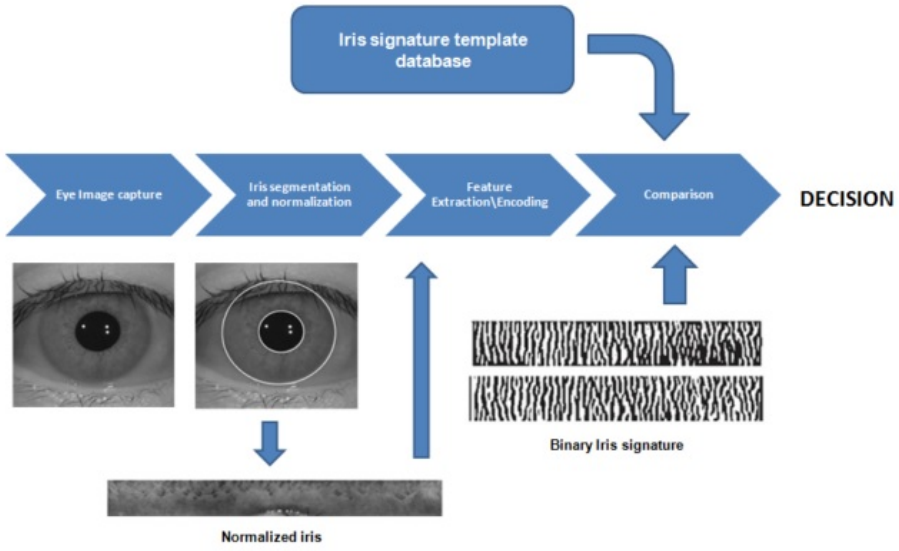


Figure 3.3: Schematic diagram of Daugman's iris recognition method [172].

In his work, Wildes considered that, since the iris' dimension is so small (about 1 cm of diameter), the acquisition of its image should be a major concern when developing iris recognition systems. The author also emphasizes the importance of obtaining images with high resolution, sharpness and good contrast in the iris pattern. Besides these two aspects, Wildes do emphasizes that the images should be well framed and the artefacts and obstructions should be eliminated.

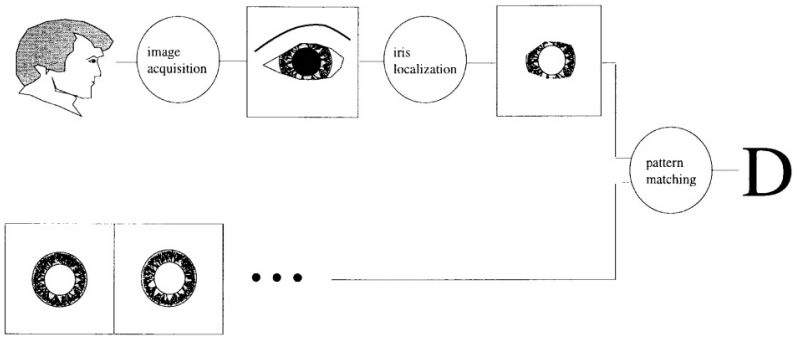


Figure 3.4: Schematic diagram of Wildes' iris recognition method [289].

3.2.1 Iris Segmentation and Normalization

Iris segmentation consists on the detection of the two defining contours of the iris region. Two main contours can be defined as the separating boundaries of the three aforementioned regions: the *limbic contour* separates the iris from the sclera, and the *pupillary contour*, the iris from the pupil. The detection of these contours is the main goal of segmentation and an essential step in the development of high accuracy recognition systems [35]. Once the iris region is successfully

segmented from an eye image, the next stage is to transform the iris region so that it has fixed dimensions in order to allow comparisons [165].

The original approach to the segmentation task was proposed by Daugman [52] and consisted in the maximization of an integro-differential operator. Some problems appear during the process of coding the information contained in the segmented iris region in a way that made it possible to compare between individuals when different iris' sizes are observed or when the dilation or contraction of the iris (as a result of non-uniform illumination) is variable [129]. To overcome these limitations, Daugman suggested a normalization step, known as the *rubber sheet model*. In this process, every location on the iris image was defined by two coordinates, relative to the previously detected iris center: an angle between 0 and 360 degrees and a radial coordinate ranging between 0 and 1, normalized to the radius of the iris. For each radial coordinate, a number of discrete points are chosen along the radial line that goes from minimum value of the radial coordinates to the maximum (i.e., from 0 to 1). Using this coding technique, regardless of the size of the iris or its contraction level, a rectangular image (with fix dimensions equal to the number of points chosen along the radial line and 360) containing all the iris information is obtained [27, 52, 208, 221].

In a different approach, Wildes [289] suggested a method involving edge detection followed by circular Hough transform (CHT). After the detection of contours, a normalization step is conducted that consists in an image registration process, where a mapping function is applied to the original image to compensate translational and scaling differences between acquired images and database templates.

For years, several works in the iris biometrics field focused on Daugman's and Wilde's algorithms, presenting variations at many levels.

Some examples of methods inspired in the classical ones are: the CHT-based method used for the segmentation step in Masek's algorithm [165]; the system that mixed both the CHT segmentation approach and the rubber sheet model normalization proposed by Ma *et al.* [154]; Krichen *et al.*'s [133] based on Gabor filter phase response for iris localization in which the segmentation step is an altered version of the CHT-based Masek algorithm; the work of Abhyankar and Schuckers [9] which begins with the transformation of the iris image into the wavelet domain followed by an enhancement of image contours and then finalized with the application of a Canny edge detector for enhancement and a CHT for the detection of both limbic and pupillary boundaries. Radman *et al.* [211] addresses a simple solution for the problem of high computational complexity of the integro-differential operator and the CHT by localizing the initial center of the pupil using a circular Gabor filter (CGF).

Active contours methods are applied such as in the work of He *et al.* [101], through an Adaboost-cascade iris detector a rough position of the iris centre is extracted and then the centre and radius of the circular iris are localised by employing an elastic model named *"pulling and pushing"*; Roy *et al.* [223] consider the iris as a non-circular structure and use an elliptic fitting model to fit both the limbic and pupillary contours perfected by a geometric active contour procedure based on an energy minimization process. Also, the segmentation of the pupil and iris by fitting a rotated ellipse was proposed by Zuo and Schmid [316].

Since iris boundaries are often not circular or elliptical, curve fitting techniques can be valuable to approximate real iris contours [207]. Several methods attempted to use active contour models to accurately localize irregular iris boundaries [60, 105, 240, 270]. An illustrative example for limbic and pupillary contour detection was presented by Lu and Lu [150]: first they used a deformable model to detect the pupillary contour, followed by the integro-differential operator to detect the limbic boundary. The approach taken by Chen *et al.* [37] consisted in detecting the sclera region of the eye, thresholding and filtering the image to detect a rectangular region for iris localization. An edge map of the region of interest is then obtained with a horizontal Sobel operator, and a dynamic programming variation of the CHT algorithm was implemented to detect the limbic boundary. This method corrects the non-circularities of the off-angle iris and combines the intersection of circles obtained by the two CHT algorithms and a linear Hough transform to perform eyelid detection. Pawar e*et al.* [197] applied geodesic active contours to perform segmentation.

Texture analysis can be used to perform segmentation as in the method proposed by Sanchez-Avila *et al.* [226] based on dyadic wavelet transform zero-crossing as iris signature where images were pre-processed by histogram stretching to aid the limbic boundary detection and then, the same algorithm is used inside its area to detect the pupillary boundary. Nabti and Bouridane's method [178] is based in a multiscale approach, using Gabor filters and wavelet transform coefficients, to improve edge detection process that determines the success of iris segmentation. Based on intensity gradient and texture difference is the method of Guo and Jones [94].

In the method of Tan *et al.* [259], first, a rough position of the iris is extracted by performing a clustering-based scheme to distinguish between iris candidates and the remaining image. Then, the regions resulting from this iterative process are analysed for specific iris characteristics, such as roundness and relative position to other regions. The second step consists in iteratively finding the shortest path that maximizes the Daugman integro-differential operator so that the limbic and pupillary boundaries can be detected.

Gradient vector field based methods have appeared in literature such as in the work of Chen *et al.* [35] where gradient flow around the iris centre plays an important role in the segmentation of the limbic contour.

When analyzing most of the methods cited in the literature, it is possible to detect some main drawbacks. In almost all of these methods, inner and outer boundaries, eyelashes and eyelids are detected in different steps, causing a considerable increase in processing time of the system. Usually, the inner and outer boundaries are detected by circle fitting techniques. This is a source of error, since the iris boundaries are not exactly circles and in noisy situations, the outer boundary of iris may not present sharp edges [20].

The competition NICE.I[1] was a good indicator of the methods for segmentation of the iris in unconstrained iris recognition with emphasis in the work with VW images considering the characteristics of the given database UBIRIS.v2 [207]. The top performing method [259] used a clustering-based iris localization scheme to perform a rough iris localization followed by an integro-differential constellation approach for fine pupillary and scleric border detection, which

[1] http://nice1.di.ubi.pt.

not only accelerates the traditional integro-differential operator but also enhances its global convergence. Finally, parametric models are learned to deal with eyelids and eyelashes. The second classified method [228] localized and filled the reflections in a YIQ (luma-chrominance) colour space and concluded that this approach has significant benefits for the subsequent processing. Then, it models iris boundaries by Daugman's classical integro-differential operators, followed by a parametric modeling of both eyelids. The third method [14] presented a knowledge-based approach inspired by the expert system paradigm, which directly encodes a set of "decision rules". The fourth method [142] performed a rough, fast eye detection and further normalized their region of interest using a c-means clustering technique. Their subsequent processing combines traditional iris segmentation techniques with RANSAC-like techniques. The method classified in the fifth position [117] used an Adaboost-based technique to roughly localize the iris and to compensate errors that result from the circular modeling of both iris borders after what the search was constrained for the iris boundaries exclusively within small stripes of the image, by means of iris linearization. Eyelashes and reflections were removed exclusively in these stripes and finally remapped the resulting boundaries into the original domain. The sixth method [152] proposed a very original strategy that detects the iris center through projection techniques and uses the center to translate the region of interest into a polar coordinate system, where morphological operators are used to roughly segment the outer iris border, which is then projected back to the Cartesian space in order to suppress the eyelids and eyebrows.

3.2.2 Feature Extraction and Matching

In order to provide accurate recognition of individuals, the most discriminating information present in an iris pattern must be extracted. Only the significant features of the iris must be encoded so that comparisons between templates can be made. Most iris recognition systems make use of a band pass decomposition of the iris image to create a biometric template. The template that is generated in the feature encoding process will also need a corresponding matching metric, which gives a measure of similarity between two iris templates. This metric should give one range of values when comparing templates generated from the same eye, known as intra-class comparisons, and another range of values when comparing templates created from different irises, known as inter-class comparisons. These two cases should give distinct and separate values, so that a decision can be made with high confidence as to whether two templates are from the same iris, or from two different irises. [165]

The original approach to the feature extraction and matching tasks proposed by Daugman [52] consisted in the use of 2D Gabor filters for texture analysis and feature extraction. To optimize computing times and lower calculation complexity, the resulting phase response to each Gabor filter was summarized in 2 bits: each pixel is assigned 1 to the first bit if the real part of the phase response is positive and 0 to the second bit if the imaginary part is negative. Thus, for every iris image a simple binary code was obtained, and the matching process against iris templates was performed by simple bitwise operations. The dissimilarity measure used by Daugman was the normalized Hamming distance, which measures the fraction of bits where the two binary codes

from iris signature and iris template disagree [27, 52, 208]. Such a simple way to quantitatively measure dissimilarity was only possible due to the binarization step of the Gabor filter response.

The method suggested by Wildes [289] proposes a different approach. In this method the feature extraction is accomplished by a multi-spectral analysis of the segmented iris using Laplacian-of-Gaussian (LoG) filters with distinct sizes. The matching is accomplished by normalized correlation between the tested database image and the normalized signature image, the result of which will express a similarity value between the two matched images.

For these steps of iris recognition process many different methods appeared after the classical works of Daugman and Wildes.

Focusing on the work done in unconstrained scenario, specially dealing with VW images we will focus our attention in the top classified methods in the competition NICE.II[1]. This competition was a good indicator of the state-of-the-art in methods for unconstrained iris recognition with emphasis in the work with VW images considering the characteristics of the given database UBIRIS.v2 [207]. The best performing approach was proposed by Tan *et al.* [260] which used both iris and periocular data. Global color-based features and local ordinal measures were used to extract discriminating data from the iris region, later fused to periocular data extracted from texton representations. Finally, fusion is performed by the sum rule using the normalized scores generated for the different types of features. Wang *et al.* [275] used an adaptive boosting algorithm to build a strong iris classifier learned from a set of bidimensional Gabor-based set of features, each operating locally and corresponding to a specific orientation and scale. Given the fact that the pupillary boundary is especially difficult to segment in VW data, the authors trained two distinct classifiers: one for irises deemed to be accurately segmented and another for cases in which the pupillary boundary was not accurately segmented. Santos and Hoyle [229] fused a set of recognition techniques that can be divided in two main categories: wavelet-based textural analysis methods applied to the iris region, complemented by distribution- based (histogram of oriented gradients and local binary patterns) and scale invariant feature transforms that analyze the periocular region, which was recently suggested as an important addition for handling degraded samples, essentially because it is less vulnerable to problems resulting from deficient illumination or low-resolution acquisition. Shin *et al.* [248] started by classifying the left and right eyes by their eyelash distributions,which they used to reduce the search space. Further, they coupled two encoding and matching strategies based in color and textural analysis to obtain multiple distance scores fused by means of a weighted sum rule, which is claimed to improve the separation between match and nonmatch distributions. Li *et al.* [143] used a novel weighted co-occurrence phase histogram to represent local textural features. This method is claimed to model the distribution of both the phase angle of the image gradient and the spatial layout, which overcomes the major weakness of the traditional histogram. A matching strategy based on the Bhattacharyya distance measures the goodness of match between irises. Finally, the authors concluded that the performance is improved when a simple image registration scheme accounts for the image deformation. Marsico *et al.* [61] proposed the use of implicit equations to approximate both the pupillary and limbic iris boundaries

[1] http://nice2.di.ubi.pt.

and perform image normalization. Next, they exploited local feature extraction techniques such as linear binary patterns and discriminable textons to extract information from vertical and horizontal bands of the normalized image. Li and Ma [144] introduced an image registration method based on the Lucas-Kanade algorithm to account for iris pattern deformation. Operating on the filtered iris images, this method divides the images into small sub-images and solves the registration problem for each small sub-image. Later, a sequential forward selection method searches for the most distinctive filters from a family of Gabor filters, concluding that a very small number of selected features is able to obtain satisfactory performance. Finally, Szewczyk *et al.* [255] presented a semi-empirical approach based on a reverse bi-orthogonal dyadic wavelet transform, empirically selecting a compactly supported bi-orthogonal spline wavelet for which symmetry is possible with FIR filters and three vanishing moments. The authors concluded that such a method produces a short biometric signature (324 bits) that can be successfully used for recognition under such challenging conditions, improving its reliability.

3.3 Databases for Iris Recognition

Several databases were constructed for testing iris recognition algorithms. Some are freely and publicly available, others have more restricted access. In this section we focus on the most commonly used databases and some very recent ones presenting new imaging devices. We present the databases describing its main characteristics, including the quality of the images, the conditions of imaging acquisition and the various kinds of noise factors that are observed in these images. Iris databases aim to promote the development of iris recognition and assess the technology's current level of performance therefore are a crucial tool for researchers.

CASIA

Apart from being the oldest iris database, this is clearly the most known and widely used by the majority of the researchers. The latest version of CASIA database, CASIA-Irisv4 [186], contains a total of 54,607 iris images. Some examples of these images are shown inf Figure 3.5. All iris images from the CASIA-Irisv4 database are 8 bit gray-level JPEG files, they present homogeneous characteristics and their noise factors are related with iris obstructions. It comprises six data subsets, which were collected or synthesized at different times: CASIA-Iris-Interval, CASIA-Iris-Lamp, CASIA-Iris-Distance, CASIA-Iris-Thousand, CASIA-Iris-Twins and CASIA-Iris-Syn.
This database has several similarities with the BATH database (Subsection 3.3) since its images were also captured under very constrained circumstances thus conditioning the resultant images. The images were also filled, in the pupil regions, with black pixels, which some authors used to facilitate the segmentation task. These facts significantly decreased the utility of the database, making the CASIA database less suited to develop algorithms to be used under unconstrained environments. [206]

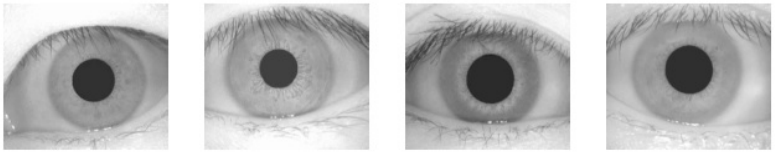


Figure 3.5: Examples of iris images from CASIA database [208].

BATH

The University of Bath iris image [265] database is a constantly growing database currently composed of over 16,000 iris images derived from 800 eyes of 400 individuals [208]. The images were taken from students and staff of the University, they are presented in gray scale and have very high quality. This database results of a project which aims to build an "high quality iris image resource". A series of acquisition and post-processing constraints assure the good image quality with the main sources of noise being only obstruction by eyelids and eyelashes. Furthermore, the main characteristics of these images are quite homogeneous, clearly resultant from a cooperative imaging setting. These facts makes this database not suitable for the development of iris recognition algorithms under unconstrained settings [206]. Some examples are presented in Figure 3.6.

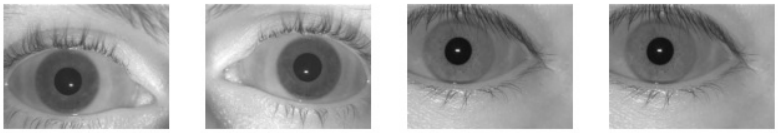


Figure 3.6: Examples of iris images from BATH database [208].

WVU

The WVU database [47] was developed by the West Virginia University and it consists of 1,852 images from 380 different eyes.
The images from this DB were captured under less constrained conditions and thus incorporate assorted types of noise, such as iris obstruction, poorly focused images and off-angle images. A few images have some regions affected by specular and lighting reflections which result from their acquisition under a natural environment. Some examples of the images from this database are shown in Figure 3.7.

MMU and MMU2

The MMU database [264] was developed by the Multimedia University and is constituted by 450 images of iris from 95 volunteers with different ages and nationalities. Afterwards, a new dataset with 995 images was made, MMU2. The images were captured from 100 subjects with

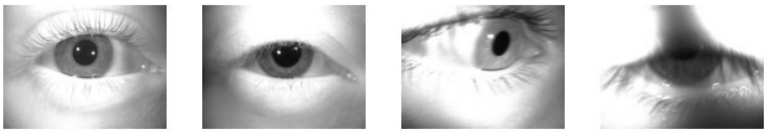


Figure 3.7: Examples of iris images from WVU database [208].

different ages and nationalities that provided five images from each eye. They come from Asia, Middle East, Africa and Europe. The images are highly homogeneous and their noise factors are exclusively related with small iris obstructions by eyelids and eyelashes and eye rotation [206]. Some examples are depicted in Figure 3.8.

Figure 3.8: Examples of iris images from MMU database [208].

UBIRIS.v1

The UBIRIS.v1 database [205] contains 1,877 colored images from 241 subjects and was collected in Universidade da Beira Interior in 2004, in two distinct sections. The set of 10 images from each subject is composed by 5 images from each session. The major purpose of UBIRIS.v1 database is the evaluation of robust iris identification methodologies. When it was created, this database constituted the world's largest public and free iris database for biometric purposes. Also, the existing iris databases were noise free so this database appeared as a new trend opposed to the constrained acquisition scenarios. Therefore, these images can be used to test and develop segmentation and recognition algorithms that are able to work with images captured under near perfect conditions. Its most relevant characteristic is to incorporate images with several noise factors, simulating less constrained image acquisition environments and the capture with and without user's cooperation. This enables the evaluation of the robustness of iris recognition methods. In Figure 3.9 are shown some examples of UBIRIS.v1 images.

ICE

The ICE database [187] is the database of the contest Iris Challenge Evaluation [187]. The database is constituted of 2,954 images, with a variable number of images per individual. The quality of the images was the main concern in the creation of this database. Therefore, the noise factors that the ICE database contains are mainly related with iris obstructions and poorly focused images. Another drawback with this database is that it was only made available for researchers and

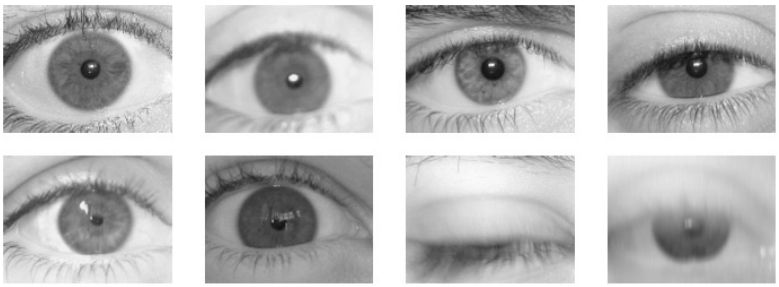


Figure 3.9: Examples of iris images from UBIRIS.v1 database [205].

entities that showed interest in participating in the competition [206]. Examples of ICE database are shown inf Figure 3.10.

Figure 3.10: Examples of iris images from ICE database [208].

UPOL

The UPOL database [64] is an iris image database that contains 384 images extracted from 64 subjects, with three iris images per eye. This database was built within the University of Palackého and Olomouc and its images were captured with optometric equipment, leading to very high quality images and maximum homogeneity. Some examples are depicted in Figure 3.11. This database's images have the singularity of being captured through an optometric framework (TOPCON TRC50IA) and, due to this, are of extremely high quality and suitable for the evaluation of iris recognition in completely noise-free environments. Also, these images have maximum homogeneity and inclusively the iris segmentation is facilitated by the dark circle that surrounds the region corresponding to the iris. Obviously, these characteristics make this database the less appropriate for the non-cooperative iris recognition research. [206]

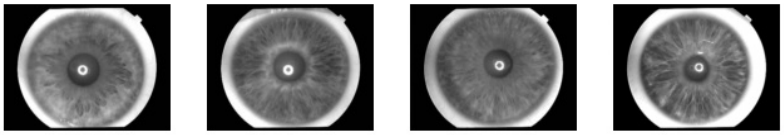


Figure 3.11: Examples of iris images from UPOL database [208].

UBIRIS.v2

The UBIRIS.v2 database [207] has 11,102 colored images from 522 subjects and, similarly to the first version of this database, it was collected in Universidade da Beira Interior in two sessions. The distinguishing points of this database from the previous ones are the distance used between the capture equipment and the user, the unconstrained acquisition conditions and the several types of noise in the images, such as iris obstructions, lightning and specular reflections, poor focus, partially captured or out-of-image iris, off-angle iris and motion blurred images. Examples of images with several types of noise are shown in Figure 3.12. The major purpose of the UBIRIS.v2 database was to constitute a new tool to evaluate the feasibility of visible wavelength iris recognition under far from ideal imaging conditions. In this scope, the various types of non-ideal images, imaging distances, subject perspectives and lighting conditions existent on this database could be of strong utility in the specification of the visible wavelength iris recognition feasibility and constraints. This database quickly became a milestone in the iris recognition field since it allowed to evaluated the methods in a unconstrained scenario.

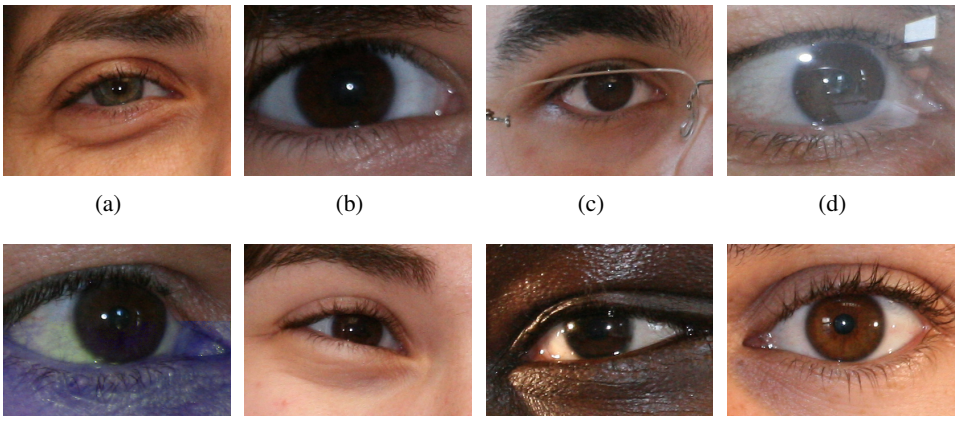


Figure 3.12: Examples image classes in the UBIRIS.v2 database: a) Heavily occluded; b) Heavily pigmented; c) Glasses occlusion; d) Reflection occlusion; e) Off-angle; f) Rotated eye; g) Black subjects and h) Normal.

VCMI

The VMCI database was created in a collaboration of two students from ESEIG (Escola Superior de Estudos Industriais e de Gestão) with the VCMI group. This database was created with the goal of testing recognition algorithms in images acquired with mobile devices [172]. The images from this database were captured with a Nokia smartphone and a Panasonic digital photographic camera and were acquired in uniform yet uncontrolled conditions. Ten images per device and per person were taken from a hundred individuals, providing 2,000 images. Some noise factors can be found in the images, such as occlusion and off-angle, as can be seen in Figure 3.13.

Figure 3.13: Examples of iris images from VCMI database [172].

Iris subset of MobBIO Multimodal DB

The MobBIO Multimodal DB [234] database was created by a group of researchers of the Visual Computing and Machine Intelligence (VCMI) group at INESC Porto and contains biometric data from the face, iris and voice of 105 individuals. The images are colored and were captured by the back camera of an Asus Transformer Pad TF 300T, with a resolution of 8 mega pixels and auto-focus. Some types of noise factors can be found in this database's images such as iris obstructions, glasses reflection and occlusion, reflection occlusion and off-angle iris as depicted in Figure 3.14. These images are all manually annotated providing ground truth data for both iris contours detection, limbic and pupillary.

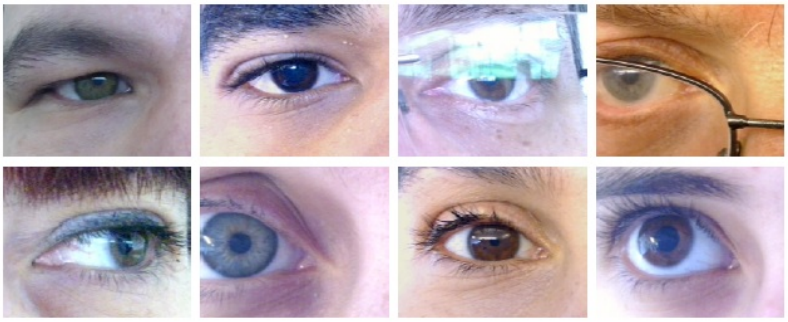


Figure 3.14: Examples of iris images from MobBIO database [234].

GUC - Visible Spectrum Smartphone Iris Database

The GUC - Visible Spectrum Smartphone Iris (VSSIRIS) Database [215] consists of iris images obtained from 28 different subjects in visible spectrum. The iris images are obtained using two different smartphone cameras (Nokia Lumia 1020 and iPhone 5S). The images are obtained in the unconstrained conditions under the mixed illumination consisting of natural sunlight and artificial room light. Each unique iris has 5 samples in the database and hence a total of 560 iris images. Some examples are shown in Figure 3.15.

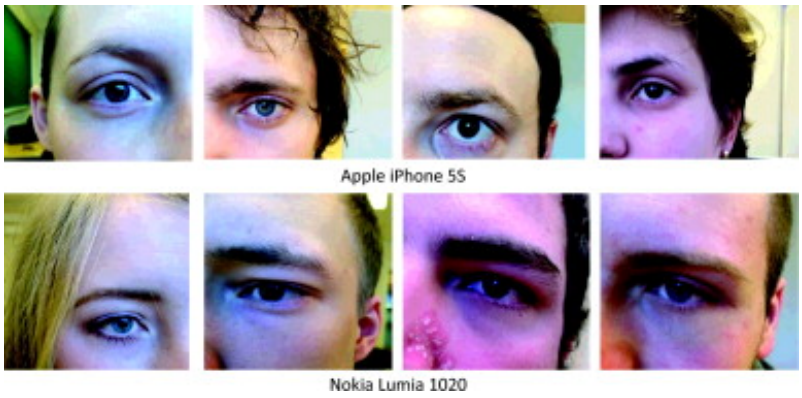


Figure 3.15: Sample images from the VSSIRIS database acquired using two different phones, from [215].

GUC Light Field Visible IRIS Database

The GUC Light Field Visible IRIS Database [1] database is comprised of 55 subjects that can represent 110 unique eye pattern. There are 5 samples captured for each eye using Lytro light field camera by simulating real life scenario. However announced, this database is not yet available, its authors are in the process of obtaining the permission in order make it available to the public.

MICHE Databases

The MICHE [266] competition originated several recent databases of iris images acquired with smartphones. These databases are not yet all of them publicly available, however, some are already being at disposal under request and others have appeared in publications of the same authors that have built them. One example is presented:

BIPLab Database

The smartphone iris database from BIPLab [267] consists of images captured from two different smartphones - iPhone 5 and Samsung Galaxy S4. The images are captured using frontal and rear camera in two different illumination - indoor and outdoor conditions. The indoor illumination corresponds to artificial light and outdoor illumination represents the varying daylight conditions. Although the database consists of 75 unique iris from two different smartphones, we have considered 50 unique iris which are captured from both smartphones. The set of iris employed in this work is captured in all different acquisition conditions representing indoor and outdoor illumination from both frontal and rear camera of smartphones. Further, each unique iris is accompanied by 4 samples making the database employed in this work consists of total number of 1600 images (50 Subjects $\times$ 4 Samples $\times$ 2 Smartphones $\times$ 2 Illumination $\times$ 2 Camera).

[1] http://www.nislab.no/biometrics_lab/vssiris_db

Table 3.1: Some characteristics of the images from iris images datasets.

Dataset	Sensors	Type	Resol. (dpi)	Size	No. imag.	No. eyes	Acquisition conditions
CasiaI. Interval	Casia CloseUp Iris Camera	GS	n/a	320×280	2,639	395	Very constrained, indoor
CasiaI. Lamp	OKI IrisPass-H	GS	n/a	640×480	16,212	819	Very constrained, indoor, lamp on/off
CasiaI. Twins	OKI IrisPass-H	GS	n/a	640×480	3,183	400	Outdoor
CasiaI. Distance	Casia Long range Iris Camera	GS	n/a	2352×1728	2,567	284	Indoor
CasiaI. Thousand	Irisking IKEMB100	GS	n/a	640×480	20,000	$2,000$	Indoor, lamp on/off
CasiaI. Syn	Casia Iris Image synthesis	GS	n/a	640×480	10,000	$1,000$	n/a
BATH	ISG LightWise	GS	n/a	1280×960	$16,000$	800	Very constrained
WVU	OKI IrisPass-H	GS	n/a	640×480	$1,852$	380	Less constrained
MMU	LG Iris Acces2200	GS	n/a	n/a	450	200	Constrained
MMU2	Panasonic BM-ET100US Authenticam	GS	n/a	n/a	995	200	Constrained

Table 3.2: Some characteristics of the images from iris images datasets (cont.).

Dataset	Sensors	Type	Resol. (dpi)	Size	No. imag.	No. eyes	Acquisition conditions
Ubiris.v2	Canon EOS 5D	RGB	72	400×300	11102	522	Less constrained with/without cooperation
VCMI	Nokia Smartphone; Panasonic Digital Camera	RGB	96	1601×1201	$2,000$	200	Less constrained with cooperation
Iris MobBIO Multimodal	Asus Transformer Pad TF 300T	RGB	72	250×200	800	200	Variable lighting cond.; indoor/outdoor; cooperation
GUC VSSIRIS	Nokia Lumia 1020; iPhone 5S	sRGB		7712×5360 / 3264×2448	560	56	Variable lighting cond.; indoor/outdoor; cooperation
GUC LFVIris	Lytro light field camera	RGB	n/a	n/a	550	110	Simulation of real life scenario

3.4 Summary of characteristics of databases

Table 3.2 summarizes some of the main characteristics of the most well known iris databases.

3.5 Commercially Available Solutions

Nowadays, the implementation of iris recognition systems is becoming usual and some of these systems can already be seen at airports, border control and social security services or even at private companies, which use them to give access to a room or area. For example, in bank account management we can find a system that requires the user to look at a camera for only few seconds to perform enrollment allowing subsequent rapid matching to access to bank accounts replacing the need of introducing a PIN [151]. Another example, is the commercial applications provided by *L1 Identity Solutions*[1] in many fields like *civilian identification management*, *criminal identification management* or even for *military* and *law enforcement* purposes. One example of aeroports application of iris recognition is the system in the Amesterdam Schipol[2].

In the remaining of this section, some commercially available systems are presented in more detail.

IRIS - Iris Recognition Immigration System

IRIS [6] was implemented by the United Kingdom border agency in order to regulate the flow of people entering the UK, namely frequent travelers. This system was used by the following airports: Heathrow, Manchester, Birmingham and Gatwick; however, since this is already an old system, the UK border agency decided to decommission it.

IrisGuard's Homeland Security Border Control

This system is used by the United Arabian Emirates' Expellee Tracking and Border Crossing Control System in all its 17 air, land, and sea ports since 2001, revealing some optimistic results relative to false matches as none was reported, from 2001 to 2004, despite of 2.7 thousand millions iris daily comparisons [59].

India's UID Program

The Unique ID program [2] is the Indian's unique identification number which is linked with biometric details – fingerprints and iris. They use this system to eliminate redundant records from their database and simplify the authentication process since the lack of some identification documents by the poorest residents is quite common.

CANPASS Air and NEXUS

[1] http://www.l1id.com/pages/17-biometrics
[2] http://www.cl.cam.ac.uk/ jgd1000/ibm.html

NEXUS pass is a binational boarding crossing pass for faster US Canada and USA border crossing [5]. Travellers with NEXUS pass can cross the border without being subject to regular questioning by customs and immigration officers [5]. People applying for the card must have an iris scan during an interview and some high-resolution iris images are taken [1].

CairoAmman Bank System

CairoAmman was one of the the first companies in the banking section to use an iris recognition system in their offices and ATMs. This iris recognition system is available at more than sixty ATMs and more than thirty seven offices [3].

Bank United System

The Bank United Corporation from Houston, USA, installed an iris scanning system at three of its ATMs and researched whether it was well accepted by the clients or not. The response to the survey, made by an independent research, was positive, as 98% of the users reported their experience as a good one [4].
Based on the positive response by the clients, they decided then to install the system at sixty ATMs in supermarkets across Texas [7].

Venerable Bede School - Impact

Venerable Bede School in Ryhope, England, uses an iris recognition system instead of ID cards for its students. The system used is called “Impact”, it was implemented in the fall of 2003 and is composed by an iris recognition camera which is integrated into a catering system. This way, students are identified and their meals are automatically charged to an account. They can also borrow library books or access restricted areas in the school, if they have permission [7].

3.6 Vulnerabilities of iris recognition systems

The specific characteristics of the human iris, in terms of universality, uniqueness, permanence and colectability make of it a good choice as a trait for a biometric system, as already analyzed in section 2.1. Nevertheless, as any other automated recognition technique, systems using iris as a biometric trait may be object of several attacks and in particular may be vulnerable to spoofing attacks. He et al. [100] refer some of the most common ways of forging an iris recognition system:

- *Use of an eye image* - e.g. photographs, video signal, screen images or paper print images;
- *Use of an artificial eye* - e.g. eye made of plastic, glass, silicon;
- *Use of a natural eye (user)* - i.e. forcing an individual to use the system;
- *Capture/Replay attack* - e.g. eye image or iris code template;
- *Use of a natural eye (impostor)* - i.e. eye removed from body or printed contact lenses.

Two of the most common and simple to perpetrate attacks are the use of printed images and contact lenses. In the first case, if a printed image from an iris of an authorized user is presented to the sensor the access is granted to the malicious agent, this is illustrated in Figure 3.16. The most common and simple approaches are those carried out with high quality iris printed images [224]. Other more sophisticated threats have also been reported in the literature such as the use of contact lenses [286]. If an individual enrolls into the system using cosmetic contact lenses, anyone wearing the same lenses can be authorized into the system, even unintentionally. Besides that, the texture of someone's iris can be printed into contact lenses with the purpose of illegally accessing an iris recognition system [286]. Two examples of patterned contact lenses are depicted in Figure 3.17.

Figure 3.16: Use of printed iris images in a biometric system [224].

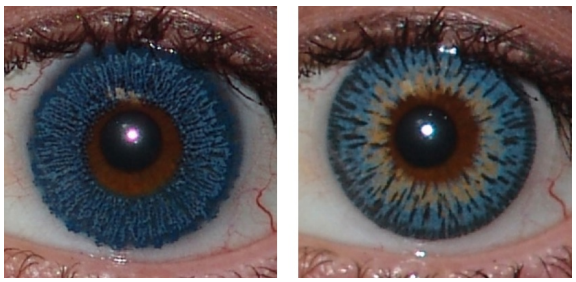


Figure 3.17: Examples of cosmetic contact lenses [273].

3.7 Iris Liveness Detection Methods

Among the different existing biometric traits, iris has been traditionally regarded as one of the most reliable and accurate. Also the imaging properties of the handheld devices makes this trait instinctive to use. The development of iris liveness detection techniques is crucial for the deployment of iris biometric applications in daily life. The evolution in the use of mobile devices in our society also raises the urge for liveness solutions in the mobile biometric field.

The feasibility of some attacks have been reported by some researchers [54, 58, 137] who showed that it is actually possible to spoof some iris recognition systems with printed iris and well-made color iris lens. Therefore, it is important to detect the fake iris as much as possible [100].

Several liveness detection methods have been presented through the past recent years. In fact, anti-spoofing techniques were presented that use physiological properties to distinguish between real and fake biometric traits. This is done in order to improve the robustness of the system against direct attacks and to increase the security level offered to the final user.

Iris liveness detection approaches can broadly be divided into:

- **Software-based techniques:** fake irises are detected once the sample has been acquired by the sensor and the detection of its liveness is done using the captured image;

- **Hardware-based techniques:** an additional device is added to the sensor in order to detect the liveness of the iris by measuring some particular properties. This type of detection can be done based in two types of eye behavior – voluntary (where the user moves or blinks the eyes according to system indication) or involuntary behavior (the eye moves without the user's demand, as the pupil oscillation, dilatation or contraction as a response to light) [55].

Even though hardware-based approaches usually present a higher detection rate, the software-based techniques have the advantage of being less expensive, as no extra device in needed, and less intrusive for the user which is a very important characteristic for a practical liveness detection solution. In general, a combination of both type of anti-spoofing schemes would be the most desirable approach to increase the security level of biometric systems. [83]

Detecting whether an iris is alive or not, promotes the robustness and reliability of a recognition system against direct attacks and helps obtaining a system with a high security level [83]. The main point of a software-based liveness detection process is to identify a set of discriminant features that allow the construction of an appropriate classifier that provides the probability of an iris being alive or not, based on the extracted set of features.

Daugman [57] and Tan *et al.* [153] proposed a software-based method of detecting iris liveness via frequency analysis, through FFT's – Fast Fourier Transform. This technique can only be used for printed iris detection as it uses the knowledge about the frequency characteristics of a printed iris image and a living iris. He *et al.* [99] suggested another method for iris liveness detection via statistical texture analysis for detecting the use of contact lenses. In this method, four features based on gray level co-occurrence matrix (GLCM) and properties of statistical intensity values of image pixels are extracted and a support vector machine is used for classification. Detecting iris edge sharpness is another possible measure for iris liveness detection. When contact lenses are used, fake iris edge is sharper than the edge of a living one [286]. Wei *et al.* [286] also proposed the use of texture analysis to detect contact lenses but in their work Iris-Textons are learned and used for texture representation. Some quality based features have been used individually for liveness detection in traits such as iris [127, 286] or face [140]. A strategy based on the combination of several quality related features has also been used for spoofing detection in fingerprint based

recognition systems [80] as well as in iris liveness detection [83]. In this latter work, a set of quality measures are used as iris liveness detection features to aid the classification of fake or real iris images included in a framework of feature selection. We find in literature that works concerning the quality of iris images are often the starting point to iris liveness detection techniques. One example is the assessment of the iris image quality based on measures like occlusion, contrast, focus and angular deformation [9], other is the use of texture analysis of the iris [99], among others like, for example, the analysis of frequency distribution rates of some specific regions of iris [153]. Other work that dealt with printed iris images was based on iris image frequency analysis [49]. In this work the author discuss how to select frequency windows and regions of interest to make the method sensitive to "alien frequencies"resulting from the printing process and the method showed to favorable compares to the results of commercial equipment used in the database development in the same work, as the referred device accepted all the printouts used. Sun et al. [301] recently proposed a general framework for iris image classification based on a Hierarchical Visual Codebook (HVC) which achieved state-of-the-art performance for iris spoofing detection, among other tasks. The HVC encodes the texture primitives of iris images and is based on two existing bag-of-words models. In a recent work, Menotti *et al.* [170] note that commonly iris anti-spoofing methods have explored hardcoded features through image-quality metrics, texture patterns, bags-of-visual-words and noise artefacts due to the recapturing process. The performance of such solutions vary significantly from dataset to dataset. Therefore, the authors propose to automatically extract vision meaningful features directly from the data using deep representations and regardless of the input type (image, video, or 3D masks). The method is inspired by the recent success of Deep Learning Methods in several vision tasks, and by the ability of the technique to leverage data, and focus on two general-purpose approaches to build image-based anti-spoofing systems with convolutional networks for several attack types not only for iris but also face and fingerprint. The first technique is hyperparameter optimization of network architectures and the second consists of learning filter weights via the back-propagation algorithm. A popular trend is to combine ocular features to iris features to detect spoof attacks. Tan and Kumar [258] developed a spoof iris detection approach that analyzes the image features such as intensity distribution, randomness of the iris texture, edge strength which are computed from the localized iris and periocular/ocular regions. To overcome the segmentation limitations, the developed approach also exploits the texture spectrum computed from the entire eye image, which can provide more effective descriptor for eye images which have failed to pass through the segmentation stage.

With the spreading of biometric applications in handheld devices, the spoofing detection problem has great importance in the mobile biometrics field. The appearance of new databases allow the development and evaluation of methods specific for these devices. However, as stated in a recent work [11], in the mobile field the problem is still unresolved owing to high level difficulty in determining efficient features with low computational cost to detect the spoofing attacks. Existing methods are not particularly targeted for liveness detection in mobile biometric applications, thus mainly inapplicable for portable devices. Hence, the authors present a multi-biometric approach, that can detect face, iris and fingerprint spoofing attacks in mobile applications, by employing a

novel real-time feature description based on order permutations, named Locally Uniform Comparison Image Descriptor (LUCID) [315]. The LUCID is computable in linear time with respect to number of pixels and does not require floating point computation, therefore the proposed approach is simple, fast and efficient, making it thus highly suitable for mobile devices. Moreover, contrary to existing schemes, this method utilize the same lone image descriptor technique and only one image for liveness detection, which can also be used for recognition. Experiments on publicly available face, iris and fingerprint data sets with real spoofing attacks show promising results. A work [12] combining the LUCID image descriptor with two other descriptors present the "MoBio LivDet" (Mobile Biometric Liveness Detection), a novel approach that analyzes local features and global structures of the biometric images using a set of low-level feature descriptors and decision level fusion. The LUCID image descriptor encodes the local information and CENsus TRansform hISTogram (CENTRIST) [294] global information. In addition, Patterns of Oriented Edge Magnitudes (POEM) [274] provides additional local and global structure information. The system allows user to balance the security level (robustness against spoofing) and convenience that they want. The proposed method is highly fast, simple, efficient, robust and does not require user-cooperation, thus making it extremely apt for mobile devices. Experimental analysis on publicly available face, iris and fingerprint data sets with real spoofing attacks show promising results.

A different approach relying in the advantages of fusing multiple traits is presented in [225], the proposed integrated multi modal biometric systems utilize two individual modalities, like IRIS and Electroencephalogram (EEG). In this work it is stated that: an EEG signal of each personality differs in such a way that they are not the same even if they do the same work or task; certain biometrics are vulnerable to noisy or bad data, such as dirty fingerprints and noisy voice records; even identical twins are not easy to be distinguished by face recognition systems; therefore the authors claim that authenticating users based on their EEG is more accurate than other biometric technologies.

The way forward seems to be the development of techniques for iris liveness detection that work well independently of the particular characteristics of the input image. In order to develop, improve and test iris liveness detection methods more adaptable to the new imaging scenarios, databases with fake samples are needed. A necessity was felt for databases of iris images captured with novel devices and constructed in less constrained conditions. Existing databases for iris liveness detection are presented in section 3.8. It can be observed that after the construction of MobBIO*fake* other databases captured with handheld devices appeared. Therefore, as exposed, new methods more adapted to this imaging scenario have been developed and evaluated in the new databases.

3.8 Databases for Iris Liveness Detection

In order to test iris liveness detection methods, databases with false samples are needed. The available databases for liveness detection include real samples and fake samples build either with printed images or by the use of contact lenses.

Clarkson

This database was made available for the contestants of the LivDet-2013 challenge [48]. It contains 270 images of real iris and 400 of fake ones. All the fake samples are images of iris with contact lenses, with a total of 14 types of lenses. The data was acquired through video (100 frames with focus variation) and two types of lighting are present in the database.

Figure 3.18: Examples of iris images from Clarkson database. Images a) and b) are real and c) and d) are fake images.

Notre Dame

The Notre Dame [66] database contains iris images: with cosmetic contact lenses; with clear soft lenses; and without contact lenses.

All images are 480×640 in size and were acquired under near-IR illumination, in two different periods – the soft lens and no lens data was acquired from 2008 to 2010 and the cosmetic contact lens images were acquired in 2012.

The total image database contains 1000 images without contact lenses, 1000 images with soft contact lenses and 1000 with cosmetic contact lenses, leading to a total of 3000 images.

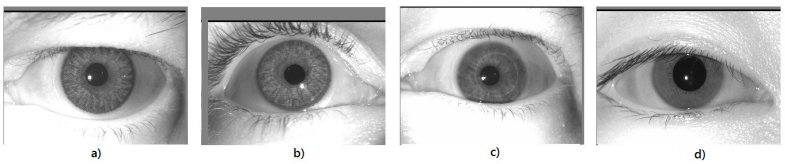


Figure 3.19: Examples of iris images from Notre Dame database. a) and b) correspond to images with cosmetic contact lenses; image c) has soft contact lenses and d) corresponds to an image without lenses.

Warsaw

The Warsaw [49] database contains 1274 real images and 729 fake images. This dataset was built using two different commercial acquisition equipments and two different printers. The fake images were obtained by printing the original ones. The database gathers images of only those printouts that were accepted by an example commercial camera, i.e. the iris template calculated for an artefact was matched to the corresponding iris reference of the living eye. This means that the quality of the employed imitations is not accidental and precisely controlled. An iris printout database of such properties was the first worldwide published. This database was partially made available for the participants in the LivDet-2013.

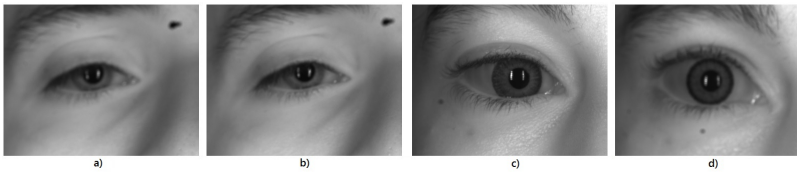


Figure 3.20: Examples of iris images from Warsaw database. a) and b) correspond to real images; images c) and d) corresponds to images with cosmetic lenses.

Biosec

The Biosec database [73] was created in the Polytechnic University of Madrid (UPM) and Polytechnic University of Catalunya (UPC) and contains both real and fake iris images. The images were taken in unconstrained conditions, in order to simulate a real situation, with a LG IrisAccess EOU3000 sensor. To build the fake part of the DB, the original images were preprocessed and printed in a paper, using a commercial printer. Then, the printed images were presented to the sensor, obtaining the fake samples. This study considered different combinations of pre-processing, printing equipment and paper type [224]. All images are in greyscale and theirs dimensions are 640×480 [83]. The two eyes of the same individual are considered as different users. The Biosec database contains 800 fake images and its correspondent real images. Fifty people participated in the image acquisition process which took place in two different occasions [224].

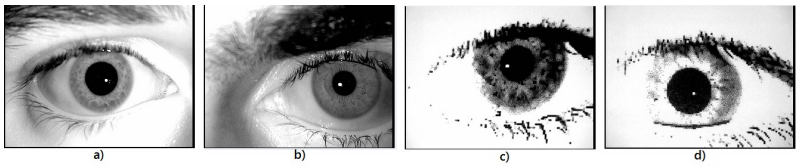


Figure 3.21: Examples of iris images from Biosec database. a) and b) correspond to real images and c) and d) to fake images.

IIITD Iris Spoofing Database

The IIITD Iris Spoofing (IIS) Database [95] was built upon the IIIT-Delhi Contact Lens Iris Database and consists of over 4800 images from 101 subjects with variations due to contact lens, sensor, and print attack. The IIIT-D CLI database [132, 295] comprises of 6570 iris images pertaining to 101 subjects . Both left and right iris images of each subject are captured and therefore, there are 202 iris classes. The lenses used in the database are soft lenses manufactured by either CIBA Vision or Bausch and Lomb. For textured lenses, four colors are used. To study the effect of acquisition device on contact lenses, iris images are captured using two iris sensors: (1) Cogent dual iris sensor (CIS 202) and (2) VistaFA2E single iris sensor.

GUC Visible Spectrum Iris Artefact Database

The GUC Visible Spectrum Iris Artefact (VSIA) Database [214] is comprised of eye images cap-

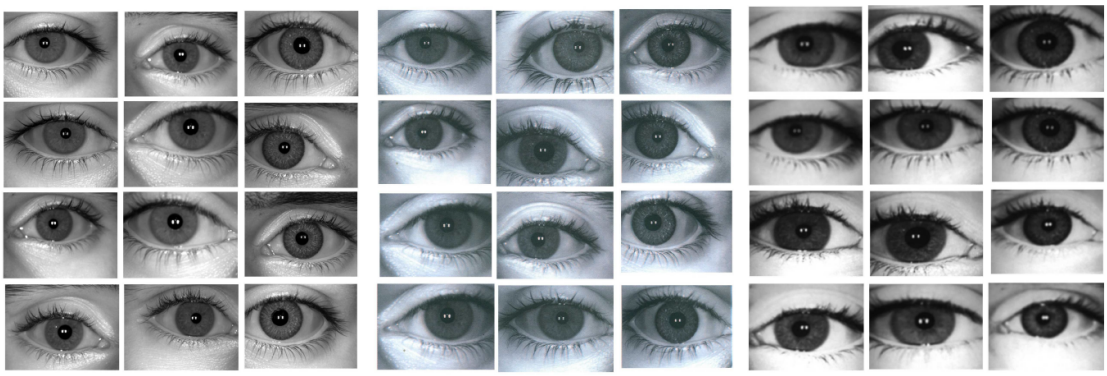


Figure 3.22: Examples of iris images from IIS database. (a) Images from the IIIT-D CLI database, (b) Print+Scan images from the IIS database, and (c) Print+Capture images from the IIS database. For each block, first column is iris image without lens, second column is with transparent lens, and third column is with textured lens (from [95]).

tured from 55 subjects (29 males and 26 females) that result in 110 unique eye patterns. For each subject, 5 samples for each eye were captured in five different instances resulting in $110 \times 5 = 550$ normal iris samples. All the artefacts are generated by considering the real-life attack scenarios that one can perform with visible iris recognition system. The artefacts generated are corresponding to 5 different kinds of static presentation attacks. In addition, this database also comprises a high quality artefact of visible eye image that can be used to attack the visible iris recognition system.

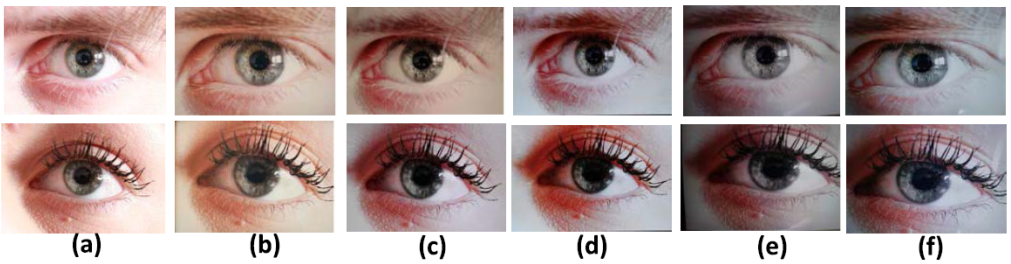


Figure 3.23: Examples of iris images from VSIA database. (a) normal eye sample, (b) Attack 1 samples, (c) Attack 2 samples, (d) Attack 3 samples, (e) Attack 4 samples, (f) Attack 5 samples. From [214].

GUC Light Field Visible Spectrum Iris Artefact Database

The GUC Light Field Visible Spectrum Iris Artefact Database (GUC-LF-VIAr-DB) [213] is composed of images from 52 subjects with 104 unique eye patterns. Since normal (or real) samples are captured using both light field and canon DSLR camera, this database has 520 (104×5 samples/eye) high quality samples and 4327 light field samples. The artefacts are generated by simulating 5 different kinds of attacks that resulted in 7607 samples with artefacts. The images were captured in an indoor scenario that has both natural (sun) and artificial (room) lighting. The data

collection was carried out over a period of 6 months in a single session. The real images were captured using two different cameras namely: (1) Lytro LFC with 1.2 Megapixe and (2) Canon EOS 550D DSLR camera with 18.1 Megapixel. For each subject, were collected 5 different samples using both Lytro LFC and Canon DSLR camera independently. The whole database consists of 520 (104×5) high resolution DSLR eye images and 4327 normal (or real) light field eye images (including multiple depth images). The artefact visible spectrum iris database was captured using only Lytro LFC. For the real samples captured using the Lytro LFC camera camera, the artefacts were generated using two different methods namely: (1) Screen artefact using iPad; and (2) Screen artefact using Samsung Galaxy Note 10.1, some examples are shown in Figure 3.24. For high resolution samples captured using the DSLR camera, the artefacts were generated using three different methods namely: (1) Photo print artefact; (2) Screen artefact using iPad; and (3) Screen artefact using Samsung Galaxy Note 10.1, some examples are shown in Figure 3.25.

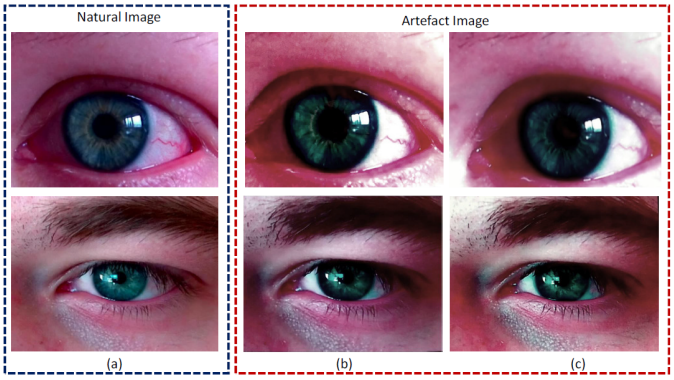


Figure 3.24: Examples of images from GUC-LF-VIAr-DB. Illustration of artefacts generated from best focus images obtained from the Lytro LFC (a) normal (or real) image, (b) iPad, (c) Samsung tablet. From [213].

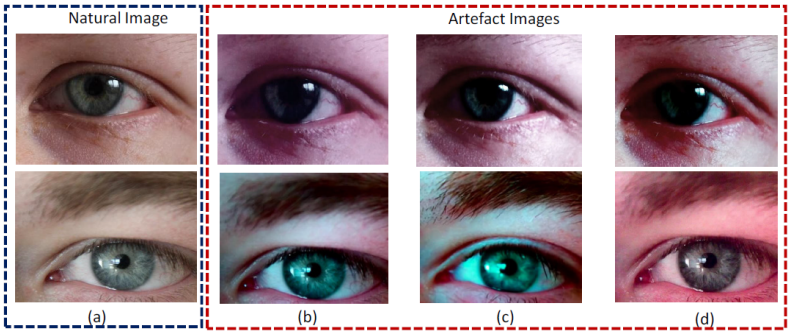


Figure 3.25: Examples of iris images from GUC-LF-VIAr-DB. Illustration of artefacts generated from DSLR images (a) Normal (or real) images, (b) iPad, (c) Samsung tablet, and (d) Photo print. From [213].

Table 3.3: Some characteristics of the images from iris liveness detection datasets.

Dataset	Sensors	Type of fake samples	No. Live	No. Fake	Size	Acquisition conditions
Clarkson	Acquired through video	14 types of contact lenses	270	400	780×600	Two lighting conditions
Notre Dame	Cross Sensor	Cosmetic and Clear Soft contact lenses	1000	2000	680×480	Constrained conditions
Warsaw	Commercial acquisition equipments	Printed images	1274	729	680×480	Constrained conditions; high quality images
Biosec	LG Iris AccessEOU3000	Printed images	800	800	680×480	Preprocessed fake images
IIITD	CIS202; VistaFA2E	Soft lenses; Printed Images	n/a	n/a	n/a	Preprocessed fake images
GUC VSIA	Lytro LFC; Canon EOS 550D DSLR; iPad; Samsung Galaxy	5 presentation attacks	550	7607	680×480	High quality artefacts

3.9 Summary of characteristics of databases

Table 3.3 summarizes some of the main characteristics of the most well known iris liveness detection databases.

3.10 Conclusion

The literature review shows beyond any doubt that iris recognition is nowadays a well studied field. This effort resulted in several real-world applications that are the evidence of the success of protocols comprising data acquisition with controlled conditions. However, the feasibility of iris recognition under unconstrained conditions is still a challenge, in fact, a tempting one due to the possible advantages of such technology in the economic and security areas.

In some of the algorithms mentioned in the literature review, there are a lot of implicit or explicit assumptions about the acquisition process, which are no longer valid in unconstrained acquisition scenarios. The most relevant of these conditions are:

- Image acquisition uses NIR illumination so that illumination can be controlled without human perception. Near-infrared illumination also helps revealing the detailed structure of heavily pigmented irises. Melanin pigments absorb much of visible light, but reflect more of the longer wavelengths of light (such as IR) [27, 53, 204].
- Subjects have to position their eye within the camera's field of view and stand still as the iris photographs are acquired [27].
- The iris and pupil are considered to always present a circular or elliptical shape [254].

Therefore, some of the promising results reported in the literature must be taken with caution and reassessed under new, more challenging, conditions. In recent years it has been recognized that the path forward, regarding iris recognition, is the development of algorithms that can work independently of subject collaboration and proper NIR illumination conditions, in order to achieve robust (i.e. accurate even with noisy images) and unconstrained (i.e. accurate for several sets of acquisition conditions: distance, movement, illumination) iris recognition and, in this way, become a real-world applicable method [221]. This paradigm shift led to the rise of new trends in the research of iris recognition, for example, exploring VL illumination instead of NIR and detecting iris contours with non-circular shape.

Till recently, the works in iris recognition continue to focus mainly on images captured with NIR illumination. For example, the latest iris recognition competition launched with the International Conference on Biometrics (ICB2013) uses the CASIA database (several of its versions, in fact) which is composed by images captured under highly constrained capturing conditions, yielding very homogeneous characteristics, their noise factors are exclusively related with iris obstructions by eyelids and eyelashes and the images were also filled, in the pupil regions, with black pixels. As referred, these facts significantly decreased the utility of the database in the development of algorithms to be used under unconstrained environments. Several databases continue to

appear that are composed by images acquired under these lightening conditions. Although they seek to comprise the so-called "noisy" images they still don't capture the complete unconstrained context as far as concerning the illumination conditions. Anyway, it is still a fact that the VW scenario will broaden the application of iris recognition to new contexts with special interest in situations that do not require cooperation of the user (like forensics and surveillance) and benefit from the user's ignorance that are being object of iris recognition. Although there are databases of images acquired under VW light there are still aspects of unconstrained scenarios to explore. One is for example, the acquisition of images in mobile devices.

With the rising of mobile biometrics the main concern is again security and circumvention of the iris recognition systems. The appearance of new databases of iris images acquired with hand-held devices both for iris recognition and iris liveness detection allow to evaluate methods more adapted to this emerging scenarios and to explore a new world in the iris biometric field. Nevertheless with new opportunities come new threats and new challenges to overcome the limitations posed by the rising scenarios.

Although the liveness detection field has registered significant advances, the fact is that in the mobile field the problem is still not resolved in part due to the difficulty in developing features that require low computational cost and are efficient to detect the spoofing attacks. Till recently, most of the existing methods were not particularly targeted for liveness detection in mobile biometric applications, thus mainly inapplicable for portable devices. In general, the success of an anti-spoofing method is usually connected to the modality for which it was designed. In fact, such systems often rely on expert knowledge to engineer features that are able to capture acquisition telltales left by specific types of attacks. However, the need of custom-tailored solutions for the myriad possible attacks might be a limiting constraint. The development of new solutions is an urge and new trends such as deep learning may bring powerful new tools to tackle this problem.

It is important to note that the search for reliability is not circunscript to the mobile biometrics. Any iris biometric application should take in account the vulnerabilities of these systems and the first point required is a step to assess the reliability of the sample obtained by the sensor. Therefore, iris liveness detection methods are crucial for the spreading of iris recognition systems with the desired confidence. So the main goal is to develop liveness methods that are more and more generally aplicable and independent of the intrinsic characteristics of images.

Chapter 4

Fingerprint Liveness Detection and Recognition Literature Review

In this chapter we present the basic concepts of fingerprint recognition and fingerprint presentation attack detection. A brief explanation about the fingerprint anatomy and its main features is given in the first section, then an overview of the types of sensors used for acquiring fingerprint images is given followed by a listing of some of the most well know databases for fingerprint recognition. After, we summarize some of the most relevant techniques for fingerprint recognition. The second part of the chapter is dedicated to the vulnerabilities of the fingerprint recognition systems, in particular, to spoofing attacks and to the presentation attack detection methods. In this context, the available databases for fingerprint liveness detection are presented.

4.1 Fingerprint Anatomy

Fingerprints are small lines/ridges and valleys in the skin of fingertips. Their configuration do not change throughout life, except if an accident, such as a burnt, happens. Fingerprints are formed at around the seventh month of fetus development due to a combination of genes and environmental factors [160, 227]. The environmental factors that influence the fingerprint formation result in such a high degree of variations that it is considered impossible to find two fingerprints exactly alike [160, 300]. Nevertheless, fingerprint patterns are not completely random and can be broadly divided in three types: arches, loops and whorls as depicted in Figure 4.1.

These three types of patterns can also be sub-divided into more specific groups such as: right/left loops, plain/narrow arches or spiral/concentric whorls [300]. Fingerprints can also be observed at three levels of detail: the global, local and very-fine levels. At the global level, singularity points such as core and delta can be found, examples of these are shown in Figure 4.2, but their distinctiveness is not sufficient for accurate matching. External fingerprint shape, orientation image, and frequency image also belong to the set of features that can be detected at the global level. At the local level, a total of 150 different local ridge characteristics, called minute details, have been identified [171]. Some of the minutiae points that can be observed are: ridge

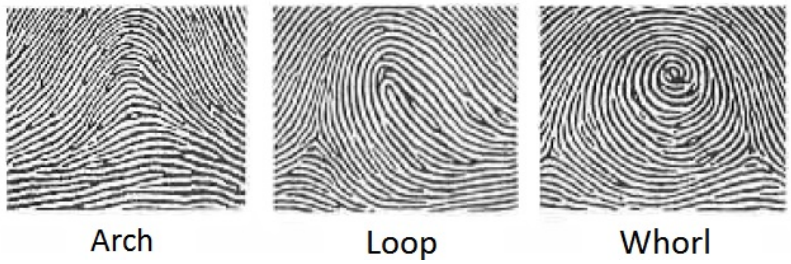


Figure 4.1: The three main fingerprint pattern types (from [227])

termination, ridge bifurcation, independent edge, dot, lake, spur and crossover, and examples of all these can be seen in Figure 4.2. The two most common ones are ridge termination and ridge bifurcation. Minutiae in fingerprints are generally stable and robust to fingerprint impression conditions. Although a minutiae-based representation is characterized by a high saliency, reliable automatic minutiae extraction can be problematic in extremely low-quality fingerprints devoid of any ridge structure. [160] Finally, at the very-fine level, intra-ridge details can be detected. These include width, shape, curvature, edge contours of ridges as well as other permanent details such as dots and incipient ridges [160]. One of the most important fine-level details whose positions and shapes are considered highly distinctive are the sweat pores, also shown in Figure 4.2. However, extracting very-fine details including pores is feasible only in high-resolution (e.g., 1,000 dpi) fingerprint images and therefore this kind of representation is not practical for non-forensic applications [160].

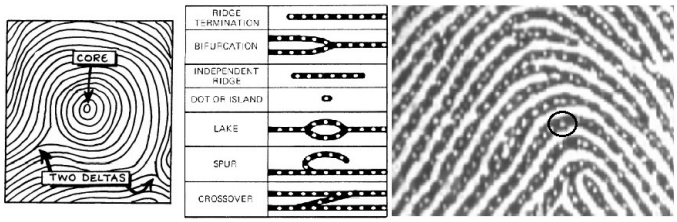


Figure 4.2: Core and delta points, types of fingerprint minutiae details and sweat pores (Adapted from [160, 227]).

4.2 Fingerprint Sensors

The traditional way of obtaining a fingerprint is the called *off-line* fingerprint acquisition, also called "ink-technique". In this technique a person's finger is stained with black ink and pressed or rolled on a paper. That paper is then scanned by a regular scanner and then a digital image of the fingerprint is obtained [160]. However, with the growth of demand and technological development, the most used technique is the live-scan acquisition in which digital images are obtained directly through a scanner. This development has also allowed the creation of smaller scanners,

with lower prices, that can even be integrated into laptops, smartphones or mouses [227].

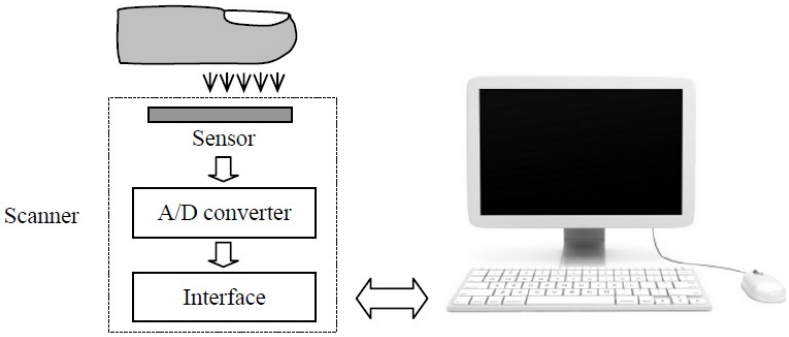


Figure 4.3: Typical structure of a fingerprint scanner [160].

Typically, the way of using a sensor follows these steps: it reads the ridge pattern in the surface of a finger, converts the analog reading to a digital form and an interface module communicates and sends the images to external devices (such as a personal computer or others). This structure is shown in Figure 4.3. Currently available scanners can be classified as *multi-finger*, if more than one finger can be acquired at the same time, or *single-finger*, if only one finger at a time can be acquired, as depicted in Figure 4.4.

Figure 4.4: Example of fingerprint scanners. The left image represents a multi-finger scanner and the right image represents a single-finger scanner [160].

Fingerprint images can also be stated as *plain*, *rolled* or *swept* considering the type of scanner used, such examples are shown in Figure 4.5. Although most scanners acquire plain impressions, some can also obtain rolled ones, which provide more information than plain images. On the other hand, swept images can be obtained using sweep scanners, which have the width of a finger but only a couple of pixels of height, reducing the cost of its manufacturing. However, this type of scanner has some drawbacks such as the learning time that a user needs to learn how to sweep the finger correctly and the time consumed to reconstruct the fingerprint image from the slices acquired.

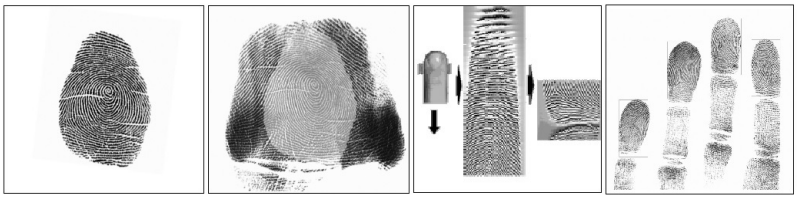


Figure 4.5: Examples of plain, rolled, swept and multiple fingerprints (Adapted from [160]).

In terms of live-scan fingerprint sensing, they can be sub-grouped in three categories: *optical*, *solid-state* and *ultrasound* sensors, being the optical and solid-state the most popular ones [160, 227].

4.2.1 Optical Sensors

Optical sensors can stand temperature fluctuations, are low cost and have better image quality than other types of sensors. However, they have some issues when dealing with damages, dirt or latent prints.

Frustrated Total Internal Reflection (FTIR)

When a subject places his finger in this type of sensor, the finger will touch the top side of a prism made of plastic or glass, as illustrated in Figure 4.6. Only the ridges will be in contact with the surface and there will be a gap between the sensor surface and the valleys.

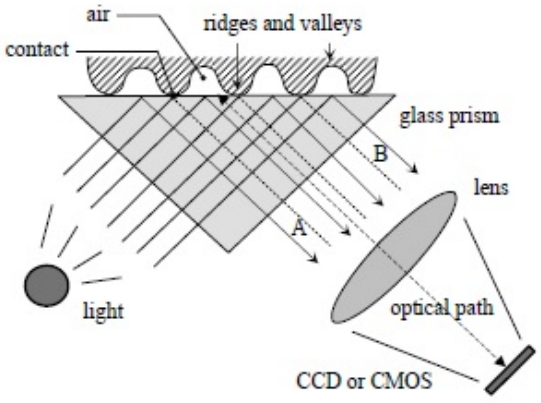


Figure 4.6: FTIR fingerprint acquisition [160].

One side of the prism is usually illuminated through a diffused light (e.g. LEDs). This light enters the prism and is reflected by the valleys (and absorbed by the ridges). As only the valleys are reflected, it is possible to acquire an image of the fingerprint through the caption (with a CCD or CMOS image sensor) of the light rays that exit the prism at its right side. It is also possible to use

a sheet prism made of a number of "prismlets" adjacent to each other instead of a single prism, in order to reduce the size of the sensor, nevertheless, this solution reduces the quality of the images obtained. The main advantage of this sub-type of sensor is that since it uses three dimensional information, it is difficult to spoof the sensor by using a flat printed image of a fingerprint. [160,227]

Optical Fibers

Instead of using a prism and a lens, this sub-type uses a fiber-optical plate. The finger is in contact with the upper side of the plate, having a CCD or CMOS on the opposite side, which receives the light conveyed through the fiber-optical plate, as depicted in Figure 4.7. Although the size of the scanner can be smaller, the size of the sensor has to cover the whole sensing area, which may result in increased costs of production [160, 227].

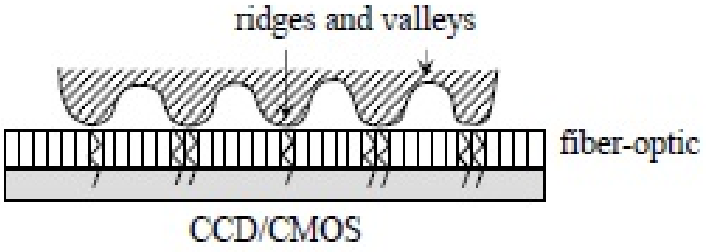


Figure 4.7: Optical-fibers fingerprint acquisition [160].

Electro-optical

This type of sensor have two main layers. The first layer contains a light - emitting polymer that when polarized with the proper voltage, emits light that depends on the potential applied on one side as illustrated in Figure 4.8. When a finger is placed on the sensor, the potential applied by the ridges and valleys (and consequent emitted light) is not the same, as ridges touch the polymer and valleys do not. That allows a luminous representation of the fingerprint pattern. The second layer consists of a photodiode array which receives the light emitted by the polymer and converts it into a digital image. Scanners using this technology are still behind FTIR in terms of image quality [160, 227].

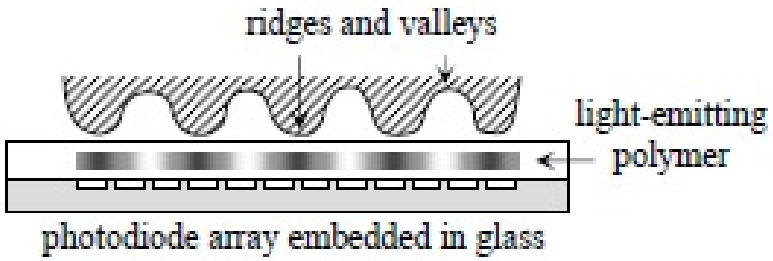


Figure 4.8: Electro-optical fingerprint acquisition [160].

Direct Reading

These sensors are called touchless as the finger is not in contact with any surface. These devices use a high quality camera to focus on the fingerprint but it is very challenging to obtain well-focused and high-contrast images [160, 227].

Multispectral imaging

This sensor captures several images of the same finger using different wavelengths of light, illumination orientation and polarization conditions. The resulting data can generate a single fingerprint image. This type of sensor is considered more robust than others but it is more complex and expensive and thus it is not the conventionally adopted scanner [160].

4.2.2 Solid-state Sensors

Solid-state sensors are also called silicon sensors and generally consist of an array of pixels, where each pixel is a very small sensor itself. They were designed to overcome some problems such as size and production cost.

Capacitive sensors

It is a two-dimensional array of micro-capacitor plates embedded in a chip. In this case, the skin of the finger works as the other plate of the micro-capacitor, one example is shown in Figure 4.9. Small charges of electricity are created between the finger's surface and each of the silicon plates. The magnitude of these electrical charges depends on the distance between the ridges or valleys and the capacitance plates. These differences can then be used to obtain an image of the fingerprint. These sensors, like the optical ones, cannot be spoofed by photographs or printed images. However, there are a number of disadvantages that should be taken into account, such as: they have a small sensor area which require more careful enrollments; electrostatic discharges

from the fingerprint may cause large electrical fields that can damage the device: the silicon chip needs to be protected from some chemical substances that may be present in finger perspiration [160, 227].

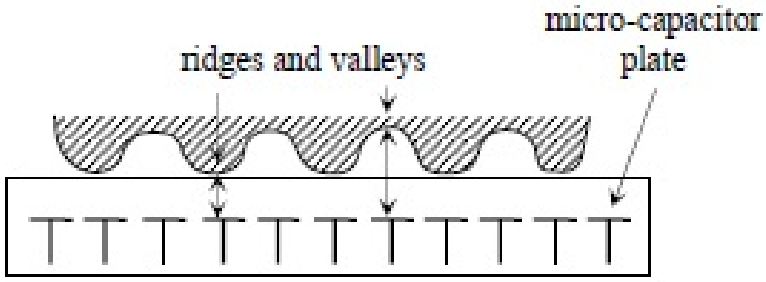


Figure 4.9: Capacitive fingerprint acquisition [160].

Thermal sensors

Thermal sensors are made of pyro-electric material. This material generates current based on temperature differentials. The difference between the temperature of the skin (of the ridges) and the air (in the valleys) is used to obtain the fingerprint image. These sensors are usually maintained at a high temperature to increase the difference between the sensor surface and the skin of the finger [160, 227].

Electric Field sensors

It creates a fingerprint image from below the top layer of the skin, avoiding the calluses, cuts and dirt that the finger might have. It consists of a drive ring and a matrix of active antennas. The ring generates a radio frequency sinusoidal signal and the antennas receive the signal transmitted by the drive ring and modulated by the subsurface of the skin [160, 227].

Piezoelectric sensors

Piezoelectric sensors are also called pressure-sensitive sensors. Its surface is made of a non-conductive dielectric material that generates current according to the pressure applied by the finger. The distinct pressure applied by the valleys and the ridges results in different amounts of

current [160, 227].

4.2.3 Ultrasound Sensors

An ultrasound sensor is based on sending acoustic signals toward the fingerprint and capturing the echo signal. This echo signal is then used to compute the depth image of the fingerprint. In order to operate that, this type of sensor contains two main components: a transmitter and a receiver, as can be seen in Figure 4.10. Even though it is known that ultrasound is probably the most accurate type of sensor for fingerprint recognition, its cost and size still largely influences its use [160, 227].

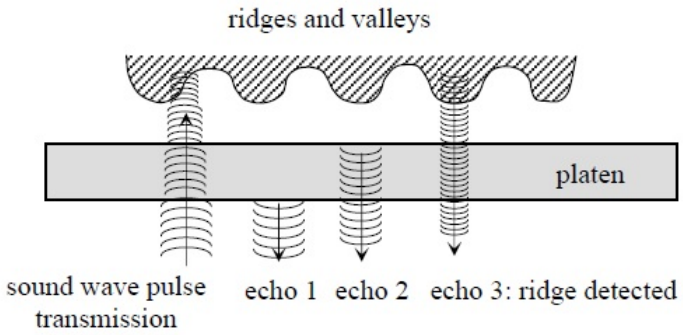


Figure 4.10: Ultrasound fingerprint acquisition [160].

4.3 Fingerprint Recognition Systems

Even though some fingerprint recognition methods directly compare images through correlation-based techniques, a representation based on the sensed gray-scale image intensities is not robust. Therefore, most of the fingerprint recognition and classification algorithms employ a feature extraction stage for identifying salient features [160]. Among the features extracted some have a direct physical counterpart, like singularities or minutiae, but others are not directly related to any physical trait as the orientation image or the filter responses. So, the several features extracted along the recognition process may be used either for matching or serve as an intermediate step for the derivation of other features. For example, some of the preprocessing and enhancement steps are often performed in order to simplify the task of minutiae extraction. Also, the segmentation step is done to avoid performing feature extraction in noisy areas such as the background. Therefore, fingerprint recognition methods can be broadly divided in three categories: correlation-based [21, 145, 290], ridge feature-based [8, 71, 114, 238, 239] and minutiae-based [136, 155, 285, 306] methods. In correlation-based method, two fingerprints are superimposed and the correlation between corresponding pixels is computed for different displacement and rotation values. According to the ridge feature-based method, features of fingerprint ridges

like pattern shapes, frequency, and orientation are used for matching process. Minutiae-based methods which are the most frequently used method, use minutiae points which are extracted from fingerprint images for matching.

Figure 4.11 presents a generic fingerprint recognition method with its several steps and the relation between them.

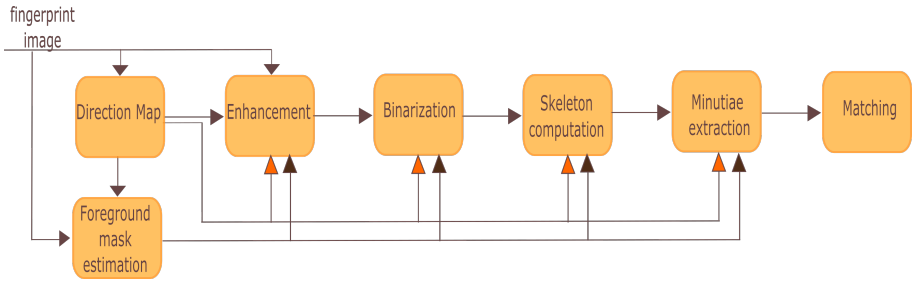


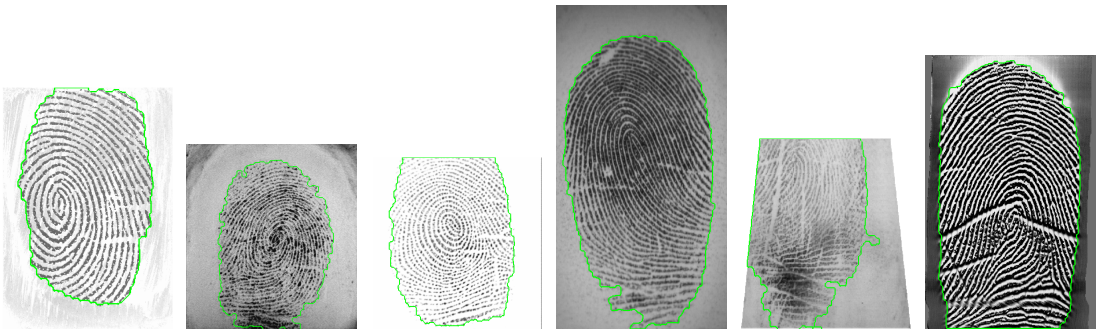
Figure 4.11: Fingerprint recognition system block diagram.

4.3.1 Fingerprint Segmentation

A fingerprint image is usually composed by two distinct regions: the foreground and the background. The foreground area, which contains effective ridge information, is caused by the contact of a fingertip with the sensor. The noisy area at the border of the image, with no effective information, corresponds to the background. The term segmentation is generally used to denote the separation of fingerprint area (foreground) from the image background. The separation of the content into these two groups allows to remove spurious minutiae located at the boundary of the foreground and to speed up the following tasks by restricting the processing to the foreground area, making the fingerprint segmentation one of the most relevant steps of an automatic fingerprint recognition system [160]. Separating the background is useful to avoid extraction of features in noisy areas that is often the background. Some examples of segmented fingerprints are shown in Figure 4.12.

A number of fingerprint segmentation methods are known from literature, which can be roughly divided into block-wise methods and pixel-wise methods [184]. Block-wise methods first partition a fingerprint image into non overlapping blocks of the same size, and then classify the blocks into foreground and background based on the extracted block-wise features. Pixel-wise methods classify pixels through the analysis of pixel-wise features. The commonly used features in fingerprint segmentation include gray-level features, orientation features, frequency domain features among others [184].

Several fingerprint segmentation methods have been proposed based in several different approaches. Mehtre *et al.* [168] proposed a method based on the local histograms of ridge orientations. In this method, the ridge orientation is estimated at each pixel and a histogram is computed for each 16×16 block. The presence of a significant peak in a histogram denotes an oriented pattern, whereas a flat or near-flat histogram is characteristic of an isotropic signal. Following



(a) (b) (c) (d) (e) (f)

Figure 4.12: Results of fingerprint segmentation in images of the FVC databases acquired with different sensors types, from [72].

the latter, Mehtre and Chatterjee [167] proposed a composite method that, besides histograms of orientations, computes the gray-scale variance of each block and, in the absence of reliable information from the histograms, assigns the low-variance blocks to the background. In the same line, Ratha *et al.* [217] proposed to assign each 16×16 block to the foreground or the background according to the variance of gray-levels in the orthogonal direction to the ridge orientation. A gradient based approach was proposed by Maio and Maltoni [155] which discriminated foreground and background by using the average magnitude of the gradient in each image block. The efficacy of this method is due to the fact that the fingerprint area is rich in edges due to the ridge/valley alternation; therefore the gradient response is high in the fingerprint area and small in the background. The method proposed by Shen *et al.* [241] performs Gabor filters convolution with each image block, and the variance of the filter responses is used both for fingerprint segmentation and for the classification of the blocks. Bazen *et al*. [21] suggested a pixel-wise method, in which three features (coherence, mean and variance) are computed for each pixel and then a linear classifier associates the pixel with the foreground or the background. The method presented by Chen *et al*. [36] uses a block cluster degree along with a linear classifier. Learning-based segmentation techniques were proposed in the works of Bazen and Gerez [21], Chen *et al*. [36], Yin *et al* [299] and Zhu *et al*. [312]. Chikkerur *et al*. [39] propose a method relying on the measure of local energy in the Fourier spectrum. The method uses the fact that ridges and valleys locally exhibit a sinusoidal-shaped plane wave with a well-defined frequency and orientation, whereas background regions are characterized by very little structure and hence very little energy. Wu *et al*. [292] proposed a fingerprint segmentation method based on the Harris corner detector. The image points with the strongest Harris response are considered as foreground points and the final segmentation is obtained as their convex hull. Liu *et al*. [146] suggested a classification using the AdaBoost classifier with two novel features (block entropy and block gradient entropy) and several commonly used features (coherence, mean, variance and Gabor features). More recently, Fahmy *et al*. [68] proposed a fingerprint segmentation method based on morphological operations. A range filter is applied to the fingerprint image followed by an adaptative thresholding binarization. A

final morphological post-processing step is performed in order to remove holes in both foreground and background. Ferreira *et al*. [72] proposed an extension of Fahmy's method [68] by introducing a new binarization, using a Fuzzy C-means (FCM) approach, process in the framework with the aim of overcoming some problems inherent to the image capture with various sensors. The effectiveness of the proposed algorithm was validated by experiments performed on all available databases of the Fingerprint Verification Competition (FVC).

4.3.2 Direction Map Computation

The fingerprint direction map computation step is crucial for the recognition process. The local ridge orientation at a pixel $[x, y]$ is the angle θ_{xy} that the fingerprint ridges form with the horizontal axis. As fingerprint ridges are not directed, the angle θ_{xy} corresponds to an non-oriented direction in $[0°...180°[$ [160]. Instead of computing local ridge orientation at each pixel, most of the fingerprint methods estimate the local ridge orientation at discrete positions, reducing computational effort and allowing estimates at other pixels to be obtained by interpolation [160]. Therefore, the fingerprint direction map (also called fingerprint *orientation map*, *orientation image* or *directional image*) is a matrix whose elements encode the local orientation of the fingerprint ridges. In Figure 4.13 an example is shown where a fingerprint image is faded into the corresponding direction map. Often an additional value is associated with each element of the matrix to denote the reliability of the orientation. This value is low for noisy and corrupted regions and high for good quality regions.

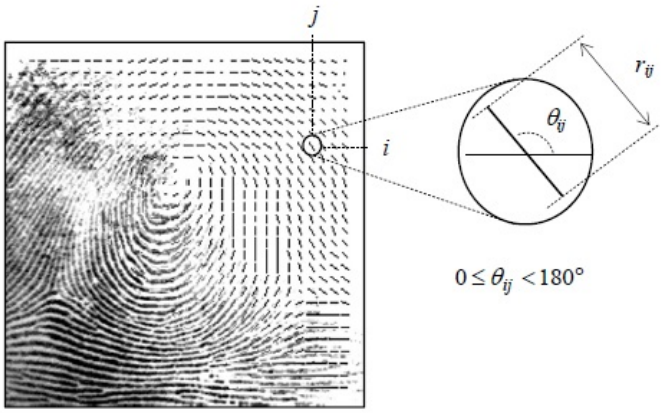


Figure 4.13: A fingerprint image faded into the corresponding direction map, from [160].

Methods for ridge orientation estimation are often gradient-based approaches, slit-and projection-based approaches, orientation estimation in the frequency domain, among others.

Gradient based methods

The simplest and most natural approach for extracting local ridge orientation is based on the computation of gradients in the fingerprint image. It is well known that the gradient phase denotes

the direction of the maximum intensity change. Therefore, the direction of a hypothetical edge that crosses the region centered in a point is orthogonal to the gradient phase at that point. This method, however simple and efficient, has some limitations. There are drawbacks caused by the non-linearity and discontinuity around 90°. Also, on one hand, there are the problems caused by the noise in the fingerprint image and its effect on a single orientation estimate; and on the other hand to the fact that simply averaging gradient estimates is not meaningful due to the circularity of angles. Kass and Witkin [128] proposed a simple but elegant solution to this limitation which allows gradient estimates to be averaged; their idea is to double the angles and encode an orientation estimate as a two dimensional vector. Based on the solution given, an effective method may be derived for computing the fingerprint orientation image used in other works. For example, in the method proposed by Ratha *et al.* [217] the dominant ridge orientation was computed by combining multiple gradient estimates within a 17×17 window centered at each pixel. Another example is the work of Bazen and Gerez [22], where the authors showed that the previous method is mathematically equivalent to the principal component analysis of the autocorrelation matrix of the gradient vectors. Donahue and Rokhlin [65] have before and independently proposed another gradient-based method which relies on least-squares minimization to perform the averaging of orientation estimates which leads to similar expressions.

It should be take in account that the major flaw of gradient-based orientation estimators is their failure in the near-zero gradient regions, namely ridge tops and valley bottoms due to the fact that in these regions the small values of both components of the gradient imply high noise sensitivity. [160] Therefore, some authors recommend to look beyond the first-degree derivatives [134]. Using second degree derivatives only partially solves the problem since the high noise sensitivity is moved to the zero crossing regions (i.e., inflexion points) where all the second order derivatives and the Hessian are null. The method by Da Costa *et al.* [50] is based on both first- and second-degree derivatives, for each region, a binary decision on which operators to use is taken according to the local coherence of the two operators.

Slit- and projection-based methods

The first slit-based approach dates back to 1960s, but some variants have been recently introduced. The basic idea is to define a fixed number of reference orientations or slits and to select the best slit based on the pixel grey-values along the slits. The local orientation at a given point is the orientation computed in a local window centered at that point. [160]

Stock and Swonger [253] proposed to sum the pixel gray-values along eight slits and select the minimum-sum slit or maximum-sum slit for ridge- or valley-pixels, respectively. For pixels lying on ridges (dark) the sum of gray-values along the ridge orientation is small. For pixel lying on valleys (bright) the sum is high.

The method proposed by Sherlock [242] suggests projecting the ridge lines inside a local window along a number of discrete orientations. The projection that exhibits the smallest variation corresponds to the orientation of the ridges within the window.

Ji and Yi [118] proposed a method where all the ridge lines except the central one are removed from the local window before computing the projections.

In the method proposed by Oliveira and Leite [195] the standard deviation of the gray-scale of the pixels corresponding to each slit is computed and the optimal slit according to the maximum standard deviation contrast between a slit and its orthogonal slit is selected.

The slit- and projection-based methods have usually higher computational complexity than gradient-based techniques and quantization might produce a coarser angular resolution. However, these methods allow to assign a probability value to each quantized orientation that can be useful to further process noisy regions. In other words, a gradient-based technique leads to a winner-take-all decision where just the optimal orientation is carried over, whereas in slit and projection-based methods one can also exploit the probability of the non-winning orientations for subsequent regularization or post processing [160].

Orientation estimation in the frequency domain methods

Kamei and Mizoguchi [126] proposed a method based on the application of 16 directional filters in the frequency domain. The optimal orientation at each pixel is then chosen according to the highest filter response taking local smoothing into consideration.

Analogous results can be achieved in the spatial domain by using Gabor-like filters, as proposed by Hong *et al.* [103] and Nakamura *et al.* [179].

The method proposed by Larkin [134] comprises two energy-based operators that provide uniform and scale-invariant orientation estimation. The second operator, the most robust one, is based on spiral phase quadrature. Although both the operators can be applied also in the spatial domain through convolution, the most natural and simpler implementation of these operators is in the frequency domain.

A method proposed by Chikkerur, Cartwright, and Govindaraju [39] is based on Short Time Fourier Transform (STFT) analysis. The image is divided into partially overlapped blocks whose intensity values are cosine tapered moving from the center toward the borders. For each block the Fourier Transform is computed and its spectrum is mapped to polar coordinates. The probability of a given angular value (within the block) is then computed as the marginal density function. The expected value of the same angular value for the block is estimated depending on the density function value.

Orientation image regularization

The orientation image corresponding to the fingerprint when computed from poor quality fingerprints may contain several unreliable elements due to creases, local scratches or cluttered noise. In this situation, a local smoothing can be very useful in enhancing the orientation image. This can be done by (re)converting the angles in orientation vectors and by averaging them. However, such a simple averaging has some limitations. [160]

To overcome the undesired effects described above, more elaborate approaches than a simple average have been proposed.

Jiang, Liu, and Kot [120] proposed a method in which, to prevent smoothing out high curvature regions, the size of the smoothing window is chosen according to a hierarchical coherence analysis. The authors of this method argued that since the noise is often caused by scars and breaks, it can be modeled as an impulse function. To suppress such kind of noise a non-linear approach, similar to a median filtering, performs better.

Zhu *et al.* [313] trained a neural network to classify orientation image elements into two classes: correct and incorrect, based on an 11-dimensional feature vector. After classification, the incorrect ridge orientations are corrected using orientation of the neighboring elements.

Zhang and Yan [308] proposed a method which define "invalid regions" in the foreground as the sets of connected elements with low coherence value, and use the contours of these regions to build a constrained Delaunay triangulation that is used to correct the orientations through interpolation.

In the method proposed by Oliveira and Leite [195] the correction is based on multi-scale analysis. They compute the orientation image at two different scales (fine scale and coarse scale) and correct only the elements whose value substantially differs between the two representations; in case of no substantial difference, the fine scale value is retained; otherwise the coarse scale value is used to correct the fine scale orientation image.

For very noisy images a local correction based on the above described techniques is often unsatisfactory and only the superimposition of a global model may provide a more effective improvement such as done in methods based in global mathematical models.

Lee and Prabhakar [138] proposed a method to compute the orientation image based on an MRF (Markov Random Field) made up of two components; one incorporates a global mixture model of orientation fields learned from training fingerprint examples and the other enforces a smoothness constraint over the orientation image in the neighboring regions. Although the current implementation is computationally intensive, it demonstrates the effectiveness of model-based estimation techniques. [160]

Global models of ridge orientations

A global "mathematical" model for ridge orientation can be very useful for several purposes such as orientation image correction, fingerprint data compression, and synthetic fingerprint generation. [160]

Sherlock and Monro [245] proposed a mathematical model to synthesize a fingerprint orientation image from the position of loops and deltas alone. This approach is also known as zero-pole model since it takes a loop as a zero and a delta as a pole in the complex plane. But the model makes some simplifying assumptions and it does not cover all possible fingerprint patterns. [160]

Later, Vizcaya and Gerhardt [272] improved the zero-pole model by using a piecewise linear approximation around singularities to adjust the zero and pole's behavior. This method also proposes an optimization technique, based on gradient descent, to determine the model parameters starting from an estimation of the orientation image.

In the methods based on the zero-pole model, the influence of a singularity on the orientation of a given point does not depend on the distance of the point from the singularity and this can cause errors in regions far from singular points. Furthermore, these models cannot deal with fingerprints with no singularities such as Arch type fingerprints. [160]

Therefore, another generalization of the zero-pole model was proposed by Zhou and Gu [310] by adding some pseudozeros and pseudopoles as the control points. In this method, the pseudozeros are the roots of the additional polynomial in the numerator and the pseudopoles are the roots of the additional polynomial in the denominator. The model parameters are derived by Weighted Least Square optimization.

The same authors Zhou and Gu [309] have introduced a combination model which models the real part and the imaginary part of the vector field with two bivariate polynomials, and improved the modeling in the noncontinuous regions around the singular points by imposing a Point-Charge Model for each singular point. Again as in the previous method they proposed, the model parameters are derived by Weighted Least Square optimization.

An iterative scheme based on a Bayesian formulation for the simultaneous extraction of orientation image and singularities was proposed by Dass [51] .

Li, Yau, and Wang [141] proposed a method which uses a first order phase portrait approach to compute the predicted orientation. The initial estimation of the orientation fields (computed through a gradient-based approach) is refined by replacing the unreliable orientations with the predicted orientations. The refined orientation is then used to obtain the final orientation model using a constrained non-linear phase portrait approach.

The method proposed by Wang, Hu, and Phillips [277] uses a 2D Fourier series expansion that has the advantage of not requiring the detection of singular points. However, similar to most existing models, it is closer to an approximation method rather than to a real modeling, since the natural variability of fingerprint patterns is not encoded in the model, and, when a large region is dominated by the noise, the recovery ability is limited.

Huckemann, Hotz, and Munk [107] argued that most of the global models proposed after Sherlock and Monro's one are controlled by too many parameters and therefore it can be critical to extract stable parameter values from a given orientation image. Therefore, they proposed a method which extends the zero-pole model, by using as control parameters five values with clear geometric meaning, and when comparing it with previous methods such as the one from Sherlock and Monro [245] and the one from Zhou and Gu [310] the authors concluded that their model better fits real fingerprint data extracted from NIST Special Database 4.

4.3.3 Ridge Frequency Estimation

The local ridge frequency (or density) at a point is the number of ridges per unit length along a hypothetical segment centered at that point and orthogonal to the local ridge orientation of the point. Given a fingerprint image, a corresponding frequency image can be defined if the frequency is estimated at discrete positions and arranged into a matrix. The local ridge frequency varies across different fingers, and may also noticeably vary across different regions of the same fingerprint. [160]

In the method proposed by Maio and Maltoni [156] , the ridge pattern is locally modeled as a sinusoidal-shaped surface, and the variation theorem is exploited to estimate the unknown frequency.

Hong *et al.* [104] estimate local ridge frequency by counting the average number of pixels between two consecutive peaks of gray-levels along the direction normal to the local ridge orientation. The method proposed is simple and fast, however, it is difficult to reliably detect consecutive peaks of gray-levels in the spatial domain in noisy fingerprint images. In this case, the authors suggest using interpolation and low-pass filtering.

The method proposed by Jiang [119] also computes the local ridge frequency by accumulating, for each column, the gray-levels of the corresponding pixels in the oriented window centered at the given point. However, instead of measuring the distances in the spatial domain, the author makes use of a high-order spectrum technique called mix-spectrum which enhances the fundamental frequency of the signal by exploiting the information contained in the second and third harmonic.

Zhan *et al.* [303] compared frequency estimation approaches operating in the spatial domain versus Fourier domain, and concluded that the former can be implemented more efficiently but the latter seems to be more robust to noise.

Chikkerur, Cartwright, and Govindaraju [39] proposed a method based on Short Time Fourier Transform (STFT) analysis.

4.3.4 Enhancement

The performance of minutiae extraction algorithms and other fingerprint recognition techniques relies heavily on the quality of the input fingerprint images [160]. An ideal fingerprint image, presents ridges and valleys which alternate and flow in a locally constant direction. In such situations, the ridges can be easily detected and minutiae can be precisely located in the image. However, in practice, due to skin conditions (like wetness or dryness in the acquisition moment, cuts, and bruises), sensor noise, incorrect finger pressure, and inherently low-quality fingers (as happens in the case of elderly people or manual workers), a significant percentage of fingerprint images is of poor quality [160]. Commonly, the same fingerprint image contains regions of good, medium, and poor quality where the ridge pattern is very noisy and corrupted.

Several types of degradation can be found associated with fingerprint images:

- the ridges are not strictly continuous; that is, the ridges have small breaks (gaps);

- parallel ridges are not well separated, caused by the presence of noise which links parallel ridges, resulting in their poor separation;
- existence of cuts, creases, and bruises on the finger.

These three types of degradation make ridge extraction extremely difficult in the highly corrupted regions. This leads to the following problems in minutiae extraction: (i) a significant number of spurious minutiae are extracted, (ii) a large number of genuine minutiae are missed, and (iii) large errors in the location (position and orientation) of minutiae are introduced. In order to ensure good performance of the ridge and minutiae extraction algorithms in poor quality fingerprint images, an enhancement algorithm to improve the clarity of the ridge structure is necessary.

The enhancement of a fingerprint image allow to improve the clarity of the ridge structures in the recoverable regions and mark the unrecoverable regions as too noisy for further processing. Usually, the input of the enhancement algorithm is a gray-scale image and the output may either be a gray-scale or a binary image, depending on the algorithm used and the goal of the method.

The most common methods can be broadly divided in pixel-wise methods; application of contextual filtering; multi-resolution enhancement and crease detection and removal [160].

Pixel-wise methods

In a pixel-wise image processing operation the new value of each pixel only depends on its previous value and some global parameters (but not on the value of the neighboring pixels). Pixel-wise techniques do not produce satisfying and definitive results for fingerprint image enhancement. However, contrast stretching, histogram manipulation, normalization [104] and Wiener filtering [93] have been shown to be effective as initial processing steps in a more sophisticated fingerprint enhancement algorithm.

Contextual filtering methods

In the contextual filtering, a set of filters is pre-computed and one of them is selected for each image region. The context is often defined by the local ridge orientation and local ridge frequency. Then the image is low-pass filtered along the ridge direction with the aim of linking small gaps and filling impurities due to pores or noise. After, a bandpass effect is performed in the direction orthogonal to the ridges to increase the discrimination between ridges and valleys and to separate parallel linked ridges.

The method proposed by O'Gorman and Nickerson [188, 189] was one of the first to use contextual filtering for fingerprint enhancement; in this method the authors defined a mother filter based on four main parameters of fingerprint images at a given resolution: minimum and maximum ridge width, and minimum and maximum valley width. Once the mother filter has been generated, a set of 16 rotated versions is derived, then the image enhancement is performed by convolving each point of the image with the filter in the set whose orientation best matches the local ridge orientation. Depending on some input parameters, the output image may be gray-scale

or binary. Sherlock, M onro, and Millard [243, 244] performed contextual filtering in the Fourier domain.

Hong *et al.* [104] proposed an effective method based on Gabor filters; these filters have both frequency-selective and orientation-selective properties and have optimal joint resolution in both spatial and frequency domains. Several posterior works made improvements in the use of Gabor filters for enhancement of fingerprint images. Greenberg *et al.* [93] introduced a change which allows the filter to better tolerate errors in local frequency estimates. The works of Erol *et al.* [67] and Wu and Govindaraju [291] relate the filter bandwidth to the local orientation coherence. The approach made by Bernard *et al.* [24] reduces the filter bandwidth if none of the responses to an initial set of filters exceeds a certain threshold. Yang *et al.* [297] argue that the fingerprint ridge and valley pattern does not always resemble a pure sinusoidal pattern, mainly because of the different values of ridge and valley width in some regions so they propose Gaborlike filters whose positive and negative peaks can have different periods and contextually adjust the two periods based on the local ridge width and local valley width, respectively. Zhu *et al.* [314] note that implementing Gabor-based contextual filtering with squared mask can lead to artifacts that can be removed if the mask support is circular. Wang *et al.* [276] suggest replacing standard Gabor filter with Log-Gabor filter to overcome the drawbacks that the maximum bandwidth of a Gabor filter is limited to approximately one octave and Gabor filters are not optimal if one is seeking broad spectral information with maximal spatial localization. Chikkerur, Cartwright, and Govindaraju [39] proposed an efficient implementation of contextual filtering based on short-time Fourier transform (STFT) that requires partitioning the image into small overlapping blocks and performing Fourier analysis separately on each block. The bandwidth adjustment proposed in this work seems to be more effective than the approach by Sherlock, Monro, and Millard [244]. A similar approach to the one by Chikkerur *et al.* [39] was introduced by Jirachaweng and Areekul [123], but their block-wise contextual information computation and filtering is performed in the DCT (Discrete Cosine Transform) domain instead of in the Fourier domain.

Multi-resolution methods

Multi-resolution analysis has been proposed to remove noise from fingerprint images. The multi-resolution enhancement consists of decomposing the image into different frequency bands (or sub-images) which will allow the compensation of different noise components at different scales: in particular, at higher levels (low and intermediate frequency bands) the rough ridge-valley flow is cleaned and gaps are closed, whereas at the lower levels (higher frequencies) the finer details are preserved. The enhanced image bands are then recombined to obtain the final image.

In the literature, different approaches have been proposed for multi-resolution: shape-adapted smoothing based on second moment descriptors and automatic scale selection [13]; based on wavelet decomposition [106]; or dyadic space scale [38]; or even methods that use a Laplacian like image-scale pyramid to decompose the original fingerprint into three smaller images corresponding to different frequency bands and then process each image through contextual filtering [77, 78].

Crease detection and removal methods

Some fingerprints are affected by the presence of a large number of creases, the incidence of such cases is higher in elderly people [160]. The presence of creases adversely influences the computation of orientation image and can lead to the detection of false ridge-ending minutiae.

Provided that contextual information, such as local orientation and frequency, have been correctly estimated, contextual filtering techniques are usually capable of filling the small ridge-gap produced by creases, nevertheless, if the image is very noisy or the creases are too wide, contextual filtering can fail. Furthermore, in some cases creases themselves could be used as features to drive or improve the fingerprint matching and therefore their deletion during enhancement may be undesirable.

The method proposed by Vernon [271] detect creases by the analysis of the Hough transform space derived from the ridge ending minutiae.

Wu *et al.* [293] and Zhou *et al.* [311] both modeled a crease by using a parameterized rectangle, followed by a multi-channel filtering framework to detect creases at different orientations. then a Principal Component Analysis is applied to estimate the crease orientation, length and width.

The method proposed by Oliveira and Leite [195] identify crease points by looking at the discordance between local orientations computed at two different scales. An approach based on Watershed transform is then applied to remove the creases.

4.3.5 Binarization and Skeleton computation

The simplest binarization approach uses a global threshold and works by setting the pixels whose gray-level is lower than the threshold to 0 and the remaining pixels to 1. However, different portions of an image may have different contrast and intensity and thus a single threshold for the entire image may not be sufficient for a correct binarization. Some binarization methods were proposed by Abutaleb *et al.* [10] and Zhang *et al.* [307]. Some binarization processes greatly benefit from an a priori enhancement. Also, some enhancement algorithms directly produce a binary output. The binary images are usually submitted to a thinning stage which allows for the ridge line thickness to be reduced to one pixel, resulting in a skeleton image.

Usually contextual filtering-based methods produce the most regular binary ridge patterns. With the aim of improving the quality of the binary images before the thinning step, some researchers have introduced regularization techniques which usually work by filling holes, removing small breaks, eliminating bridges between ridges, and other artifacts and are based in mathematical morphology tools.

In resume, after binarizing the image, a skeleton image is obtained by applying thinning operations to the binarized image. One example of a set of the fingerprint image, the corresponding binary image and the obtained skeleton is depicted in Figure 4.14.

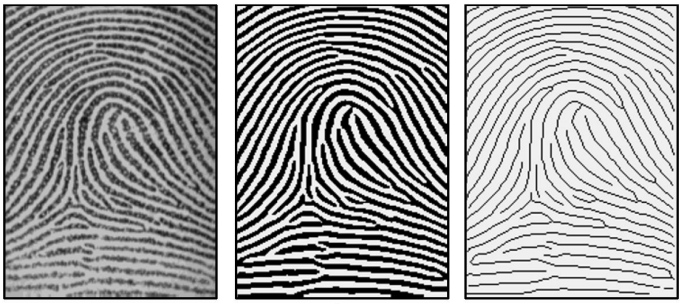


Figure 4.14: Fingerprint image, corresponding binarized image and skeleton image obtained after thining (from [160]).

4.3.6 Minutiae Extraction

While some authors propose minutiae extraction methods that need previous binarization, others use approaches that work directly with gray-scale images, as binarization is time consuming and some information may be lost during its process.

The binarization-based methods can use global or local threshold. Leung *et al.*. [139] proposed a method which extracts the minutiae from thinned binary images using a three-layer perceptron neural network. More recently, in the work of Gamassi *et al.* [84], for each point the algorithm counts the black-white transitions along a square path centered at that point and large enough to touch two ridges. Another work, proposed by Shin *et al.* [247] presents a method in which thick binary ridges are encoded with Run Length Code (RLC) and the minutiae are extracted by searching for the termination or bifurcation points of ridges in the RLC.

With the aim of overcoming some of the problems related to fingerprint image binarization and thinning (e.g., the presence of spurious minutiae in the case of irregular ridge edges), some authors have proposed direct gray-scale extraction methods. Usually the direct gray-scale extraction methods are based on ridge-line tracking and classifiers. Methods that perform direct gray-scale extraction are among others the ones proposed by Maio and Maltoni [155], Jiang *et al.* [121, 122], Liu *et al.* [147], Chang and Fan [34].

4.3.7 Matching

A fingerprint matching algorithm compares two given fingerprints and returns either a degree of similarity or a binary decision [160]. There are several factors that should be taken into account when performing fingerprint matching and that can influence its result, like displacements, rotations, partial overlaps, distortions, pressure, skin conditions and noise. Fingerprint matching techniques can be divided into three groups: correlation-based, minutiae-based and non-minutiae feature-based.

In the correlation-based methods, the correlation between two fingerprint images' pixels is computed for different alignments.

Regarding minutiae-based methods, minutiae are extracted from the two fingerprint images and stored as sets of points in a two dimensional plane. By comparing the two stored matrices, the method tries to find an alignment between them so that it results in the maximum number of minutiae pairings.

The non-minutiae feature-based matching are based in features other than fingerprints minutiae. Therefore, is performed a comparison of fingerprints in terms of features extracted from the ridge pattern. The most commonly used features are: size of fingerprint and external silhouette; number, type and position of singularities; global and local texture information; geometrical attributes and spatial relationship of the ridge lines; level 3 features (such as sweat pores).

4.4 Databases for Fingerprint Recognition

As the fingerprint was probably the first biometric trait to be used in a systematic way, a considerable number of fingerprint databases have been built along the time. Nevertheless, most of them are unavailable or private due to security concerns. That is, for instance, the case of the civil registration databases or police identification databases. Fortunately, there are several databases available for academic and research purposes. However, it is important to note that they are made available for researchers under some conditions and for user privacy respect usually these images are not depicted in scientific works without previous authorization from the databases owners.

4.4.1 National Institute of Standards and Technology Databases

The National Institute of Standards and Technology (NIST) has built several fingerprint databases over the last years. These datasets comprise images of all fingers and are not well suited for the evaluation of algorithms operating on live-scan images. These databases are available for purchase. (http://www.nist.gov/itl/iad/ig/fingerprint.cfm).

- NIST DB 4 [281]: 8-bit Gray Scale Images of Fingerprint Image Groups.
- NIST DB 9 [282] and NIST DB 10 [279]: 8-bit Gray Scale Images of Mated Fingerprint Card Pairs.
- NIST DB 14 [283]: NIST Mated Fingerprint Card Pairs 2; 8-bit Gray Scale Images.
- NIST DB 24 [280]: database of grayscale fingerprint images, corresponding minutiae and selected latent fingerprints corresponding to fingerprint images in the data set.
- NIST DB 27 [86]: this database was released to test the performance of latent fingerprint identification. This DB includes latent fingerprint images of varying quality together with their corresponding rolled impressions taken from cards; and minutiae data manually extracted by human experts.
- NIST DB 29: this database comprises Plain and Rolled Images from Paired Fingerprint Cards.

- NIST DB 30: this database contains dual resolution images from Paired Fingerprint Cards.

The NIST Databases 4, 9, 10 and 14 contain a large sets of images scanned from rolled inked impressions. NIST DB 24 contains 100 video sequences from 10 individuals. This database was mostly used to study the effect of finger rotation and plastic distortion. NIST DB 27 contains latent fingerprints and their corresponding rolled impressions. Minutiae data was manually extracted by experts and is also provided with the database [160].

4.4.2 Fingerprint Verification Competition Databases

On the scope of the "Fingerprint Verification Competitions" (FVC) that have been taking place over the years, 4 fingerprint databases were built: FVC2000, FVC2002, FVC2004 and FVC2006.

The FVC databases are all constituted by four subsets DB1, DB2, DB3 and DB4. In every database: the DB1 and DB2 subsets comprise images of the forefinger and the middle finger of both the hands (four fingers total) of each volunteer; the DB3 subsets is constitued by two images of six fingers (thumb, fore and middle on left and right hands) and four impressions were acquired of each of the six fingers of each volunteer; finally the DB4 subset is a synthetic database.

Fingerprint Verification Competition 2000 Database

The FVC2000 database [158] was released for the Fingerprint Verification Competition 2000 competition and comprises 1600 images that were divided equally in training and test set for the competition. The images of this database were captured in two sessions without quality check. The DB1, DB2 and DB4 subsets are considered Low/Medium difficulty and the DB3 dataset is considered to be of Medium/High difficulty. The volunteers were mainly students.

Fingerprint Verification Competition 2002 Database

The FVC2002 database [158] was released for the Fingerprint Verification Competition 2002 competition and comprises 1600 images that were divided equally in training ands test set for the competition. The images of this database were captured in three sessions without quality check. The DB1, DB2, DB3 and DB4 subsets are considered of Low difficulty. The volunteers were mainly students. In the capture of these images some perturbations were voluntarily exaggerated such as displacement, rotation, wetness and dryness.

Fingerprint Verification Competition 2004 Database

The FVC2004 database [159] was released for the *Fingerprint Verification Competition 2004* competition and comprises 1600 images that were divided equally in training ands test set for the competition. The images of this database were captured in three sessions without quality check. The DB1, DB2, DB3 and DB4 subsets are considered Medium difficulty. The volunteers were mainly students. In the capture of these images some perturbations were voluntarily exaggerated such as displacement, rotation, wetness and dryness.

Fingerprint Verification Competition 2006 Database

The FCV2006 database [29] was made available for the participants in the *Fingerprint Verification Competition 2006* and comprises 1800 images that were divided in 1680 for training set and 120 for the test set for the competition. The images of this database were selected from a larger database by choosing the most difficult fingerprints according to a quality index. The DB1 subset is considered to be of High difficulty; DB3 subset is considered of Medium difficulty; DB2 and DB4 are considered to be of Low difficulty. The set of volunteers was a heterogeneous population including manual workers and elderly people. Data collection was performed without deliberately introducing exaggerated distortion, rotations, etc. and the volunteers were simply asked to put their fingers on the acquisition device.

Figure 4.15: Examples of fingerprint images from FVC2006 database [227].

4.5 Vulnerabilities of Fingerprints Biometric Systems

Between acquiring biometric data and delivering a result, there are various points where attacks may occur and compromise the overall security of a fingerprint biometric system. Several weak links and vulnerabilities are identified by Ratha *et al.* in [218] as depicted in Figure 4.16:

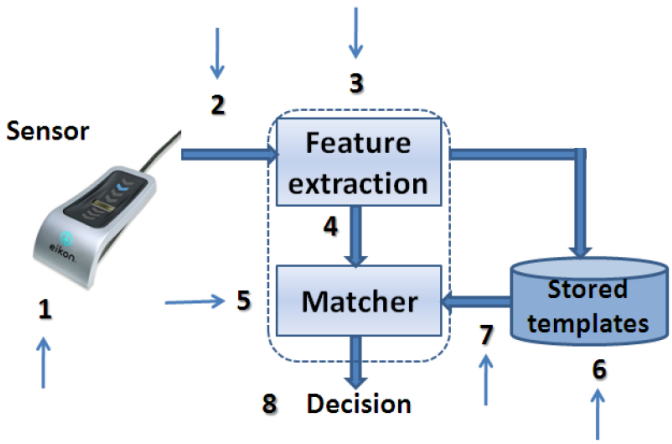


Figure 4.16: Vulnerable points of attacks in a fingerprints biometric system (from [162]).

These vulnerable points allow the folowing attacks to a fingerprint recognition system:

1. *Presentation Attack*. A reproduction of the biometric modality is presented as in- put to the sensor.
2. *Biometric Signal Replication*. The sensor is bypassed and biometric data previously stored or intercepted is resubmitted (e.g., copy of a fingerprint image).
3. *Feature Modification*. The feature extractor is substituted with a trojan horse in which features are preselected by the attacker.
4. *Replacing Features*. The set of features extracted from the input biometric trait is replaced with a fraudulent set of features.
5. *Overriding the Matcher*. The matcher is corrupted and forced to output match scores preselected by the attacker.
6. *Replacing Templates*. One or more templates are modified by an attacker such that an authorized identity is associated with a fraudulent template.
7. *Modifying Data through the Channel*. The templates transmitted through the channel are intercepted and corrupted.
8. *Altering the Decision*. The final match result is overridden by an attacker.

One of the main problems of fingerprint recognition systems, or biometric recognition systems in general, is their vulnerability to the presentation attack referred as number 1 in Figure 4.16. These systems can be spoofed by using an altered sample or a fake sample of the biometric trait used in a specific system.

One of the groups of techniques used by the impostors to spoof the fingerprint biometric sensors are fingerprint alterations. In this case, the impostor attempts to change his own fingerprint pattern by various means, such as cutting the finger, applying acids or even performing fingerprint transplantations. The main goal of the impostor is to destroy or alter his fingerprints to such an extent that the automated system is unable to find a match in the identification process. In such a case, he might avoid the consequences of being on the blacklist (e.g. because of recorded criminal activities, etc.). [251]

4.5.1 Fake Fingerprints Acquisition

The fake samples could be acquired with or without user cooperation. Naturally, with user cooperation, an authorized user may help an hacker to create a clone of his fingerprint. In order to obtain a fingerprint without the cooperation of the correspondent individual, it is necessary to obtain his/her print from a glass or other surface. Those marks left on surfaces are called latent fingerprints [160, 268]. Latent fingerprints can be painted with a dye or powder and then "lifted" with tape or glue. In spite of the advantage of not requiring cooperation from the user, these prints are, usually, low quality as they can be incomplete or smudged and thus are not very accurate. Either way, once the fingerprint is obtained, the easiest way of creating a fake sample is by printing

the fingerprint image into a transparent paper. However, a more successful method is to create a 3D fake model with the fingerprint stamped on it. This can be done by creating a mold that is then filled with a substance (silicon, gelatin, Play-Doh, wax, glue, plastic). This mold is used to create a thick or thin dummy (Figure 4.17) that an intruder can use. concerning the denomination used, although commonly the fake used in spoofing are referred as spoof or fake, in the International Standardization project ISO/IEC 30107, instead of the term fake, the term artefact is defined as follows: artificial object or representation presenting a copy of biometric characteristics or synthetic biometric patterns.

Figure 4.17: Play-Doh Finger mold and Silicon Finger Model.

4.6 Fingerprint Presentation Attack Detection

As already stated, biometric recognition systems in general, and fingerprint recognition systems in particular, can be spoofed by presenting fake or altered samples of the biometric trait to the sensor used in a specific system. Liveness detection techniques and alteration detection methods are both included in the *presentation attack detection* (PAD) methods [251].

Concerning spoofing attacks with fake samples, the fake samples can be acquired as described in section 4.5.1. Liveness detection methods are used to detect the use of fake samples in recognition and can be categorized as hardware or software-based whether the detection is performed through additional hardware or by processing the obtained image [45]. Hardware-based solutions work by measuring some physical characteristics and have the disadvantage of being expensive and requiring intervention at the device level. Memon *et al.* [169] make a detailed classification of the several methods of hardware based fingerprint liveness detection methods dividing them in three categories: *biological or involuntary signals*, *skin physiology characteristics* and *stimulation response*. Software-based liveness detection methods can be divided in two categories based on whether they work with a single static 2D scan (static methods), or need 2D scans at different time points during the acquisition process that support the observation of the dynamic properties (dynamic methods) [251]. The static methods comprise the analysis of textural features; sweat pores; ridge and valley structure; perspiration; and surface coarseness [251]. A schematic view of this categorization can be observed in Figure 4.18.

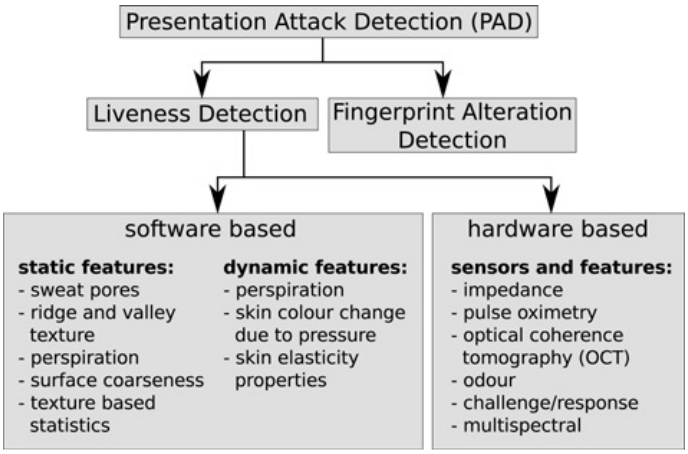


Figure 4.18: PAD methods (from [251]).

4.6.1 Fingerprint Liveness/Presentation Attack Detection Methods

Nikam and Agarwal proposed several methods based on the use of statistical analysis of the fingerprint scans, particularly, they propose the use of Local Binary Patterns (LBP) along with the wavelet transform [181]. It is known that a LBP's histogram can be a powerful texture feature and thus can be used to determine whether a fingerprint is real or fake. The same authors published several works proposing the use of Gray Level Co-occurrence Matrices (GLCMs) combined with diversified methods such as Gabor filters [182], wavelet transform [183] and curvelet transform [180]. In a recent work, Nogueira *et al.* [76] propose two different methods, one performing the feature extraction with LBP and another using Convolutional Networks (CN) for the first time in this task. The methods perform dimension reduction using Principal Component Analysis (PCA) and use a Support Vector Machines (SVM) with Radial Basis Function (RBF) kernel to perform classification. Ghiani *et al.* have proposed a method for liveness detection by using Local Phase Quantization (LPQ) [88]. The LPQ is a blur insensitive texture classification method. As it is able to represent all spectrum characteristics of images in a compact feature representation, avoiding redundant or blurred information, the authors believe that it could be used in this field. They used the four data sets collected for the Second International Fingerprint Liveness Detection Competition (LivDet11) [296] to test the algorithm. Ghiani *et al.* [87] also proposed the use of Binarized Statistical Image Features (BSIF) to detect the vitality of fingerprints. This approach has already been tested for face recognition and texture classification. Their goal is to obtain statistically meaningful representation of the fingerprint data by learning a fixed set of filters from a small set of images. They also claim that through learning, it is possible to adjust the descriptor length to the unusual characteristics of a fingerprint. Ghiani *et al.* tested this algorithm with the four LivDet2011 [296] datasets, obtaining promising results. However, there are still some problems with this algorithm, such as finding the right window size or the length of the binary string that results from the application of the filters to the image. In these two latter works, the classifier used was a linear support vector machine. Gragnaniello *et al.* [92] investigate the use of a local discrimi-

natory feature space, called Weber Local Descriptor (WLD) for fingerprint liveness detection. This descriptor consists of two blocks: differential excitation and orientation. These are then evaluated for each pixel of the image and concatenated into an histogram that is used to build the discriminatory features. A linear support vector machine classifier is then used to classify the images. These authors have tested this method with both LivDet2009 [164] and LivDet2011 [296] datasets and propose the combination of this method with the Local Phase Quantization (LPQ) [88] in order to obtain better results. Galbally *et al.* [80] use quality related features in their liveness detection work. The extracted features are ridge strength, ridge continuity and ridge clarity. They claim that those features can be extracted from the following quality measures: local angle, power spectrum and pixel intensity. The classification step is performed using linear discriminant analysis. This study presented an overall rate of 90% correctly classified samples, tested on a challenging database comprising over 10,500 real and fake images. This large database is created from the images of LivDet2009 [164] and ATVS [81] databases. Tan and Schuckers [257] proposed a new method for fingerprint liveness detection based on ridge signal analysis and valley noise analysis. They aim to quantify perspiration patterns along ridge in live samples and quantify noise patterns along valleys in fake samples. They present the performance of several standard pattern recognition classifiers including SVM. Their results show that the performance can reach 99.1% of correctly classified images. They have also proposed another method based on the statistics of Wavelet signal processing [256] aiming to detect the perspiration phenomenon using only a single image. In this work, the authors use classification trees for the classification task. Manivanan et al. [161] proposed a method to detect pores as a sign to fingerprint vitality using only one fingerprint image then applying two filtering techniques: highpass and correlation. The main reason of using highpass filter was to extract active sweat pore, then a correlation filter was used to locate the position of pores. Recently, Johnson and Shuckers [124] proposed a pore analysis method which still classifies the pores using their perspiration activity even if they are well represented in high quality fake fingers. A new invariant descriptor of fingerprint ridge texture, histograms of invariant gradients (HIG), is proposed by Gottschlich *et al.* [91] and again a SVM is used for classification. Warwante *et al.* [278] studied how the wavelet transform can be applied to fingerprint liveness detection. In this work, it is stated that wavelet analysis can help minimizing the effect of ridge and valley pattern when estimating the surface coarseness because it allows the study of the input image at different scales. They have created a high resolution database to which they then applied the proposed algorithm. Although they obtained positive results, one cannot say that the same would occur with images with less quality. In a recent work, Menotti *et al.* [170] proposed two general-purpose approaches to build image-based anti-spoofing systems with convolutional networks for several attack types in three biometric modalities, namely iris, face, and fingerprint. One technique is hyperparameter optimization of network architectures and the second lies at the core of convolutional networks and consists of learning filter weights via the well-known back-propagation algorithm. In this latter work the authors note that most of the state-of-the-art methods rely on hard-coded features sometimes exploring quality metrics directly related to the biometric trait that is being treated, such as directionality and ridge strength; or general texture patterns such

as LBP- or LPQ-based methods. In an innovative approach, their work is inspired by the recent success of *Deep Learning* techniques in several vision tasks. Some previous works that explore filter learning through natural image statistics, such as BSIF [87], or another that uses CN [76], are examples of this new research trend, which seeks to model the problem by learning features directly from the data. One basilar idea of this approach is to assume little a priori knowledge about acquisition-level biometric spoofing and explore deep representations of the data.

4.7 Databases for Fingerprint Liveness Detection

LivDet 2009 - Fingerprint Liveness Detection Competition 2009

This database was made available for the contestants of the LivDet 2009 [89]. Its images were acquired from three different devices: Biometrika FX2000 (Flat optical, 569 dpi, image size 312×372); CrossMatch Verifier 300CL (Flat optical, 500 dpi, image size 480×640); and Identix DFR2100 (Flat optical, 686 dpi, image size 720×720). The following materials were used in order to build the fake part of the database: Play-Doh, Gelatin and Silicon. The train and test datasets comprise over $5,000$ samples coming from around 100 different fingers.

LivDet 2011 - Fingerprint Liveness Detection Competition 2011

This database was made available for the contestants of the LivDet 2011 [296]. Its images were acquired from four different devices: Biometrika FX2000 (Flat optical,569 dpi, image size 312×372); Italdata ET10 (Flat optical, 500 dpi, image size 640×480); CrossMatch L SCAN Guardian (Flat optical, 500 dpi, image size 640×480); Atmel Fingerchip (Thermal sweeping, 96 dpi, image size not available). The fake fingerprints were generated following a consensual procedure using seven different materials: body-double skin-safe silicone rubber, ecoflex platinum catalysed silicone, gelatin, latex, modasil, playdoh and wood glue. The train and test sets comprise over $8,000$ samples coming from around 200 different fingers (depending on the dataset).

LivDet 2013 - Fingerprint Liveness Detection Competition 2013

This database was made available for the contestants of the LivDet 2013 [89]. Its images were acquired from four different devices: Biometrika FX2000 (Flat optical,569 dpi, image size 312×372), CrossMatch L SCAN Guardian (500 dpi, image size 640×480), Italdata ET10 (Flat optical,500 dpi, image size 640×480). and Swipe - Atmel Fingerchip (96 dpi, image size not available). More than 4000 images were taken with each of the aforementioned devices. The following materials were used in order to build the fake part of the database: Body Double, Latex, Play-Doh, Wood Glue, Gelatin, Silicon and Modasil. The train and test sets comprise over 8,000 samples coming from around 200 different fingers (depending on the dataset).

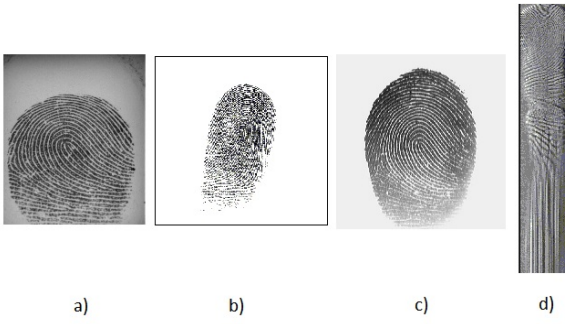


Figure 4.19: Examples of live fingerprints from the LivDet2013 datasets: a) Biometrika, b)Crossmatch, c)Italdata and d) Swipe.

Figure 4.20: Examples of fake fingerprints from the LivDet2013 datasets, using different materials a) Gelatin, b) Latex, c) Play-Doh, d) Body Double and e) Wood Glue.

ATVS database

The ATVS [81] database contains over 3000 live and spoof fingerprint images. The fake fingers were created using silicon and two methods were followed: with and without cooperation. Three devices were used to acquire the images: the flat optical sensor Biometrika FX2000, the flat capacitive sensor Precise SC100 and the thermal sweeping sensor Yubee with Atmel's Fingerchip.

4.8 Conclusion

However the amount of work done in fingerprint recognition and the continuous improvements made in the accuracy of this recognition systems the truth is that some of the steps can still be improved. The broadening of application of fingerprint recognition, specially into the mobile biometrics scenario, is arising and a transition has to be made surely based in the acquired knowledge in more traditional scenarios. In spite of the great amount of research in fingerprint recognition using traditional sensors the research using images from the new acquiring scenarios is taking its first steps.

Concerning fingerprint liveness detection, the performance of the state-of-the-art software fingerprint liveness detection methods suggests that additional hardware is necessary to develop a fingerprint liveness detection solution that would be resistant against targeted attacks. Owing to the large variety of possible artefact material and fabrication techniques, a single aspect dedicated

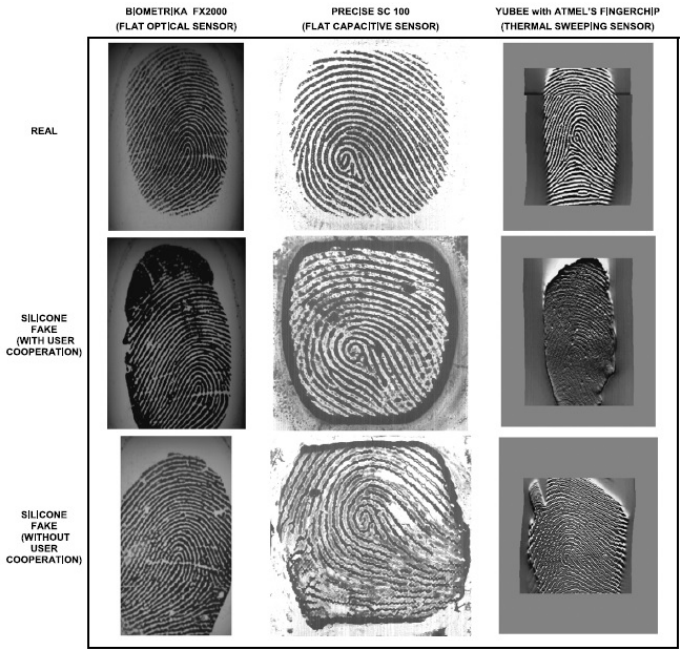


Figure 4.21: Examples of live and fake fingerprints from the ATVS database [80].

liveness detection sensor can usually be deceived if an appropriate new combination of materials and techniques has been used [232]. To increase the difficulty of producing an artefact, some manufacturers try to include a larger number of supplementary sensors that would capture information on multiple aspects of the scanned characteristic. Even though this greatly increases the difficulty of the artefact fabrication process, the large variety of the properties of the genuine fingers, as well as their artefact counterparts, requires the application of machine learning approaches to process the information from all the sensors and take the final decision whether a genuine characteristic has been presented. Since the performance of the machine learning-based classifiers depends on the training data, the sensor can still be vulnerable if an entirely new material and fabrication technique has been used to produce the artefact characteristic. [251]

Therefore, we consider that the traditional approaches to the liveness detection problem are not realistic and conduct to optimistic results. In our opinion, corroborated by recent works in literature, the path to pursue is to rely less on the prior knowledge about the fake samples and more on the knowledge about the real samples.

Part II

Biometric Databases and Biometric Competitions

Chapter 5

Mobile Multimodal Database Construction and Competitions Organization*

The existence of suitable databases is of utmost importance for the development and validation of new methods in the biometric field. More often than desired, the databases are not freely available or are not suitable for the scenarios which are intended to be tested. In the biometric field, we are facing the rising of a new acquisition scenario in which the conditions are not as nearly controlled as they used to be in the first biometric systems commercially available. This is mainly a consequence of the development of biometric applications in mobile devices. The already mentioned "Mobile Biometrics Field" is gaining more and more importance as our mobile devices are transformed in repositories of personal and professional data more than just communications goaled devices. The spread of handheld devices equipped with cameras and sound recorders has raised the opportunity of applying biometric recognition in mobile devices. Along with this phenomena, there is a necessity of adjusting the biometric systems to this new acquiring scenario and its proper limitations. Therefore, to develop and evaluate new methods there is a necessity for databases built in these conditions. In this chapter we present a multimodal database acquired with an handheld device, comprising samples of face, iris and speech.

We will present the construction of the mentioned biometric database, **MobBIO Multimodal DB**, and the organization of one biometric competition which used this database as benchmark, **MobBIO 2013**. The ambition of the competition was to become a reference event for academic and industrial researchers, especially in the area of security and personal identification. And also, establishing a public access multimodal biometric database represented an important contribution of the competition to the scientific community. We also show the impact of this event, to the present date, referring the interest that the database continues to raise in the scientific community and the works that have used it.

*Some portions of this Chapter appeared in [234]

5.1 Multimodal Biometrics

Most biometric systems deployed in real-world applications rely on a single source of information to perform recognition, thus being dubbed *unimodal*. Extensive studies have been performed on several biological traits, regarding their capacity to be used for unimodal biometric recognition. Table 2.1 summarizes the analysis performed by Jain [112] and Proença [208], regarding the qualitative analysis of individual biometric traits, considering the four factors: **Universality**, **Uniqueness**, **Collectability** and **Permanence**. A careful analysis of the advantages and disadvantages laid out in the table seems to indicate a couple of general conclusions: (1) there is no "gold-standard" biometric trait, i.e. the choice of the best biometric trait will always be conditioned by the means at our disposal and the specific application of the recognition process; (2) some biometric traits seem to present advantages that counterbalance other trait's disadvantages. For example, while voice's permanence is highly variable, due to external factors, the iris patterns represent a much more stable and hard to modify trait. However, iris acquisition in conditions that allow accurate recognition requires specialized NIR illumination and user cooperation, while voice only requires a standard sound recorder and even no need for direct cooperation of the individual.

This line of thought seems to indicate an alternative way of stating the two conclusions outlined in the previous paragraph: even though there is no "best" biometric trait *per se*, marked advantages might be found by exploring the synergistic effect of multiple statistically independent biometric traits, so that each other's pros and cons counterbalance resulting in an improved performance over each other's individual accuracy. Biometric systems that include multiple sources of information for establishing an identity are known as *multimodal biometric systems* [222]. It is generally regarded, in many reference works of the area, that multimodal biometric systems might help cope with a variety of generic problems all unimodal systems generally stumble upon, regardless of their intrinsic pros and cons [111]. These problems can be classified as:

1. Noisy data: when external factors corrupt the original information of a biometric trait. A fingerprint with a scar and a voice altered by a cold are examples of noisy inputs. Improperly maintained sensors and unconstrained ambient conditions also account for some sources of noisy data. As an unimodal system is tuned to detect and recognize specific features in the original data, the addition of stochastic noise will boost the probabilities of false identifications [115].

2. Intra-class variations: when the biometric data acquired from an individual during authentication is different from the data used to generate the template during enrolment [115]. This may be observed when a user incorrectly interacts with a sensor (e.g. variable facial pose) or when a different sensor is used in two identification approaches [222].

3. Inter-class similarities: when a database is built on a large pool of users, the probability of different users presenting similarities in the feature space of the chosen trait naturally increases [222]. It can, therefore, be considered that every biometric trait presents an

asymptotic behaviour towards a theoretical upper bound in terms of its discrimination, for a growing number of users enrolled in a database [115].

4. Non-universality: when the biometric system fails to acquire meaningful biometric data from the user, in a process known as failure to enrol (FTE) [115].

5. Spoof attacks: when an impostor attempts to spoof the biometric trait of a legitimately enrolled user in order to circumvent the system [110].

We shall note that taking advantage of the evidence obtained from multiple sources of information will result in an improved capability of tackling some of the aforementioned problems. These sources might be more than just a set of distinct biometric traits. Other options, such as *multiple sensors*, *multiple instances*, *multiple snapshots* or *multiple feature space representations* of the same biometric are also valid options, as depicted on Figure 5.1 [115].

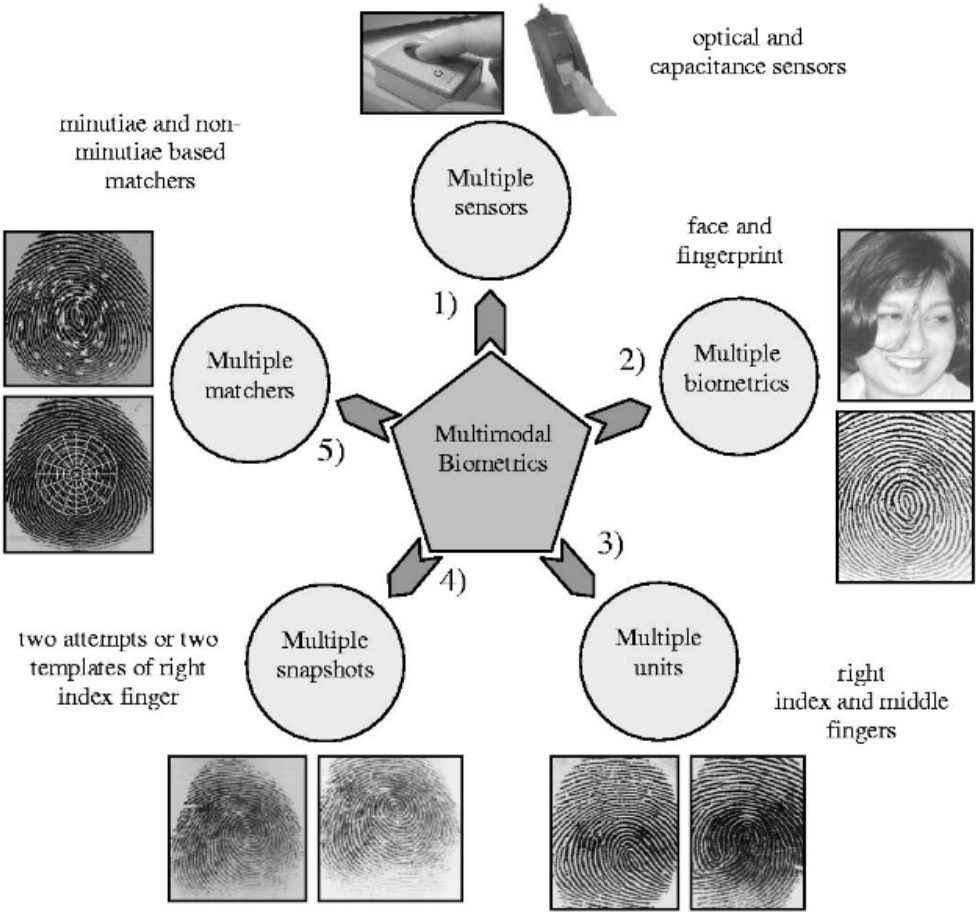


Figure 5.1: Scenarios in a multimodal biometric system. From [222].

But in this context we are focusing on the situation where we combine different biometric traits. In this case face, iris and speech.

5.2 Multimodal Databases

A strong trend observed as of late is the appearance of multimodal databases. As already referred, it seems obvious that the complementarity of some biometric traits will bring advantages and, consequently, a more accurate biometric recognition. When it comes to the choice of a biometric trait a vast list of possibilities is found, as shown in Section 2.1. This diversity gives rise, in existing multimodal databases, to many possible combinations of traits.

In 2003, the *BANCA* [19] database was released. This was considered a large, realistic and challenging multi-modal database intended for training and testing multi-modal verification systems. This database was captured in four European languages in two modalities (face and voice). For recording, both high and low quality microphones and cameras were used. The subjects were recorded in three different scenarios, controlled, degraded and adverse over 12 different sessions spanning three months. In total 208 people were captured, half men and half women.

In the same year, the first multimodal database with 5 modalities and time variability, launched by the *Multimodal Biometric Identity Verification* project, was the BIOMET [85]. The database was constructed in three different sessions, with three and five months spacing between them and contains samples of face, voice, fingerprint, hand shape and handwritten signature.

In 2003, the Biometric Recognition Group - ATVS made public and freely available the *MCYT-Bimodal Biometric Database* [74]. This database includes fingerprint and handwritten signature, in two versions containing data from 75 and 100 users, respectively offline and online signature acquisition.

Within the *M2VTS project (Multi Modal Verification for Teleservices and Security applications)* the database XM2VTS [202] was launched, comprising several datasets including face images and speech samples. According to its authors, the goal of using a multimodal recognition scheme is to improve the recognition efficiency by combining single modalities, namely face and voice features. At cost price, sets of data taken from this database are available including high quality color images, 32 KHz 16-bit sound files, video sequences and a 3d Model of each subject's head.

In the aforementioned databases there are several limitations, such as the absence of important traits (e.g., iris), limitations at sensors level (e.g., sweeping fingerprint sensors), and informed forgery simulations (e.g., voice utterances pronouncing the PIN of another user) [196]. The *BioSec Multimodal Biometric Database Baseline* [73] was an attempt to overcome some of these limitations. This database included real multimodal data from 200 individuals in two acquisition sessions including fingerprint, iris, voice and face. However the two releases of this database are now under construction and are not available at the moment. An enlarged version of the previous database is the *Multiscenario Multienvironment BioSecure Multimodal Database* (BMDB) [196] which comprises signature, fingerprint, hand and iris acquired in three different scenarios. This database is not freely accessed.

The *WVU/CLARKSON: JAMBDC - Joint Multimodal Biometric Dataset Collection* project gave rise to a series of biometric datasets, available under request and with costs. Integrated within

the aforementioned project, the West Virginia University constructed two releases of biometric data containing six distinct biometric modalities: iris, face, voice, fingerprint, hand geometry and palm print. The two releases differ only in the number of subjects. Within the same initiative, the Clarkson University created another dataset which contains image and video files for the same modalities except for hand geometry [48].

The *IV2 Multimodal Biometric Database* [198] includes Iris, 2D, 3D, Stereoscopic and Talking Face Data. Because reliable face recognition is still a great challenge to computer vision and pattern recognition researchers, the IV2 database is focused around this modality. Known variabilities that are critical for the performance of the biometric systems (such as pose, expression, illumination and quality) are present. The face and sub-face data that were acquired in this database are: 2D audio-video talking face sequences, 2D stereoscopic data acquired with two pairs of synchronized cameras, 3D facial data acquired with a laser scanner, and iris images acquired with an portable infrared camera. The database is designed for monomodal of multimodal experiments.

The multimodal database *MyIdea* covers many biometrics traits and is as such qualified as "multimodal". Some of those modalities are acquired individually, some others are acquired simultaneously, when it makes sense according to some scenarios. More precisely, the recorded domains are the following ones: video (bimodality face/voice), fingerprints, palm prints, hand geometry, signatures, bimodality signature/voice, writing, bimodality writing/voice. This database was meant to complete already-existing mono- or multimodal databases, as *BANCA*, *BIOMET* or *MCYT*. Sensors and recording protocols have been designed to guarantee data "compatibility" with these databases as much as possible. The data are acquired following realistic use scenarios of biometric technologies. For example, video sequences are acquired in a work environment context, or in a "degraded" context, as an electronic banking station. Also, sensors of different quality were used to acquire the data, in order to allow studies about the impact of those different qualities on the performance of algorithms.

The MOBIO database [166] consists of bi-modal audio and video data taken from 152 people. The speech samples and the face videos were recorded using two mobile devices: a mobile phone and a laptop computer.

The emergence of portable handheld devices, for multiple everyday activities, has created the need for mobile identity verification applications. The objective of research is to create a reliable, portable way of identifying and authenticating individuals. To pursue this goal, the availability of testing databases is crucial, so that results obtained by different methods may be compared. It is noted that the existing databases do not completely fulfill the requirements of this line of research. On one hand, there are limitations in the variety and combination of biometric traits, and on the other hand some of the databases are not publicaly accessible, limiting their usability.

5.3 MobBIO: a Multimodal Database Captured with a Portable Handheld Device

The development of biometric recognition systems is generally limited by the shortage of large public databases acquired under real unconstrained working conditions. Database collection represents a complicated process, in which a high degree of cooperation from a large number of participants is needed [194]. For that reason, nowadays, the number of existing public databases that can be used to evaluate the performance of multimodal biometric recognition systems is quite limited.

Motivated by such need we constructed a multimodal database, the *MobBIO Multimodal DB*, acquired using a portable handheld device, namely an Asus EeePad Transformer tablet. With this approach we aimed to tackle not only the ever growing need for data, but also to provide a database whose acquisition environment follows the rapid evolution of our networked society from simple communication devices to mobile personal computers. The database is composed by three modalities: iris, face and voice. A possible schematics of a multimodal system trained for the MobBIO database is presented on Figure 5.2.

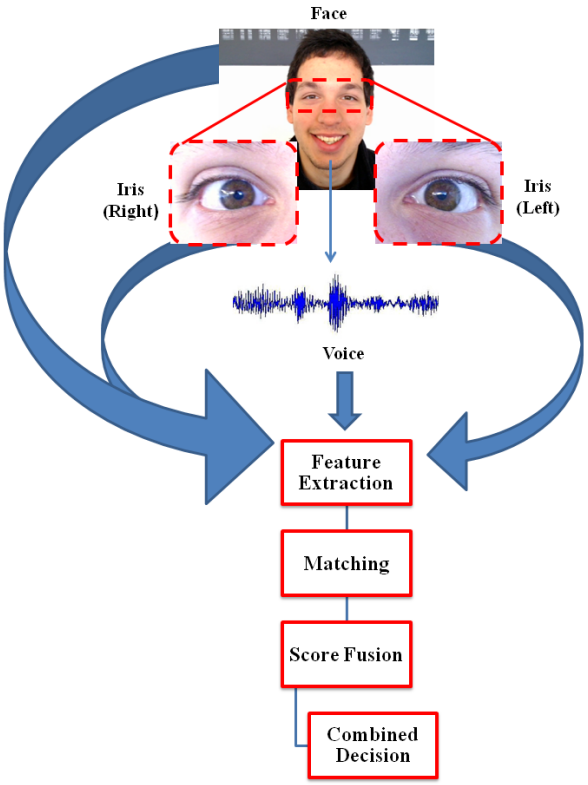


Figure 5.2: Flowchart of a generic multimodal system working on the modalities present in the MobBIO database.

The reasons to create the MobBIO multimodal database were related, on one hand, to the raising interest in mobile biometrics applications and, on the other hand, to the increasing interest in multimodal biometrics. These two perspectives motivated the creation of a database comprising face, iris and voice samples acquired in unconstrained conditions using a mobile device, whose specifications will be detailed in further sections. We also stress the fact that there was no multimodal database with similar characteristics, regarding both the traits and the unique acquisition conditions, at the date it was constructed.

As voice is the only acoustic-based biometric trait and the facial traits - face and iris - are the most instinctive regions for a mobile device wielder to photograph, we chose these three traits for the MobBIO database. In the choice of such traits it was also taken into account that the design of consumer mobile devices is extremely sensitive to cost, size, and power efficiency and that the integration of dedicated biometric devices is, thus, rendered less attractive [246]. However, the majority of the developed iris recognition systems rely on near-infrared (NIR) imaging rather than visible light (VL). This is due to the fact that fewer reflections from the cornea in NIR imaging result in maximized signal-to-noise ratio (SNR) in the sensor, thus improving the contrast of iris images and the robustness of the system [174]. As NIR illumination is not an acceptable alternative we obtain iris images with simple VL illumination, even though this results in considerably noisier images.

Mobile device cameras are known to present limitations due to their increasingly thin form factor. Therefore, these devices inherently lack high quality optics like zoom lenses and larger image sensors. Nevertheless, for most daily uses, the quality is considered "good enough" by most consumers [262]. Regarding acoustic measurements, no hardware improvements can solve the problems that harm the performance of voice recognition: environmental noise and voice alterations by external noises, such as emotional state or illness, need to be accounted for, by the algorithm [130]. Multimodal approach may help counter image-based difficulties, like low illumination or rotated images, with voice-based features or vice-versa. By exploring multiple sensors the intrinsic hardware-based limitations of each one can be balanced by the other, resulting in a synergistic effect in terms of biometric data quality.

5.3.1 Description of the database and its construction

The MobBIO Multimodal Database comprises the biometric data from 105 volunteers. Each individual provided samples of face, iris and voice. Most of the volunteers were portuguese but we also had volunteers from U.K., Romania and Iran. The average of ages was approximately 34, being the minimum age 18 and the maximum age 69. The gender distribution was 29% females and 71% males.

The volunteers were asked to sit, in two different spots of a room with natural and artificial sources of light, and then the face and eye region images were captured by sequential shots. The distance to the camera was variable ($10-50$ cm) depending on the type of image acquired: closer for the eye region images and farther away for face images. For the speech samples, the volunteers were asked to get close to the integrated microphone of the device and the recorder was activated

and deactivated by the collector. The equipment used for the samples acquisition was an Asus Transformer Pad TF 300T, with Android version 4.1.1. The device has two cameras one frontal and one back camera. The camera we used was the back camera, version TF300T-000128, with 8 MP of resolution and autofocus.

For the voice samples, the volunteers were asked to read 16 sentences in Portuguese. The collected samples had an average duration of 10 seconds. Half of the sentences presented the same content for every volunteer, while the remaining half was randomly chosen among a fixed number of possibilities. This was done to allow both the application of text-dependent and text-independent methods, which comprise the majority of the most common speaker recognition methodologies [70].

The iris images were captured in two different lighting conditions, with variable eye orientations and occlusion levels, so as to comprise a larger variability of unconstrained scenarios. For each volunteer 16 images (8 of each eye) were acquired. These images were obtained by cropping a single image comprising both eyes. Each cropped image was set to a 250×200 resolution. Some examples of iris images are depicted in Figure 5.3.

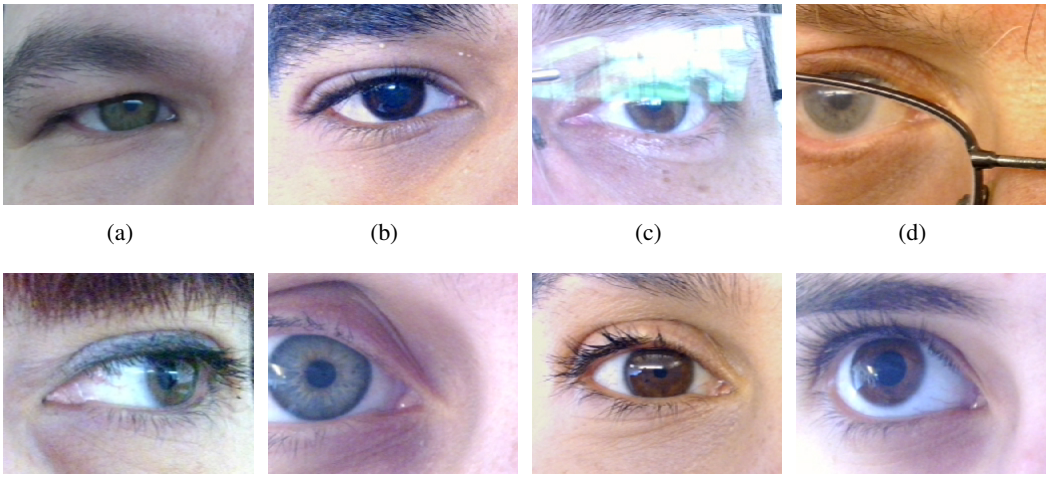


Figure 5.3: Examples of iris images from MobBIO database: a) Heavily occluded; b) Heavily pigmented; c) Glasses reflection; d) Glasses occlusion; e) Off-angle; f) Partial eye; g) Reflection occlusion and h) Normal.

The iris images can, by themselves, constitute an important tool of work concerning iris recognition in mobile devices environment. This dataset is provided with manual annotation of both the limbic and pupillary contours, enabling the evaluation of segmentation methods. An example of such annotation is shown in Figure 5.4.

Face images were captured in similar conditions as iris images, in two different lighting conditions. A total of 16 images were acquired from each volunteer, with a resolution of 640×480. Some examples are illustrated in Figure 5.5.

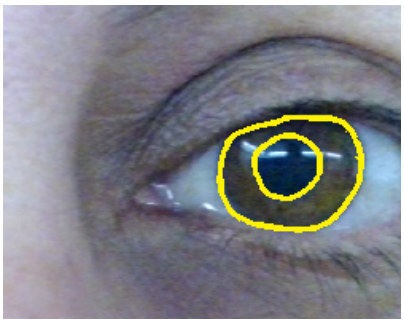
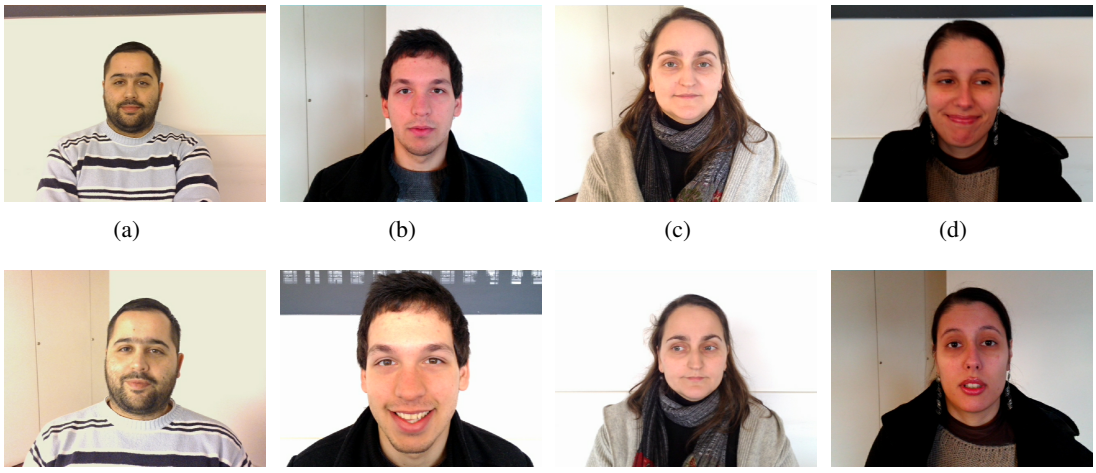


Figure 5.4: Example of a manually annotated iris image.



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

(e) (f) (g) (h)

Figure 5.5: Examples of face images from MobBIO database.

5.4 MobBIO 2013: 1st Biometric Recognition with Portable Devices Competition

Biometrics is one of the most studied subject for purposes of security and identification. The highlight of this competition was the application of biometric recognition, conditioned by the acquisition environment under which data is collected, to portable devices such as tablets and smartphones. Such a paradigm shift will involve the development of robust recognition algorithms capable of overcoming the new challenges posed by the introduction of less controlled acquisition conditions. In this competition, the participants were not limited to work on a specific trait but, instead, were allowed to work on one or multiple biometric modalities.

The MobBIO 2013 - 1st Biometric Recognition with Portable Devices Competition was embraced by ICIAR2013 - 10th International Conference on Image Analysis and Recognition, Póvoa de Varzim, Portugal, 2013. The organising committee was composed by Ana F. Sequeira, João C. Monteiro, Hélder P. Oliveira and Ana Rebelo. The main goal of the competition was to compare

different methodologies for biometric recognition and to assess the state of the art in a 'real-world' database and to foster further research in the field using data acquired with common portable devices and combining different biometric traits.

Each participant was invited to submit an algorithm in the form of an executable code. The algorithms could be developed either with respect to a single biometric trait from those provided in the training set or in a multimodal approach by fusing information from several traits. The benchmark dataset was the MobBIO Multimodal DB, described in section 5.3.1, comprising:

- Iris images, from both eyes, acquired under variable natural and artificial illumination conditions with user cooperation.
- Face images under the same assumptions as the ones presented above for iris.
- Voice recordings in noisy environments.

Each individual's data was stored under an identification (ID) value, so as to allow the evaluation of the participant's algorithms, as well as to maintain anonymous all the acquired data. The training dataset provided to the participants is composed by 4 images of each eye, 8 images of the face and 8 voice recordings per individual.
The submitted algorithms were evaluated with regard to two main functional modes of a biometric system: Verification mode (in which during test, the developed systems will receive a new input (either eye/face image(s) and/or sound file) and an identity claim and shall return as the output if the claim is either true or false); or Identification mode (in which the developed systems will receive a new input (either eye/face image(s) and/or sound file) and return as output the ID prediction for the tested input). The test dataset was composed by several triplets of files (one face; one eye and one voice recording) identified by the same filename. These triplets were separated by type in three different folders: "Face"; "Voice"; "Iris". Also a claim file was provided to be used in verification mode. On the test dataset, in identification mode, impostors (individuals that are absent from the training set) were introduced so as to test the vulnerability of the participant's algorithms to spoof attacks.
The submitted executable file was expected to:

- open the convenient folders: "Face"; "Voice"; "Iris" and read the files inside the folders;
- in verification mode: read the claims associated to the files read;
- return a file which for each triplet return "True/False", for verification mode, and an "ID-prediction", for identification mode.

To the best evaluated algorithm was attributed the "Best Biometric Recognition with Portable Devices Competition" award at ICIAR 2013.

5.5 Participants of MobBIO2013

The competition attracted several registered participants from different countries. Several participants registered for the competition as depicted in Table 5.1, however, in the submission step only two registered participants submitted their algorithms.

Table 5.1: Registered participants.

Team	Institution
Ilias Theodorakopoulos	University of Patras, Greece
SIM Group	Bethlehem University, Palestine
Jugurta Montalvão	Federal University of Sergipe, Brazil
Imago	Federal University of Paraná, Brazil
Andy Chen	Singapore
REISIER	Portugal
Rui Marques	Portugal
Alberto de Santos Sierra	Spain
Sumit Srivastava	India
Dhanesh Kumar Solanki	Experinn Research , India
Ajay Sundar Karuppasamy	Anna University, Chennai, India
Elsayed Hemayed	Faculty of Engineering, Cairo University, Egypt
Damon L. Woodard	Clemson University, School of Computing, USA
Nermin Kamal Abdel-Wahab	Cairo University, Egypt
Fernando Alonso-Fernandez	Halmstad University, Sweden
Ralph Corby	Smart Sensors Ltd, United Kingdom
Annalisa Franco	University of Bologna, Italy
António Manuel Ribeiro de Sousa	UTAD, Portugal
G.RAM SUNDAR	India
Juan-Carlos Perez-Cortes	ITI - Instituto Tecnologico de Informatica, Spain
Duan Fu, Zhang Jianli	University of Taiyuan, China
Vitor Yano	University of Campinas, Brazil
Gayathri Mahalingam	Dept. of Computer Science, UNC Wilmington, USA
M. Swaminathan, J. Karthikeyan	M.E. Biometrics and Cyber Security, India
Praveen P.	Vellore Institute of Technology, India
Huang Dong	Institute of High Performance Computing, Singapore
Jawed Akhtar Unar	University of Malaya, Malaysia

5.6 Performance Evaluation

The results were evaluated by the Equal error rate (EER), in verification mode, and False identification rate (FIR), in identification mode.

- **EER** is given by the value for which the False Acceptance Rate and the False Rejection Rate are equal. The False Acceptance Rate (FAR) and the False Rejection Rate (FRR) are as defined in Chapter 2,

- **FIR** is given by the ratio between the number of wrong recognition attempts and the total number of recognition attempts, as defined in Chapter 2.

5.7 Results of MobBIO2013

The final results of the competition are presented in Table 5.2.

Table 5.2: Participants.

Rank	Team	FAR	FRR	Average
1	Ilias Theodorakopoulos, Greece	56.48	4.52	30.50
2	SIM Group, Palestine	81.43	47.50	64.47

The winner of MobBIO2013 was Ilias Theodorakopoulos from the University of Patras, Greece.

5.8 Final remarks on the MobBIO Multimodal DB and the *MobBIO 2013* Competition

The increased use of handheld devices in everyday activities, which incorporate high performance cameras and sound recording components, has created the possibility for implementing image and sound processing applications for identity verification. The aim to produce reliable methods of identifying and authenticating individuals in portable devices is of utterly importance nowadays. The research in this field requires the availability of databases that resemble the unconstrained conditions of these scenarios.

We aimed to contribute to the research in this area by deploying a multimodal database whose characteristics are valuable to the development of state-of-the-art methods in multimodal recognition. The manual annotation of iris images is a strong point of this database as it allows the evaluation of segmentation methods with this noisy images. We hope that in the future, the other samples will also be annotated manually: the face will be identified in face images and the silence and speech will be identified in the sound recordings. It might be argued that the use of this dataset may still be somewhat limited in the research community because of its characteristics. It would be better if the voice samples were recorded both in English as well as Portuguese, and the images stored in several resolutions and even more challenging real-life conditions, such as variable environments (indoors and outdoors) and illuminations. This set of suggestions will surely be taken into consideration for future improvements over the present dataset.

Through the competition this database has been spread among the scientific community and was a valuable resource for the research in the mobile multimodal biometrics field. However, we note that its usefulness went far beyond its immediate application in the competition. On one hand, it has been requested by other researchers afterwards. On the other hand, the whole database or its subsets have been tested in several works in the multimodal, face, iris or periocular recognition area [16, 17, 170, 173, 175] and the iris dataset was used in works concerning the iris segmentation

problem [174, 175]. Also, the iris image collection has allowed the construction of an iris dataset with fake images, the MobBIO*fake*, composed by printed copies and their respective originals. This database was developed for the purpose of iris liveness detection research as described in detail in section 6.2.

The work described in this chapter focused on the multimodal mobile biometrics scenario. This database was constructed having in sight the aim to acquire biometric data with an handheld device which would then allow the evaluation of methods adapted to this more recent context of application of biometric recognition. The traits chosen were face, iris and voice due to their complementarity and the adaptability to mobile devices (which are equiped with cameras and sound recording systems).

However, it should be noted that the main focus of the present thesis is not the multimodal scenario. The construction of this database was a first approach to the mobile biometrics problem. The iris subset of the MobBIO Multimodal DB represented an important tool for research regarding iris recognition in the mobile scenario. The iris images by itself allowed the evaluation of methods for segmentation and recognition more adjusted to the limitations of images captured with this type of devices. Also, allowed the construction of an iris database with printed fake samples for iris liveness detection captured with an handheld device which at its time was a novelty.

Even though the fingerprint was not included in the multimodal database constructed, the following work in this thesis will include research in liveness detection and in recognition for fingerprint recognition systems. One motivation for the posterior work with fingerprint was the fact that this biometric trait has recently been introduced in mobile devices for authentication. Nowadays there are still not fingerprint databases of images acquired with these devices available for research but we believe that in a near future this research necessity will be suppressed.

Chapter 6

Iris Liveness Detection Database and Competition*

As already mentioned, the existence of suitable databases for the development, evaluation and validation of methods in the biometrics field is of utmost importance. Regarding iris recognition we sensed a lack of databases with images acquired with handheld devices and under less controlled imaging conditions. And consequently the same occurred in the iris liveness detection field. We observe that the Mobile Biometrics Field is gaining more and more importance since our mobile devices are transformed in repositories of personal and professional data more than just communications goaled devices. The spread of handheld devices equiped with cameras and has encouraged the application of iris biometric recognition in mobile devices. Along with this phenomena, there is a necessity of adjusting the biometric systems to this new acquiring scenario and its proper limitations, and, above all, to the security issues that are raised. Therefore, the liveness detection techniques become of crucial importance and to develop and evaluate new methods there is a necessity for databases built under these conditions. In this chapter we present a database acquired with an handheld device. This database was build upon the set of iris images of the **MobBIO Multimodal DB**, presented in the previous chapter, by adding fake samples to these images. The purpose of this database is to be used in liveness detection, in particular, to evaluate the presentation attack in which printed images are presented to the sensor.

In this chapter we present the construction of the mentioned biometric database, **MobBIO*fake***, and the organization of the biometric competition which used this database as benchmark, **MobILive 2014**. We also show the impact of this event, to the present date, referring the interest that the database continues to raise in the scientific community and the works that have used it so far up to the time of writing.

*Some portions of this Chapter appeared in [235, 237]

6.1 Iris Liveness Detection

The evolution in the use of mobile devices in our society and the growing accessibility of cameras in handheld devices encourages iris biometric applications in daily life. It is known that iris has been traditionally regarded as one of the most reliable and accurate traits among the different biometric traits. However, researchers were motivated to explore its vulnerabilities and to measure how spoofing attacks may compromise the security of iris based recognition systems. These systems may be vulnerable to attacks which consist on the presentation of a fake iris to the sensor pretending to be one of a legitimate user. A spoofing attack may be perpetrated by the use of contact lenses or the presentation of high quality iris printed images. This vulnerabilities raise an urge for liveness solutions in the **mobile biometric field** and along with this comes the need for **suitable databases** in which new methods can be tested.

6.2 MobBIO*fake*: an Iris Database with Printed Fake Samples

The MobBIO*fake* database was constructed upon the MobBIO Multimodal DB [234] and is composed by a subset of 800 iris images and 800 fake copies. The fake samples were obtained from printed images of original ones captured with the same handheld device and in similar conditions.

The aim of constructing such a database was, on one hand, to fulfill the necessity of databases and, on the other hand to broad the acquisition conditions of the images. The number and variety of databases for iris liveness detection is somewhat limited so the fact that these images were captured with a portable device and are *RGB* images came as a novelty and made it possible to evaluate liveness methods in this new upcoming scenario. After the construction of MobBIO*fake* other databases were made available with images acquired by smartphones but at its time MobBIO*fake*'s characteristics came as a novelty.

6.2.1 Description and construction of the database

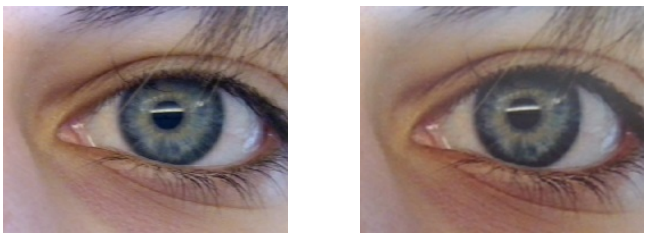
The MobBIO*fake* is composed by a subset of 800 iris images from MobBIO and 800 fake copies, in a total of 1600 iris images. The construction of the MobBIO*fake* upon the MobBIO Multimodal DB iris images subset comprised several steps. The images of each volunteer were joined in a single image, as shown in Figure 6.1.

A preprocessing (contrast enhancement) was applied using GIMP software [90] to the image. This enhancement is believed to improve the quality of the fake sample [224]. After this, the images were printed in a professional printer using high quality photographic paper. At this point we were ready to capture the images. Each individual image (a image of one single eye) was acquired using the same portable device and in similar lighting conditions as the original ones were captured, as illustrated in Figure 6.2.

Finally, the individual eye images were cropped and resized to predefined dimensions. An example of a real image and its copy is depicted in Figure 6.3.

Figure 6.1: *MobBIOfake* construction: joint images of one volunteer.

Figure 6.2: *MobBIOfake* construction: fake samples acquisition.



(a) Real image (b) Fake image

Figure 6.3: Corresponding real and fake images of *MobBIO*.

6.3 MobILive 2014 – 1st Mobile Iris Liveness Detection Competition

Biometric systems based on iris are vulnerable to several attacks, particularly direct attacks consisting on the presentation of a fake iris to the sensor. The development of iris liveness detection techniques is crucial for the deployment of iris biometric applications in daily life specially in the mobile biometric field. The *MobILive2014* - 1st Mobile Iris Liveness Detection Competition

[1] - was organized in order to record advances in iris liveness detection and took place between December, 2013 and April, 2014.
MobILive2014 was embraced by the IJCB2014 - IEEE Int. Joint Conference on Biometrics (IJCB), Clearwater, Florida, USA, 30 September – 02 October, 2014[2] and its organising committee was Ana F. Sequeira, João C. Monteiro, Hélder P. Oliveira and Jaime S. Cardoso. The challenge launched by this competition was to perform iris liveness detection in mobile applications as an anti-spoofing measure (a step prior to the iris recognition process). The goal of *MobILive* was to contribute to the state of the art of this particular subject and to become a reference event for academic and industrial researches. Providing a public access database of real and fake iris images (captured in cooperative yet unconstrained conditions using a handheld device) represented an important contribution of the competition to the scientific community.
This competition covered the most common and simple spoofing attack in which printed images from an authorized user are presented to the sensor by a non-authorized user in order to obtain access. The benchmark dataset was the MobBIO*fake* database which is composed by a set of 800 iris images and its corresponding fake copies (obtained from printed images of the original ones captured with the same handheld device and in similar conditions).
The benchmark data was divided in train and test sets each one comprising real and fake images from 50 different individuals each. The train dataset was provided to the participants. The test set was used by the organizers to perform the evaluation of the methods.
MobILive included an intermediate submissions period during which the participants submitted an executable file. The result of its evaluation (performed by the organizers) was published and updated after each new submission by evaluating the algorithms in the same randomly obtained subset of the test set composed by 200 images. These intermediate submissions were meant to stimulate interaction with the event, to give feedback to the participants about their performance compared with the other participants and to allow the refinement of the algorithms.
The final results were obtained by the evaluation of the final submission which was the last submission of each participant on the entire test dataset. The winner of the competition was the first classified in this final evaluation.

6.4 Participants

Ten participants registered in the competition from several countries such Brazil, Hong Kong, India, Italy, Norway, Sweden, United Kingdom, United States of America. Among the registered participants, six teams submitted their algorithms. In Table 6.1 we list the registered participants and indicate the teams which submitted methods to the competition.

[1] http://mobilive2014.inescporto.pt/
[2] http://ijcb2014.org/

Team	Institution	Submitted Method
Federico II	University Federico II of Naples, Italy	✔
GUC	Gjϕvik University College, Norway	✔
HH	Halmstad University, Sweden	✔
IIT Indore	Indian Institute of Technology, India	✔
IrisKent	University of Kent, United Kingdom	✔
UNICAMP	Un. Campinas, F.U. Ouro Preto, Brazil	✔
Jinyu Zuo	USA	
Rahul Soni	USA	
Tan Chun Wei	Hong Kong Polytechnic University	
Zahid Akhtar	University of Udine, Italy	

Table 6.1: Registered participants.

6.5 Briefs of the submitted methods[1]

Federico II

Submitted by D. Gragnaniello, C. Sansone and L. Verdoliva from DIETI, University Federico II of Naples, Italy. This approach is based on the use of local descriptors, which are powerful tools to describe the statistical behavior observed locally in small patches of the image. Among the best known we can count the Local Binary Pattern (LBP) [191], successfully used for different tasks, like texture and face recognition. These patterns are able to detect micro structures whose underlying distribution is estimated by histograms collected over the ensemble of all patches. In order to increase the discriminative power of LBP, in [185] the authors propose to evaluate the histograms of the co-occurrence among these micro patterns. In this way, it is possible to better represent complex patterns and capture the spatial relations in the image. Hence in this work, the authors consider the features as proposed in [185], but the main difference is that they are evaluated on the prediction-error image (also called residual image). In fact, modeling the residuals rather than the pixel values is very sensible in these low-level methods (not based on image semantic), since the image content typically does not help detecting local alterations. This consideration is especially true for the problem of iris liveness detection, when printed iris are presented to the sensor, in order to detect seemingly invisible alterations of the natural characteristics of the biometric trait. In particular, the features are extracted from the prediction-error images evaluated as proposed in [287]. Finally, a Support Vector Machine (SVM) with linear kernel was used as classifier and a leave-one-out cross-validation to find the best value of its parameters. The proposed approach can be summarized in the following steps: 1) computation of the high-pass residuals; 2) feature extraction based on co-occurrence of adjacent LBP; 3) SVM classifier with linear kernel.

[1] Presented as given by the authors.

GUC

Submitted by R. Raghavendra, Kiran B. Raja and Christoph Busch from Gjϕvik University College, Norway. GUC's Presentation Attack Detection (PAD) (or spoof detection or counter measure) Algorithm for visible iris attack detection is based on both local and global statistical features. The pool of statistical features are based on various image quality measures that captures variation at pixel level and also at the block level to reflect the rich information to identify the presence of visible iris artefact. Further, the weighted multi-classifier fusion of these features are carried to make the final decision. The weights are optimized on the training subset of MobBIO*fake*.

HH

Submitted by Fernando Alonso-Fernandez and Josef Bigun from Halmstad University, Sweden. The fake iris detection system is based on Gray-Level Co-Occurrence textural features [43,97,250] extracted from the three color (RGB) channels of the image. This method looks for the best features by Sequential Forward Floating Selection (SFFS) [209], using SVM as classifier [269]. Given n features to combine, it employs as criterion value of the SFFS algorithm the HTER (*Half Total Error Rate*) of the corresponding classifier trained with the n features. This method also localize the eye center position, which is used as input of the Gray-Level Co-Occurrence Matrix (GLCM) feature extraction algorithm, so as to extract GLCM features in the desired region of the image only. For this purpose, the authors employ their eye detection algorithm based on symmetry filters [16].

IIT Indore

Submitted by Vivek Kanhangad, Pragalbh Garg and Pranjalya Singh from the Indian Institute of Technology Indore. The IIT Indore algorithm for ILD is based on the analysis of differences in texture patterns for discriminating between real and fake iris images. Specifically, the approach is based on feature level combination of the following three texture descriptors: Local Phase Quantization [192]; Binary Gabor Pattern [305]; Local Binary Pattern [191]. The combined feature set resulting from the feature level fusion of the above descriptors is then used to train a support vector machine (SVM) classifier with linear kernel.

IrisKent

Submitted by Yang Hu, Konstantinos Sirlantzis and Gareth Howells, from University of Kent, UK. This method exploits the combination of multiple features for ILD in mobile applications. Firstly, some base level features are extracted. The base level features are then fed to a score level combiner and a feature level combiner. The final decision is made based on the output of the two combiners. In score level combiner, each base level feature is fed to a classifier. The response of each classifier is used as scores and combined to produce an output. In feature level combiner,

different selected features are combined to form a new feature vector. This feature vector is fed to a classifier whose score is used as output. The base level features are extracted using a spatial pyramid structure [135, 298]. The spatial pyramid structure partitions image into increasing finer sub-regions. Local features within each sub-region are pooled together, and the pooled features of each sub-region are concatenated to form base level features. The spatial pyramid structure captures the local and global distributions of the features. It gives us information of the iris region as well as the distributions around the iris. In the algorithm, 9 local features are extracted to form 9 base level features: sparse coding on SIFT features [148, 298], sparse coding on Histograms of Oriented Gradients (HOG) features, local binary patterns, red channel histogram, red channel correlogram, color histogram, intra-color correlogram, inter-color correlogram and multi-color correlogram [40]. Sparse coding on SIFT features encodes the SIFT descriptor by sparse coding as a local feature. Similarly, sparse coding on HOG feature encodes HOG features by sparse coding. Additionally, histogram and correlogram are computed on all RGB color channels. Correlogram is a pairwise intensity distribution. It reveals color and texture distribution of image. Intra-color correlogram concatenates the correlogram at each color channel. Inter-color correlogram concatenates the correlogram between each two color channels. It reveals the color and texture distribution across the color channels. Multi-color correlogram is the combination of intra-color correlogram and inter-color correlogram. We perform feature selection for score level combiner. A feature is selected if it can either improve the accuracy of ILD or preserve the accuracy and enlarge the gap between real and fake images. In feature level combiner, considering computational cost, we simply select the features with top performance. In the experiments, LBP and intra-color correlogram are selected for the score level combiner, while LBP and multi-color correlogram are selected for the feature level combiner.

LIV-IC-UNICAMP

Submitted by D. Menotti[1,2], G. Chiachia[1] and A. X. Falcão[1] from (1) University of Campinas and (2) Federal University of Ouro Preto. A key characteristic of the UNICAMP system is the use of a special type of convolutional neural networks (CNNs) for feature extraction. These networks are inspired in recent work on biologically-inspired computer vision [200] and have two important properties: (i) *optimal* architecture and (ii) *random* filter weights. The first property has shown to be of crucial importance for CNNs [23, 200] and the second one allows for the construction of robust, highly nonlinear feature extractors even from datasets with few training samples, while also performing surprisingly well [44, 116]. In the attempt to discover optimal CNN architectures, the number of layers and the operations considered by the authors are the same as in [200], as well as the optimization procedure, which consists of randomly sampling and evaluating thousands of candidate CNNs to choosing the best one. In order to evaluate these CNNs, the authors further divided the images made available into training and test sets such that they were disjoint in terms of person identity. In fact, the UNICAMP method can be viewed as the combination of three subsystems, each one containing one CNN for feature extraction and one linear SVM (operating on top of these features) to predict whether iris images are fake or real. The difference among

these subsystems is their corresponding CNN, which were found to perform best when allowed to output feature vectors of size in the intervals $[200, 5000]$, $[5000, 10000]$ and $[10000, 20000]$. While each of these subsystems perform quite well by their own, the authors found that combining them by majority voting led to superior performance. In addition, six (out 800) samples that were most frequently mistaken by the subsystems (*i.e.*, supposedly outliers) were removed while training the linear SVMs for the final submission.

6.6 Performance Evaluation

The metrics used to evaluate the results are *False Acceptance Rate (FAR)*, *False Rejection Rate (FRR)* and *Mean Error Rate (MER)* (which is the mean of FAR and FRR), as defined in Chapter 2.

6.7 Discussion and Results

As referred, the competition comprised a period for intermediate submissions (a maximum of four) during which every new submission was evaluated with a randomly obtained subset of 200 images from the test set. In Table 6.2 are shown the rankings for the three intermediate submissions.

Rank	1^{st}	2^{nd}	3^{rd}
1	FedericoII	FedericoII	UNICAMP / IITIndore
2	GUC	GUC	GUC / Federico II
3	IITIndore	IrisKent	IrisKent
4	IrisKent	UNICAMP	HH
5		IITIndore	

Table 6.2: Ranking of the algorithms' performance in $1^{st}, 2^{nd}, 3^{rd}$ intermediate submissions.

We remark that the intermediate submissions and the publication of the ranking motivated the healthy competition among the participants. Through the publication of updated rankings the participants could assess the performance of their methods relatively to the others and we believed that the competition benefited form this synergy between the participants. The final results were obtained in the test dataset composed by 800 images. The final version of the algorithms was the last one received till 17^{th} of March, 2014. In Table 6.3 the final results are presented for the six teams in competition.

The best result was obtained by the *IIT Indore* team with a MER of 0.25%, closely followed by the *GUC* and *FedericoII* teams. In general, the results obtained were very good and improved the published results for this database.

Rank	Team	No.	FAR	FRR	MER
1	IIT Indore	2	0.00	0.50	0.25
2	GUC	2	0.75	0.00	0.38
3	Federico II	1	1.25	0.00	0.63
4	UNICAMP	3	0.50	2.00	1.25
5	IrisKent	4	0.25	3.75	2.00
6	HH	2	29.25	7.00	18.13

Table 6.3: Final results presenting the FAR, FRR and MER (in %). "No." refers to the number of submissions.

6.8 Further analysis of results

The results obtained, by most of the participants, were far better than what was the state-of-the-art with this database, around 12% of MER [235]. Considering these results, four of them below 2%, we decided to analyze the characteristics of our images and test the robustness of the methods to a more "clever" manipulation of images.

Some intuitive research was performed and one of the attempts made was to convert the images to the *CIELab* color space and analyze the variation of the values of the three channels. Particularly, the values of the L and the b channels allowed to determine a strong separation between the two classes of images being possible to define a threshold that would separate the real and fake images quite well. Putting ourselves in the role of a malignant agent, in order to be well succeeded in the spoofing attack, we would try to make our fake images resemble the most to the real ones. Therefore, we manipulated some fake images (approximately 100 images, i.e., 12.5% of the total of fake images) so that the values of these two channels were distributed similarly to the ones of the real images. In Table 6.4 we present the results obtained by evaluating the algorithms in this dataset. We must state that these results may be considered invidious since the algorithms were not trained with these images.

Rank	Team	No.	FAR	FRR	MER
1	IIT Indore	2	0.00	0.50	0.25
2	GUC	2	0.75	0.00	0.38
3	UNICAMP	3	0.50	2.01	1.26
4	IrisKent	4	5.75	3.75	4.75
5	Federico II	1	16.25	0.00	8.13
6	HH	2	29.25	7.00	18.13

Table 6.4: Results of the evaluation in the manipulated test set (in %). "No." refers to the number of submissions.

There are some methods clearly more robust to these changes than others. The methods of *FedericoII* and *IrisKent* teams appear to be more sensible to these changes. The four remaining methods are not affected by the changes in the images.

Other similar changes could be performed which would increase the feasibility of the spoofing, however this kind of manipulation requires from the intruder to possess privileged knowledge and ease of access to the database storage.

6.9 Final remarks on the MobBIO*fake* DB and the *MobILive2013* Competition

We believe that the deployment of iris biometric applications in daily life, particularly in the mobile biometric field, has created a necessity for ILD solutions. The 1st *MobILive* Competition was organized having as main goals the possibility to record recent advances in ILD and stimulate new ones. In our view, the objectives were accomplished, considering that excellent results were achieved, exceeding the state-of-the-art results, by participants from all over the globe. However, the results obtained also encouraged us to go further. One very important aspect in this field of research is the necessity of more public available datasets with more variety of acquiring scenarios. We made our contribution and we expect to have motivated the appearance of new, more challenging, databases for ILD.

Since the competition, this database has continued to capture the attention of researchers all over the world. To the date, it has been requested by researchers from Australia, Brazil, India, Iran, Ireland and United Kingdom. Another important aspect of evidence of its usefulness is the fact that, beyond the competition, this database has continued to be used to evaluate iris liveness detection methods. MobBIO*fake* is used as benchmark in several published papers [15, 170, 212, 214, 236] (references known till may of 2015) and in some others still to appear.

Part III

Liveness Detection and Robust Recognition

Chapter 7

Iris Liveness Detection Methods in Mobile Biometrics Scenarios*

In this chapter it is presented the proposed iris liveness detection methodology. The methodology combines several feature extraction methods proposed for iris liveness detection, passes them through a feature selection process and feeds different classifiers with the best subset of features. This methodology was applied in two different situations: in the first study the iris segmentation step is manual and in the second we performed an automatic segmentation.

7.1 Iris Liveness Detection Methodology Overview

The problem of liveness detection of a biometric trait can be seen as a two class classification problem where an input trait sample has to be assigned to one of two classes: real or fake. The key points of the process are to find a set of discriminant features and then build an appropriate classifier which determines the probability of the sample vitality given the extracted set of features [83].

We propose to address the task with a methodolgy composed by a cascade of four main operations, as depicted in Figure 7.1.

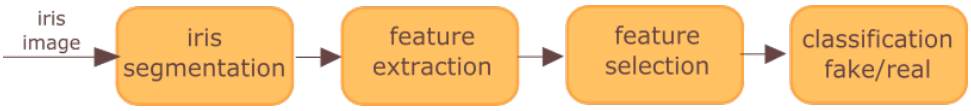


Figure 7.1: Steps of the proposed method

The first step of the method is the *segmentation*. This step is required since some of the feature extraction methods are applied only to the iris region or to the bounding box containing the iris and pupil. The methodology included whether a manually or an automated segmentation method. The advantage of the manual segmentation is that the results are not influenced by segmentation errors. However, if we aim to explore the applicability of these methods in real world applications, the

*Some portions of this Chapter appeared in [235, 236]

automatic segmentation is a necessary step to be included in process and the errors it may produce must be dealt with.

The second step of the methodology is the *feature extraction*. This stage comprises the application of the methods described in subsection 7.3.

After the feature extraction it is necessary to perform the *feature selection*. This comprises the application of the method *Sequential Forward Floating Search*, described in subsection 7.4, before applying the classifiers to evaluate the proposed methods.

The last step is the *classification* described in section 7.5. We used three state-of-the-art classifiers: Discriminant Analysis (DA), k-Nearest Neighbours (kNN) and Support Vector Machines (SVM).

7.2 Segmentation

The first step of the proposed methodology is the *segmentation of the iris region*. This step is required since some of the feature extraction methods are applied only to the iris region or to the bounding box containing the iris and pupil. We evaluated the method proposed both performing manual segmentation and applying an automated method.

Manual Segmentation

In a first approach we choose to make the segmentation process manually, since the image segmentation step is always a troublesome operation, in order to evaluate the performance of the succeding operations without being impacted by the always challenging task of segmenting iris.

The manual segmentation is done by marking three different points in the image. The first point is the eye center, i.e., we consider a single center for both pupil and iris. The second point is marked in the pupil border and the third in the iris border. With these points is possible to determine the iris and pupil radius and then approximate the contours as two concentric circles. With the manual segmentation's information we are able to map the regions of interest which will be eventually used by the liveness detection algorithms. We note that, for the sake of simplicity, this manual segmentation makes two assumptions which are not accurate, in fact we know that the center of iris and pupil are not necessarily the same and the limbic and pupillary contours are not necessarily circular.

Robust segmentation under unconstrained scenarios

The automated *segmentation* algorithm [175] performs automatic detection of the inner and outer contours of the iris. The method focuses on mutual context information from iris center and iris limbic and pupillary contours to perform robust and accurate iris segmentation in noisy images.

The simultaneous detection of the iris center and limbic contour are addressed by first over-detecting center candidates, followed by a contour detection around each of them. The center candidates are estimated using a convergence index filter methodology [131]. Next, a window

centered in each candidate is converted into the polar domain followed by shortest path algorithm to determine good closed paths around the center. Using combined data from the center and respective contour, the best pair center/contour is selected. Finally, the pupillary segmentation is performed using a new polar image around the centroid of the detected limbic contour. For more details on the method see Annex A.

7.3 Feature Extraction Methods

As refered, some of the measures are obtained from the entire eye image but others are extracted only from the iris region.

Algorithm 1 - High Frequency Power

The *High Frequency Power* algorithm, which provides feature 1, works on the whole image and measures the energy concentration in the high frequency components of the spectrum using a high pass convolution kernel of 8×8. The application of this convolution is a good Fourier Transform approximation and works as high frequency spectral analysis, which can be considered an estimator of focus [56]. The focus of a real iris, as it is a 3D volume, is different from a fake iris focus, which has a 2D surface. For more details on the method see [83].

The final measure (feature 1) is given by equation 7.1, where I_c is the convolution matrix $M \times N$.

$$Feature1: \quad r = \frac{1}{M \times N} \sum_{i=1}^{N} \sum_{j=1}^{M} I_c(i,j) \tag{7.1}$$

Algorithm 2 - Local Contrast

The *Local Contrast* algorithm, which provides feature 2, is applied on the bounding box that involves the iris and the pupil. The bounding box is divided in blocks of $P \times P$ and for each block the *Fast Fourier Transform* (FFT) algorithmit is applied to extract the medium power frequencies, which better represents the contrast. The final value is given by the number of blocks with medium values (between 20 and 60) divided by the total number of blocks. This algorithm was inspired in an occlusion estimation technique [9] and it was adapted for contrast estimation for iris liveness detection in [83]. The value of P was adapted to each database according to the image size.

Algorithm 3 - Global Contrast

The *Global Contrast* algorithm, which provides feature 3, explores the fact that parts extremely bright or dark of the image are not useful and can be considered as noise. Thus, pixels near medium value (128 in 8-bit image) are considered of best contrast [9]. In order to quantify the contrast, the original pixel's values are normalized between 0 and 25, according to the normalization function depicted in Figure 7.2. Original pixels near medium value will get higher values in the normalized

scale, while very low and very high values (< 10 and > 245) are normalized to 0. This measure was presented in [9] and it was adapted for global contrast estimation for iris liveness detection in [83].

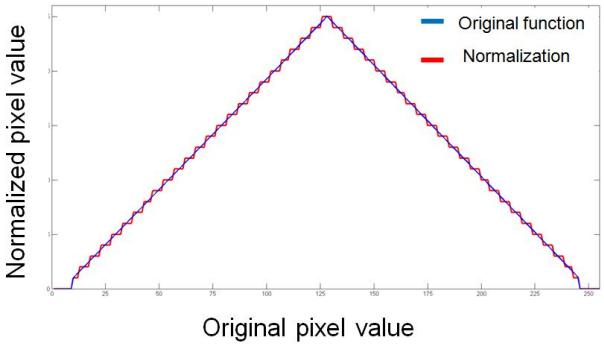


Figure 7.2: Normalization function for the algorithm 3.

Algorithm 4 - Frequency Distribution Rates

The *Frequency Distribution Rates* algorithm consists in different mathematical combinations of three different parameters which consider respectively the power of the low (F_1), medium (F_2), and high (F_3) frequencies (computed according to the 2D Fourier Spectrum) from two iris subregions in the horizontal direction. This subregions are illustrated in Figure 7.3.

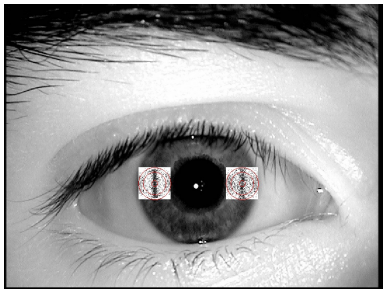


Figure 7.3: Example of the subregions used in the algorithm 4.

Each subregion is subdivided in three circular concentric regions, which determine the three different frequencies, i.e, for the first subregion, F_1^1 refers to the central circle, F_2^1 refers to the middle circular ring and F_3^1 refers to the outer circular ring, as depicted in Figure 7.4. The final F_1 is given by the average between the two regions: $F_1 = \frac{F_1^1 + F_1^2}{2}$. The same is done to F_2 and F_3. More details on the method can be found in [83, 153].

With the three final frequencies we extract seven different combinations [83, 153], represented in Table 7.1.

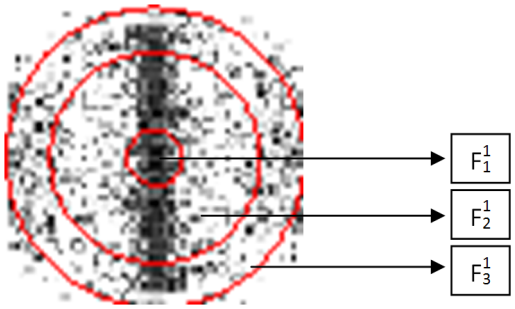


Figure 7.4: One of the regions of interest subdivided to calculate the frequencies.

Table 7.1: Extracted measures from the final frequencies.

Features no.	Combination
4	$F_1 + F_2 + F_3$
5	$F_2/(F_1 + F_3)$
6	F_3
7	F_2
8	F_1
9	$(F_1 + F_2)/F_3$
10	$(F_1 \times F_2)/F_3$

Adaptation of Algorithm 4 to non-circular contours (Adapted Frequency Distribution Rates Algorithm)

The algorithm 4 had to be adapted to the non-circular contours detected by the automated segmentation method used.

The *Adapted Frequency Distribution Rates* algorithm consists in different mathematical combinations of three different parameters which consider respectively the power of the low (F_1), medium (F_2), and high (F_3) frequencies (computed according to the 2D Fourier Spectrum) from two iris subregions. In the original method the two regions were determined by the circular contours of iris and placed in the horizontal direction, as depicted in previously presented Figure 7.3.

Considering that the contours obtained using the automated segmentation method are not circular and much less concentric as depicted in Figure 7.5, then the two regions determined by limbic and pupillary contours can not be determined as previously showed. It was necessary to adjust the estimation of the position of these two regions giving the new scenario.

One thing to take in account was the necessity of having a portion of the iris in the less occluded and noisy zone and also with the biggest area possible to help to enhance the quality of the features extracted. So, with the contours obtained in the segmentation step the iris mask was determined as well as its limits (as ilustrated in the first two images of Figure 7.6) and then the candidates to centers of the two regions were determined (as ilustrated in the third image of Figure 7.6). Therefore, we searched for the position of the centre which lead to the maximum possible radius,

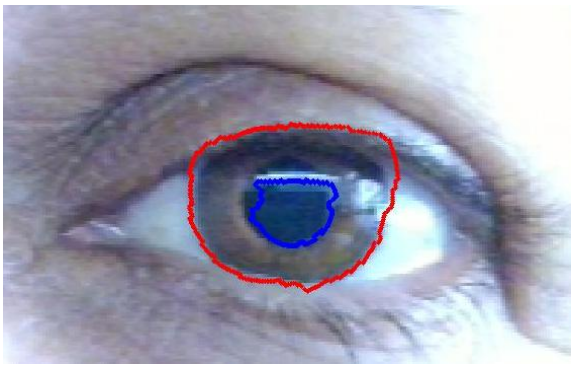


Figure 7.5: Non-circular contours obtained with the automated segmentation method.

starting on the position corresponding to an angle of 225° in the left side and to an angle of 315° in the rigt side, as illustrated in Figure 7.6, and searching in the intervals $[180°, 270°]$ and $[270°, 360°]$, respectively, as illustrated in Figure 7.6.

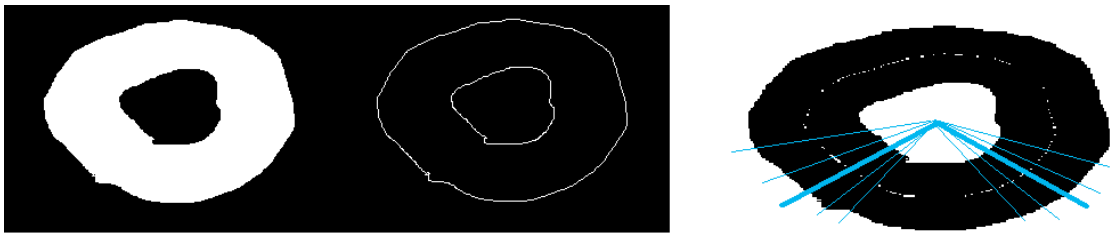
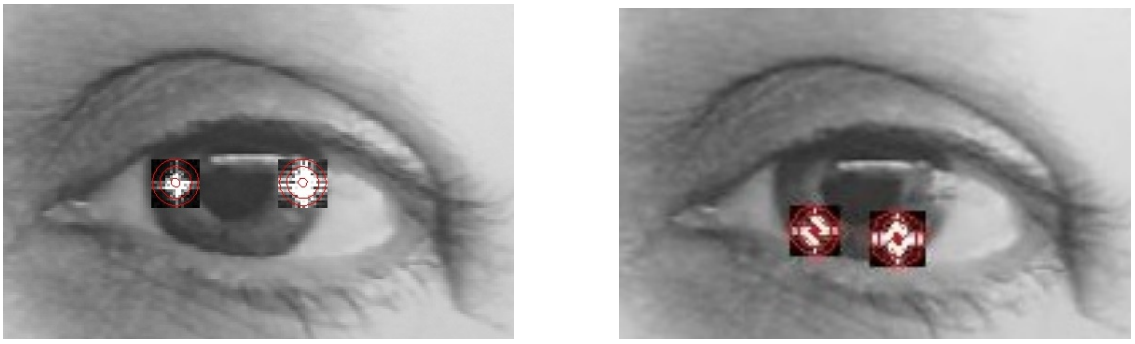


Figure 7.6: Mask and non-circular contours and the candidates to centers of the regions for feature extraction with Adapted Algorithm 4.

Therefore, in this adaptation of the algorithm, for the same image, the position and radius of the two regions may differ from the ones obtained in the manual segmentation version due to the non-circular iris contours provided by the automated segmentation process. This is illustrated in Figure 7.7.



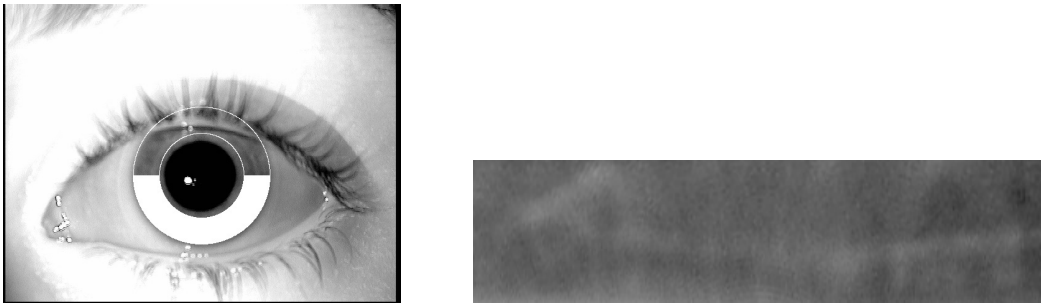
(a) Original (b) Adapted

Figure 7.7: Regions of interest used in the original algorithm and in the adapted version of Algorithm 4.

Once the two subregions are determined the extraction of features is done in the same way as in the original version.

Algorithm 5 - Statistical Texture Analysis

The *Statistical Texture Analysis* algorithm was developed as a contact lens countermeasure. The outer portion of the color contact lens (corresponding to regions closer to the outer circle) provides the most useful texture information for fake iris detection since this section of the fake iris is insensitive to the pupil dilation [99]. The region of interest is the lower part of the iris in order to minimize the occlusion by the eyelashes and eyelids, which in general occurs in the upper iris portion. In order to achieve invariance to translation and scale, the region of interest is further normalized to a rectangular block of a fixed size $W \times H$ as can be seen in Figure 7.8. (The values of W and H were adapted to each database according to the image size.)



(a) Original (b) Normalized

Figure 7.8: Region of interest used in the algoritm 5 (from [99]).

After the normalization, the Gray Level Co-occurence Matrix (GLCM) is calculated. This is one of the most proeminent approaches used to extract textural features [97]. Four measures are extracted: the mean (μ) and the standard deviation (σ), are extracted directly from the normalized region of interest, and the contrast (con) and the energy (e), are extracted from the GLCM matrix. These measures will provide features 11 to 14 and their values are given, respectively, by the equations below:

$$\mu = \frac{1}{W \times H} \sum_{i=1}^{H} \sum_{j=1}^{W} I(i,j) \tag{7.2}$$

$$\sigma = \sqrt{\frac{1}{W \times H} \sum_{i=1}^{H} \sum_{j=1}^{W} (I(i,j) - \mu)^2} \tag{7.3}$$

$$con = \sum_{i=1}^{N} \sum_{j=1}^{N} (i-j)^2 P(i,j) \tag{7.4}$$

$$e = \sum_{i=1}^{N} \sum_{j=1}^{N} P(i,j)^2 \tag{7.5}$$

Methods not requiring contour detection In these equations, I denotes the normalized iris image, W is the width of the normalized iris image, H is the height of the normalized iris image, P is the co-occurrence matrix and N denotes the dimension of the co-occurrence matrix. For more details on the method see [99].

Adaptation of Algorithm 5 to non-circular contours (Adapted Statistical Texture Analysis)

The algorithm 5 also had to be adapted to the non-circular contours detected by the automated segmentation method used. The difference from the original verson is the fact that the two contours are not circular therefore not possibly concentric, but the normalization proces is analogous to the traditional one. Once the defined region is normalized the extraction of features is done in the same way as in the original version.

7.4 Feature Selection

The algorithms implemented originated 14 different features. Due to this dimensionality it is possible that the best classification results are not obtained using all the features, but a subset of them. It is therefore convenient to search for the optimum number and set of features. We used the “Sequential Forward Floating Selection” (SFFS) [209] to perform the feature selection. The SFFS is basically a combination of search methods such as “Plus-l-Minus-r” [252] and Sequential Forward Search (SFS) [288]. The appearance of “floating" comes from the fact that the values l and r are not fixed, i.e., they can “float". Another aspect is the dominant direction of search, including (*forward*) or excluding (*backward*) characteristics [209]. As criterion function the *Mahalanobis* distance was used. The SFFS has shown to be competitive when compared to other selection techniques [113].

7.5 Classification

The classification results were obtained using three classification methods: Discriminant Analysis (DA), k-Nearest Neighbour (kNN) and Support Vector Machine (SVM).

For each cardinality, ($\aleph = 2, ..., 12$), the results of classification were obtained calculating the average of the results of 50 runs for classification of the images based on the corresponding best $\aleph$ features. The results were obtained by randomly dividing, in each run, the 1600 samples in two sets: 1000 samples for training and 600 for testing. The parameter k in kNN was optimized using cross-validation, and tested in the interval $[1,20]$ by steps of 1. For the SVM, we used a polynomial kernel and also used cross-validation for optimization of the parameters, performing a "grid-search" on the parameters of the models. Exponentially growing sequences of C were tested: $C = 2^N$ with N varying between -1 and 15. For the polynomial degree, d, values tested were: $d = 1,2,3,4,5$.

7.6 Datasets

The implemented methods were tested in five databases for iris liveness detection evaluation: Biosec [73], Clarkson [41], NotreDame [66], Warsaw [49] and MobBIO*fake* [235], presented in more detail in sections 3.3 and 6.2.

Biosec

The Biosec database [73] comprises a total of 1600 images: 800 real images and its corresponding 800 fake samples. The false database was obtained by pre-processing and printing on paper using a commercial printer the original images and then presenting these to the iris sensor, obtaining the fake copy.

Clarkson

The subset of Clarkson database [48] that we used for this study contains 270 real iris images and 400 fake iris images. The fake samples are images of eyes with contact lenses comprising 14 models of contact lenses. There are two different lighting conditions in the database, which was acquired by video (capturing 100 frames and with variation of focus).

NotreDame

The subset of images from the *ND_Cosmetic_Contact_Lenses_2013 Dataset* [66] that we used contains iris images of subjects without contact lenses, with soft contact lenses, and with cosmetic contact lenses, acquired using an LG 4000 iris sensor. The subset comprises 2000 "real" iris images and 1000 "fake" images.

Warsaw

The subset of Warsaw database [49] that we used contains 228 real images and 203 fake images. The fake samples are printed images of real iris images.

MobBIO*fake*

The MobBIO*fake* database [235] is comprised of 800 real iris images and 800 fake copies, captured with the same portable device and in similar conditions.

7.7 Experimental Results

7.7.1 Results of the proposed methodology using manual segmentation

In this section we present and discuss the results obtained by the proposed methodology for iris liveness detection using manual segmentation.

The feature extraction algorithms applied returned a set of 14 different features. The first step was to analyze individually each of the 14 different features, for each image dataset. By the analysis of the histogram obtained for fake and real images we can have a hint about which features will be good discriminative between fake and real images. For each histogram, the *threshold* obtained considering equal error rate (EER) allow us to determine the minimum error associated with that feature, for the considered dataset. In Table 7.2 are shown the minimum error values associated with each feature for each database.

Table 7.2: Minimum error associated with each feature for each database.

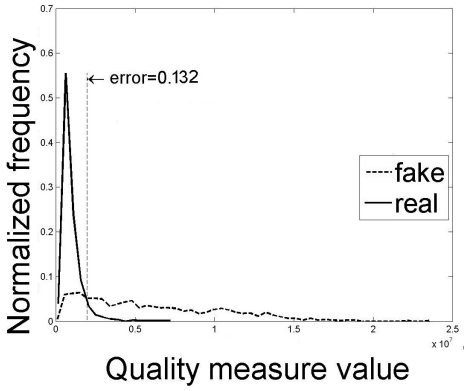
	Associated Error (%)		
Feature no.	*Biosec*	*MobBIOfake*	*Clarkson*
1 (alg1)	31.3	31.8	35.4
2 (alg2)	17.2	31.2	**26.3**
3 (alg3)	21.1	**21.9**	26.7
4 (alg4)	15.1	40.8	31.5
5 (alg4)	43.6	27.9	36.0
6 (alg4)	14.4	42.9	31.5
7 (alg4)	15.6	33.0	31.3
8 (alg4)	15.2	30.8	31.9
9 (alg4)	39.9	26.4	36.3
10 (alg4)	15.8	29.3	31.8
11 (alg5)	22.9	27.2	32.2
12 (alg5)	17.2	35.7	37.8
13 (alg5)	**13.2**	35.7	39.1
14 (alg5)	25.8	29.3	37.9

It is clear from the Table 7.2 that each dataset has a variable behavior concerning the features obtained. Simply observing the minimum errors (emphasized in the table) we may conclude that the "best" feature for one database is not necessarily the best for any of the others.

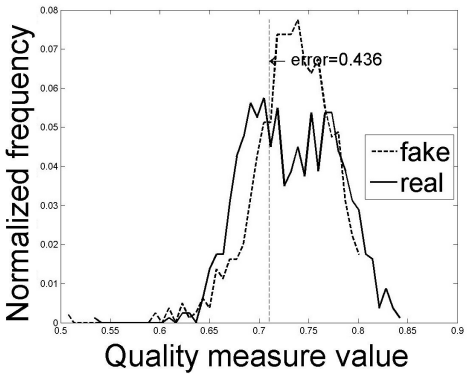
To enlighten a bit more how the discriminative power of each feature was analyzed we show in Figure 7.9 the best and worse feature for each database. These histograms illustrate clearly the efficiency of each feature in discriminating real images from fake images. For some features the lines for fake and real images are well separated while for others the overlap between the histograms is very significant.

The next step was to perform feature selection as to avoid possible redundancies in the set of features. Reducing the number of features to the strictly necessary will also improve the computational efficiency of the method. In Table 7.3 are shown the best subset of features for each cardinality, from 2 to 12, for each database.

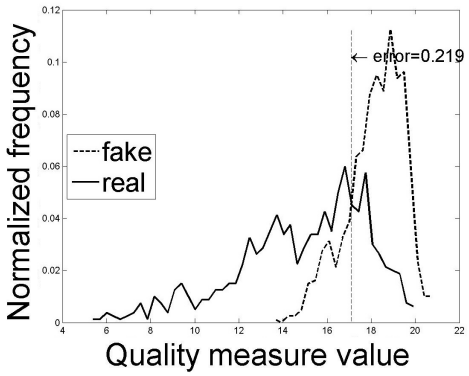
Again we observe the diversity of the results obtained for each database. Another relevant aspect is the combinations of features, in some cases we observe that features that individually do not have a good performance when combined provide the best subsets. This fact reinforces the pertinence of using a method for feature selection.



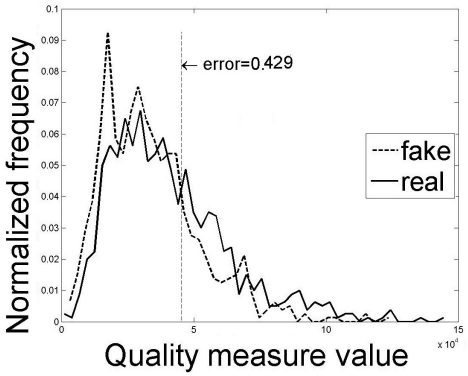
(a) *Biosec* - Feature 13 (best result).



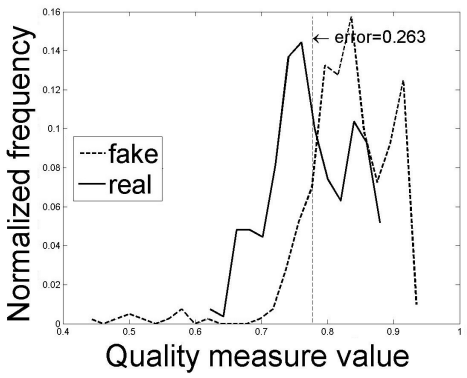
(b) *Biosec* - Feature 5 (worse result).



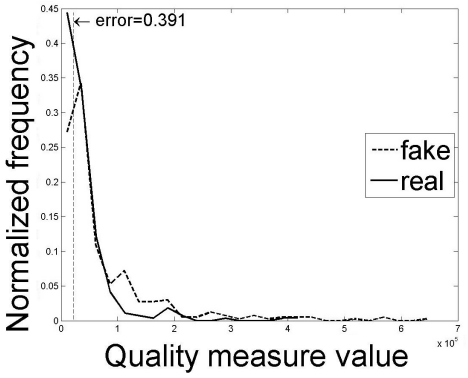
(c) *MobBIOfake* - Feature 3 (best result).



(d) *MobBIOfake* - Feature 6 (worse result).



(e) *Clarkson* - Feature 2 (best result).



(f) *Clarkson* - Feature 13 (worse result).

Figure 7.9: Histograms for the best result/smallest minimum error (left) and worse result/biggest minimum error (right) for each database.

Finally, Tables 7.4, 7.5 and 7.6 show the classification results for each cardinality, for each database, obtained using the best subset determined by the feature selection (averaged over the 50 runs).

The overall best results were obtained for Biosec database and the worst overall results were obtained for MobBIO*fake*. This is not a surprising result since MobBIO*fake* presents more chal-

Table 7.3: Best subset of features for each cardinality, for each database.

$\aleph$	Subset of features		
	Biosec	*MobBIOfake*	*Clarkson*
2	[1 6]	[3 10]	[3 14]
3	[1 2 11]	[5 8 10]	[9 11 14]
4	[1 2 6 11]	[3 5 8 10]	[3 9 11 14]
5	[1 2 6 11 13]	[3 4 7 8 10]	[2 3 9 11 14]
6	[1 2 6 11 12 13]	[3 4 7 8 9 10]	[1 2 3 9 11 14]
7	[1 2 5 6 11 12 13]	[3 4 5 7 8 9 10]	[1 2 3 9 11 12 14]
8	[1 2 5 6 7 11 12 13]	[3 4 5 7 8 9 10 13]	[1 2 3 5 9 11 12 14]
9	[1 2 5 6 7 9 10 11 13]	[3 4 5 7 8 9 10 12 13]	[1 2 3 5 9 11 12 13 14]
10	[1 2 5 6 7 9 10 11 12 13]	[3 4 5 7 8 9 10 11 12 13]	[1 2 3 4 5 6 9 11 12 14]
11	[1 2 3 5 6 7 9 10 11 12 13]	[2 3 4 5 7 8 9 10 11 12 13]	[1 2 3 4 5 6 9 11 12 13 14]
12	[1 2 3 5 6 7 9 10 11 12 13 14]	[1 2 3 4 5 7 8 9 10 11 12 13]	[1 2 3 4 5 6 7 9 11 12 13 14]

Table 7.4: Classification results for *Biosec* (classification errors in %).

$\aleph$	DA		kNN		SVM	
	μ	σ	μ	σ	μ	σ
2	10.24	0.99	10.34	1.23	10.16	1.17
3	4.36	0.92	4.43	0.83	4.68	0.67
4	0.52	0.24	0.76	0.27	0.77	0.34
5	1.14	0.48	0.89	0.34	0.78	0.29
6	0.85	0.38	0.56	0.27	0.54	0.26
7	0.92	0.32	**0.40**	0.20	0.57	0.31
8	0.93	0.30	0.47	0.25	0.56	0.33
9	1.80	0.47	1.11	0.34	0.87	0.28
10	1.28	0.39	0.46	0.28	0.73	0.26
11	1.30	0.28	0.40	0.24	0.52	0.33
12	1.68	0.59	**0.37**	0.24	0.50	0.26

Table 7.5: Classification results for *MobBIOfake* (classification errors in %).

$\aleph$	DA		kNN		SVM	
	μ	σ	μ	σ	μ	σ
2	18.03	1.31	16.52	1.47	17.29	1.13
3	29.29	1.54	17.34	1.25	20.69	1.83
4	17.50	1.42	**12.62**	1.18	14.36	0.94
5	18.03	1.30	13.00	1.26	14.18	1.20
6	18.29	1.44	12.82	1.02	14.33	1.15
7	18.88	1.35	13.27	1.35	14.55	1.27
8	18.31	1.11	13.52	1.21	13.74	1.28
9	18.34	1.28	14.44	1.25	14.11	2.22
10	17.58	1.38	13.92	1.19	13.39	1.01
11	17.15	1.35	14.51	1.44	12.53	1.39
12	17.25	1.12	14.45	1.21	**12.50**	1.21

Table 7.6: Classification results for *Clarkson* (classification errors in %).

א	DA		kNN		SVM	
	μ	σ	μ	σ	μ	σ
2	29.25	2.48	18.86	2.38	21.63	2.65
3	23.38	2.25	18.38	2.06	16.29	2.55
4	20.15	2.61	10.64	1.85	9.20	2.16
5	17.53	2.47	7.82	1.90	7.03	1.62
6	15.74	2.08	8.89	1.71	7.45	2.10
7	14.36	2.01	8.32	2.16	**6.77**	1.41
8	14.55	1.88	9.50	1.63	7.57	1.75
9	12.99	2.49	8.88	1.74	6.92	2.22
10	11.03	1.92	7.86	1.65	5.89	1.86
11	11.02	2.11	7.17	1.32	5.74	1.56
12	14.33	3.17	7.51	1.82	**5.69**	1.65

lenging characteristics. It was notorious from the study of features individually that this database presented the worse results. We interpret this fact as a sign that new databases were needed for the research of liveness in new scenarios.

Comparing the classifiers, we conclude that DA led to worse results. This fact is also not surprising since this classifier may be considered simpler than the others. The kNN achieved the overall best results.

Now, analyzing each database *per se*, we observe for the Biosec database that the best average classification rate was obtained with kNN. In terms of the cardinality of features, we note that the best average result, 0.37%, obtained with a subset of 12 features, is followed closely by the value 0.4% with only a cardinality of 7. And this again encourages the use of feature selection since the computational time and complexity is lowered if we lower the number of features.

Concerning the MobBIO*fake*, undoubtedly the classification errors obtained are higher than the other databases, what is not unexpected as we already referred. The best average results were obtained with the SVM classifier, 12.50%, but corresponding to a high cardinality, 12. Not very far form this value we find a subset with much lower cardinality, 4, for the kNN, with an average error of 12.62%.

Analyzing the Clarkson results, we note that the combination of features improved considerably the results when compared with the performance of the features individually. The best average result was obtained with SVM, 5.69%; this value is not as good as the Biosec best result but is better than the MobBIO*fake* one, but unfortunately it refers to a subset of high cardinality, 12. However, we may find a 3^{rd}-best value with a cardinality of 7.

7.7.2 Remarks on the methodology with manual segmentation

The proposed methodology combined several state-of-the-art features for iris liveness detection and was evaluated in two well known databases and a newly proposed database for iris liveness detection with images acquired in unconstrained conditions and with a handheld device. The

MobBIO*fake* database proved itself to be more challenging than the others. The results achieved in this referred database although not yet to be considered satisfactory lead to a new more challenging scenario.

Published works present methods tested with existing databases which achieve excellent results, (0% error classification rate). However, we note that some of these methods are closely connected with the particular database characteristics. The results achieved by our approach, in particular with MobBIO*fake* images, did not achieve that excellent accuracy, but we consider this justifiable by the fact that we avoided the use of methods strongly dependent on the images used, such as ratios of iris and pupil radius or areas, among others.

This study has shown the necessity of improving the existing methods and develop new ones more suitable to the new imaging scenarios. Another aspect to invest next was the segmentation step which preferably should be automatic.

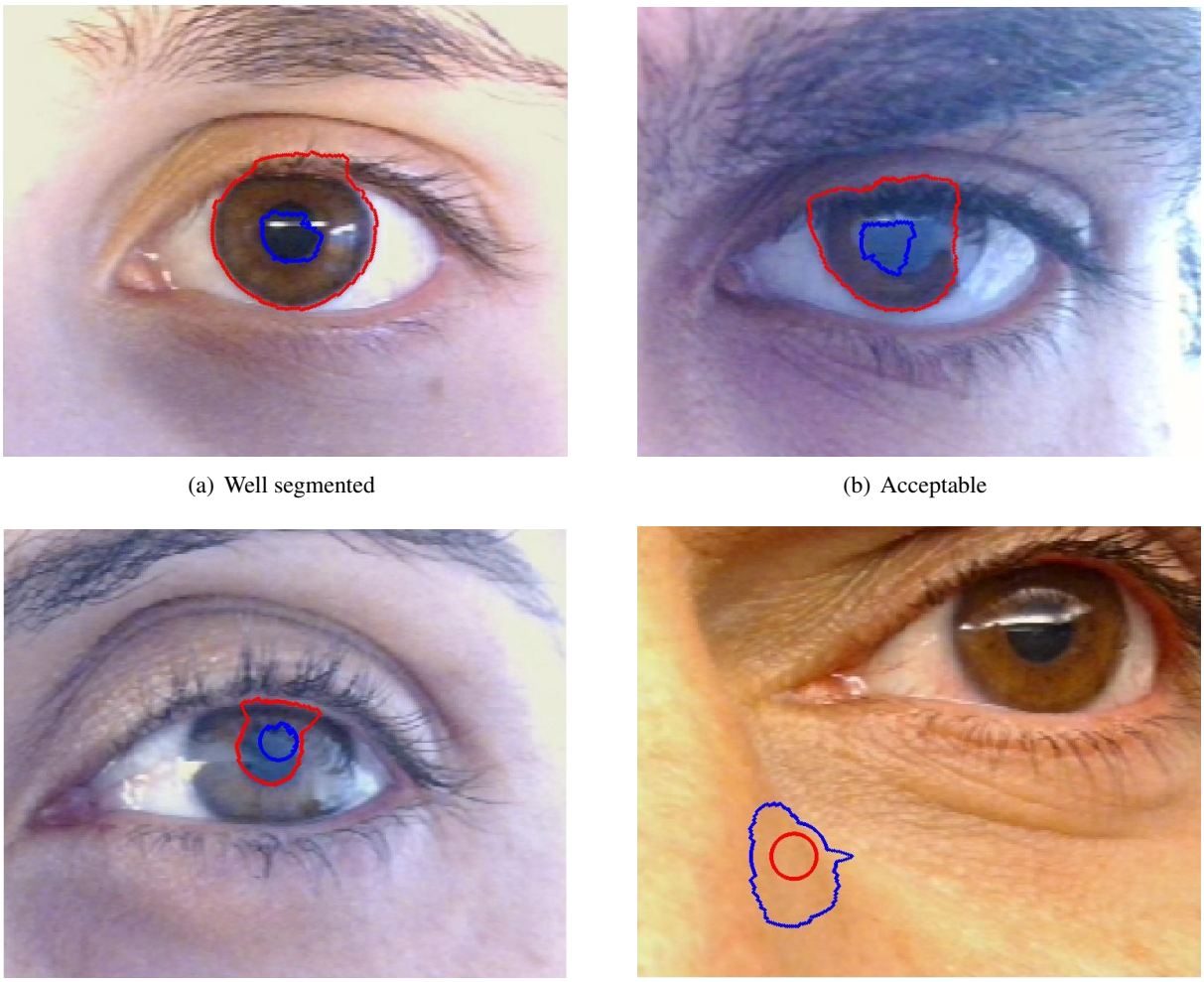
7.7.3 Results of the proposed methodology using automatic segmentation

Even though we aim to automatize our method the fact is that the segmentation step is by itself quite challenging. Although it is possible to find in the literature methods with good performance rates there is still much to improve. Specially concerning more unconstrained imaging scenarios. Due to the variable nature of images tested, the segmentation results leads to better or worse results depending on the database. It is necessary to balance the two aspects: automation versus quality. By one side, we intended to use an automatic segmentation process, but, on the other side, we wanted to avoid the excess of errors due to badly performed segmentation. Therefore, we choose to reject some of the segmented images in order to guaranty a minimum of accuracy of the method. It is almost as difficult to automatically decide about the quality of a result of a segmentation method as to perform the segmentation. This selection was made manually. Some examples are depicted in Figure 7.10.

In Table 7.7 it is shown the number of images accepted for each database and the percentage it represents in relation to the total number of images also presented in the table. These values are divided for real and fake images.

Table 7.7: Total number of images, number of images accepted after the automatic segmentation step and the corresponding percentage (%).

DB	Real images			Fake images		
	Total	Accepted	Percentage (%)	Total	Accepted	Percentage (%)
Biosec	800	686	85.75	800	524	65.50
Clarkson	270	205	75.93	400	314	78.50
MobBIOfake	800	717	89.63	800	731	91.38
Notre-Dame	2000	1709	85.45	1000	805	80.50
Warsaw	228	209	91.67	203	182	89.66



(a) Well segmented (b) Acceptable

(c) Rejectable (d) Rejectable

Figure 7.10: Examples of contours obtained in the segmentation step.

7.7.4 Methods not requiring contour detection

In Table 7.8 we show the results of classification for each database using only three features. The three features used do not require the detection of both, pupillary and limbic, contours. The algorithms 1 and 3 deal with the complete image. Algorithm 2 use a bounding box around the iris region. We included this method in this experiment taking in account that is much simpler to obtain a bounding box than performing the detection of both iris contours.

Concerning the methods that does not require iris contour detection, the best result was obtained for the Clarkson database, 5.8%. Observing the results we note that in this case the error rates are rather high compared with the results obtained with the complete set of features.

7.7.5 Methods requiring contour detection

In this section we present the results of the proposed method (comprising algorithms 1 to 5) in the two different segmentation scenarios: manual and automated segmentation. However in this new

Table 7.8: Classification results of methods 1, 2 and 3 (classification errors in %).

DB	DA		kNN		SVM	
	μ	σ	μ	σ	μ	σ
Biosec	**9.0**	1.1	9.8	1.2	9.5	1.1
Clarkson	**5.8**	1.4	7.7	1.55	7.43	2.01
MobBIOfake	20.9	1.4	20.6	1.6	**19.9**	2.2
Notre Dame	34.6	1.5	**29.2**	1.3	29.3	1.9
Warsaw	21.6	2.8	20.9	3.3	**20.6**	2.8

experiments, the sets of images we used was the subset of images selected after the segmentation step. This was made so that we can compare the results of the two scenarios.

In Table 7.9 we show the results of classification for each database using the information from manual segmentation.

Table 7.9: Classification results with manual segmentation (classification errors in %).

DB	DA			kNN			SVM		
	μ	σ	א	μ	σ	א	μ	σ	א
Biosec	0.8	0.4	6	0.5	0.3	12	**0.5**	0.3	10
Clarkson	9.1	2.1	12	6.2	1.6	10	**6.2**	1.6	11
MobBIOfake	16.9	1.6	11	16.3	1.2	8	**14.7**	1.5	11
Notre Dame	12.0	0.9	9	12.4	1.1	8	**7.7**	0.7	9
Warsaw	10.5	2.4	10	13.6	2.4	8	**10.0**	2.2	6

The best result (for the manual method) was obtained for the Biosec database, 0.5%. The worse result was obtained for MobBIO*fake*, 14.7%. The results obtained are not exactly the same of the results obtained with the methodology with manual segmentation presented before because in these results a selection of images were performed in order to allow us to compare these results with the following with automated segmentation.

In Table 7.10 we show the classification results for each database using the automatic segmentation.

Table 7.10: Classification results with automatic segmentation (classification errors in %).

DB	DA			kNN			SVM		
	μ	σ	א	μ	σ	א	μ	σ	א
B	3.4	0.6	12	**2.0**	0.6	12	2.1	0.6	11
C	9.8	2.1	11	8.3	2.1	10	**6.9**	1.8	10
Mf	16.5	1.9	9	13.0	1.4	11	**11.4**	1.2	12
ND	4.4	0.6	12	4.0	0.7	12	**2.9**	0.6	12
W	10.4	2.0	8	10.1	2.5	11	**7.8**	3.1	9

The best result (for the automatic method) was obtained for the Biosec database, 2.0%, as for the manual method. Again, the worse result was obtained for MobBIO*fake*, 11.4%. This latter is

not surprising since we have already observed in the previous results that this database has more challenging characteristics than the existing databases for iris liveness detection.

Considering the cardinalities of the best classification results, we observe that the best results are in most of the cases obtained with a high number of features. In our point of view this fact does not invalidate the usefulness of the feature selection although it may suggest to extend the set of features extracted.

Considering the classification methods, we observe that the SVM provided the best results when compared to the other classifiers: DA and kNN (with only one exception in ten results).

7.7.6 Remarks on the methodology with manual segmentation

The proposed methodology combined several state-of-the-art features for iris liveness detection and was evaluated in four well known databases and a iris liveness detection database with images acquired in unconstrained conditions and with a handheld device. The MobBIO*fake* database proved itself to be more challenging than the others.

When we compare the two scenarios, using manual or automated segmentation, we observe that for most databases the results did not significantly worsened moving from the manual to the automatic segmentation. In some situations, the error rate even lowered. We note the fact that due to the segmentation errors some images were discarded so the final results were obtained with a subset of the initial set of images images. This fact can be the cause for the difference in the results.

7.8 Participation in Iris LivDet2013 Competition

The proposed methodology combined with an automatic segmentation method [174, 175] was submitted to an iris liveness competition, the Iris LivDet2013 [41], held as part of the IEEE BTAS 2013[1] granting us the first place [42]. This participation was done in collaboration with Juliano Murari whom I co-supervised during his master thesis and with João C. Monteiro who is the first author of the cited papers on the segmentation method.

LivDet 2013 was open to all academic and industrial institutions. The competition supplied three public iris liveness databases, each including spoof and live images. Patterned contact spoof images were provided in the Clarkson University and Notre Dame datasets; printed iris spoof images in the Warsaw datasets.
Submissions from competitors included both win32 console applications as well as Matlab executable files, optimized using training datasets. Submitted algorithms were tested against separate testing datasets of images and error rates calculated.

Submissions were received from:

- ATVS - Biometric Recongnition Group, Universidad Autonoma de Madrid;

[1] http://www.btas2013.org/

- Federico, University of Naples Federico II;
- Porto, Faculdade de Engenharia da Universidade do Porto.

Each algorithm supplied a liveness score on the scale of $0-100$ for each image in the liveness databases. The error rates were determined using two metrics: FerrFake – Fake called Live; FerrLive – Live called Fake.

Observing Figure 7.11 representing the FerrLive rate and Figure 7.12 representing the FerrFake, it can be observed that across the two datasets that all three algorithms submitted against, Porto had the lowest average error rates with a 12.18% FerrLive and 9.98% FerrFake. Federico tested a lower FerrFake than the other algorithms, but their higher FerrLive created a higher average error rate. Error rates were lower for spoof images on the Warsaw dataset, which was to be expected as printed irises are easier to locate than contact lenses. The Clarkson University dataset had the overall worst error rates given that the dataset used various levels of blur in the training and test datasets.

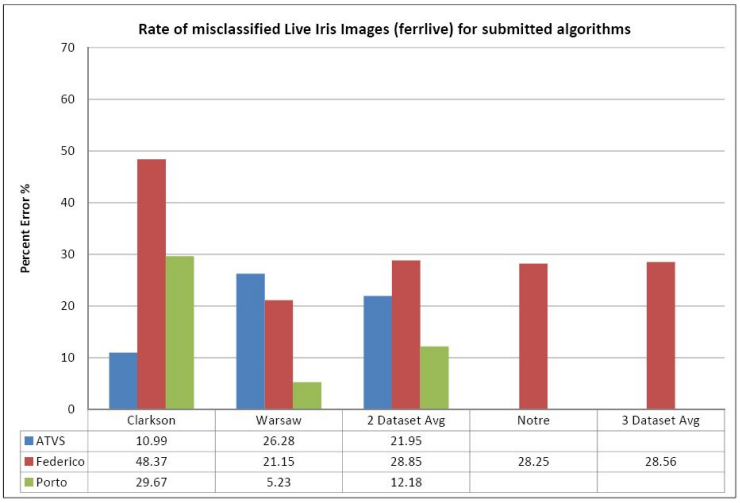


Figure 7.11: Results of Iris LivDet2015 (Ferrlive), from [42].

7.9 Conclusions

In this networked society of ours, mobile handheld devices have evolved from simple communication devices to mobile personal computers. Therefore, important personal private data as well as business data are stored on those mobile handheld systems. Naturally there is a strong need for user identification and access control. These need leads to the emerging field so called *Mobile Biometrics*. The current trend is not only to increase the recognition performance in mobile scenarios but also to increase the security of this processes. In this scenario the actuality of the iris liveness detection topic is unquestionable.

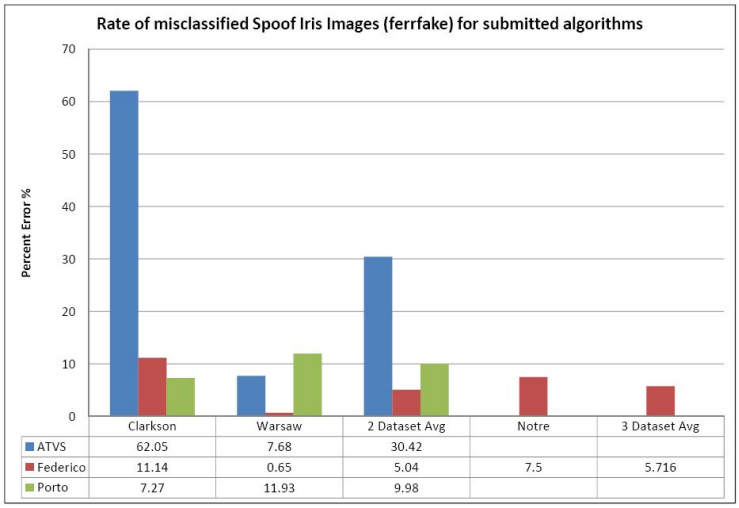


Figure 7.12: Results of Iris LivDet2015 (Ferrfake), from [42].

The proposed methodology for iris liveness detection was evaluated in four well known databases for iris liveness detection and a more recently (at the time) published database for iris liveness detection purposes with images acquired in unconstrained conditions and with a handheld device. The proposed method comprised state-of-the-art methods for iris liveness detection in a feature selection framework. Two different scenarios were tested whether the segmentation step was done manually or by an automated process. This latter proved to raise more difficulties to the liveness classification problem as the results worsened in a global evaluation. In both scenarios the MobBIO*fake* database proved itself to be more challenging than the others.

Published works present methods tested with existing databases which achieve excellent results, (0% error classification rate). However, we note that some of these methods are closely connected with the particular database characteristics. In the present work we avoid measures such as ratios of pupil and iris radius or areas, for example. Also we introduce a new variation which affects the overall results: the automatic segmentation step. Comparing our two scenarios, we observe that the impact of automatic segmentation as a source of error was not that significant, since only in some databases the classification results worsened considerably. We consider undoubtedly that this is the path to walk if we aim to pursue real-world applications in which segmentation must be done automatically and in real time.

Recently, new databases appeared with images captured with mobiles devices reinforcing the importance of such imaging scenarios in liveness detection research. So there are nowadays more tools to develop and evaluate new methods and a urge to deploy new solutions is continuously raised by the spread of mobile biometric solutions.

Chapter 8

Fingerprint Liveness Detection in the presence of Capable Intruders *

In this chapter we present a new methodology for fingerprint liveness detection. We observed that traditional approaches have been quite optimistic about the behavior of the intruder assuming the use of a previously known material. This assumption has led to the use of supervised techniques to estimate the performance of the methods, using both live and spoof samples to train the predictive models and evaluate each type of fake samples individually. Additionally, the background was often included in the sample representation, completely distorting the decision process. Therefore, we propose that an automatic segmentation step should be performed to isolate the fingerprint from the background and truly decide on the liveness of the fingerprint and not on the characteristics of the background. Also, we argue that one cannot aim to model the fake samples completely since the material used by the intruder is unknown beforehand. We approach the design by modeling the distribution of the live samples and predicting as fake the samples very unlikely according to that model. Our experiments compare the performance of the supervised approaches with the semi-supervised ones that rely solely on the live samples. The results obtained differ from the ones obtained by the more standard approaches which reinforces our conviction that the results in the literature are misleadingly estimating the true vulnerability of the biometric system.

8.1 Methodological limitations of current research

One limitation of several of the existing procedures is the inclusion of the background in the liveness decision. Center your attention in Figure 8.1(a) and Figure 8.1(b). In both cases the left fingerprint is real and the right is fake. However, in the first pair we may observe that the fingerprint occupies the most part of the image and, on the other hand, in the second pair, the background represents a significant area of the image and naturally it will dominate the representation (LPB, GLCMs, etc.). In the first pair it is expected that, for a good set of discriminant liveness detection features, the image features differ significantly between the real and the fake samples. In

*Some portions of this Chapter appeared in [233]

the second pair, as the background is very similar and occupies a considerable part of the image, the liveness system is likely to find much more challenging to discriminate the real from the fake in this second case than in the first case. Therefore, the influence of background in the decision making is not enabling the system to capture the true vulnerability of the biometric system and the security level may be underestimated in this case. Additionally, we are not interested in assessing the liveness of the background (which should always be lifeless) but only in assessing the liveness of the fingerprint. It is therefore with surprise that one verifies that fingerprint foreground segmentation is well established in the recognition works and often forgotten in the liveness detection works. We found in fingerprint liveness detection literature examples of inclusion of background and examples of manual or random crops to obtain the fingerprint partially [257, 278] but rarely an automatic segmentation step included in the process [80]. This observation shows how the segmentation step has been disregarded in most of the liveness detection methods.

(a) (b)

Figure 8.1: Example of differences in background area in two pairs of real and fake fingerprint images from two different datasets (images were degraded for privacy purposes). a) Pair of real and fake fingerprint images with a small area of background (Biometrika dataset) b)Pair of real and fake fingerprint images with a significant area of background (Crossmatch dataset)

Another methodological limitation is that models are designed and evaluated using fake samples of one type of material individually. The systems are developed and tested under the assumption that the intruder will fabricate the fake fingerprint by employing one of the materials used for training, resulting in optimistically estimating the security level of the system. At the design time, the developer assumes to possess labeled data representative of the real and fake samples and therefore resorts to standard binary classifiers (which may follow generative principles, like Naive Bayes, or be non-probabilistic, like SVM).The binary classifiers adopted to make the decision between real and fake samples implicitly assume that the training samples are representative of the complete population, with the test data to which the system is applied coming from the same distribution as the training data. Although that is a fair assumption for the real samples, it may be a crude model for fake samples created from a new material, see Figure 8.2. It may well happen that there is a mismatch between the distribution of the observations in training and testing data. In recent years, machine-learning researchers have investigated methods to handle mismatch between the training and test domains, with the goal of building a classifier using the labeled data in the old domain that will perform well on the test data in the new domain [210]. Recent

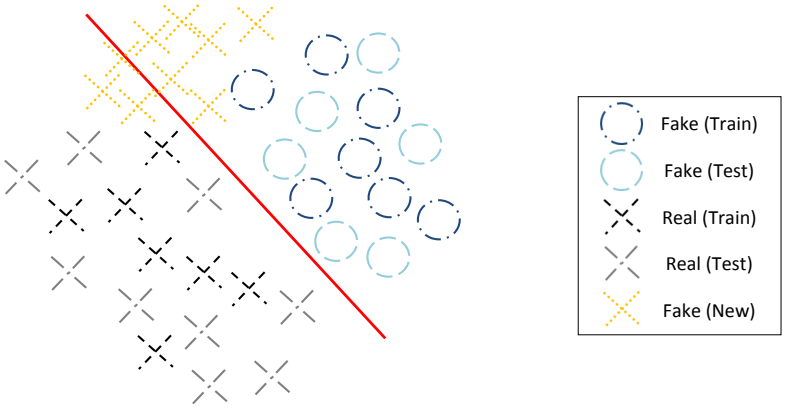


Figure 8.2: Illustrative Example. Black crosses and dark blue circles are fake and real samples in the training set. Light blue circles are real samples; gray crosses are fake samples from materials present in the training set; orange crosses are fake samples from a new material. The red curve represents the model learnt from the training samples.

works already depart from this conventional approach. Marasco and Sansone [163] performed an experimental comparison among fingerprint liveness detection approaches adopting materials for training different than those adopted for testing and concluding that the performance significantly decreases in this conditions. Although standard binary classifiers were used, this work encourages the adoption of different classification approaches in the liveness detection context. Another approach to the problem was presented by Rattani and Ross [219], consisting on designing an on-line scheme for automatic adaptation of a liveness detector to novel spoof materials encountered during the operational phase, which can lead to significant improvements in the performance. This approach is likely to accommodate well small differences (drifts) between materials, but it will likely under-perform when the new material is significantly different from the ones seen before. Still, this approach is in a sense complementary to ours.

8.2 Proposed approach

A first methodological contribution is the inclusion of an automatic fingerprint segmentation step before feature extraction for liveness detection. Instead of extracting features from all the acquired image, the image is segmented to find the mask defining the fingerprint region. The features can then be computed based solely on the information inside the mask (or the bounding box). This way of proceeding should lead to more realistic estimates of the vulnerability of the systems.

As new materials for fraudulent spoof attacks are going to continuously appear and become more and more sophisticated it is our conviction that the path to pursue is in the direction of using less and less information of spoofing materials and rely more strongly on the live samples. Our second methodological contribution includes: a) the realistic estimation of the performance in the

presence of new materials used to fabricate the fake samples; b) the use of decision models that rely only on the information from the real samples to detect the liveness.

The traditional approaches to liveness detection in general may be included in the category of **"supervised detection"** considering that they assume the availability of a training data set which has labeled instances for real and fake classes. Any unseen data instance is compared against the model to determine which class it belongs to [32]. Having in mind that in real-world solutions the system is not "aware" of the kind of attack that might be performed, in this work we explore other approaches in which we do not assume knowledge about the fake samples used in a spoofing attack. This study can be framed in a **"semi-supervised detection"** category considering that we use the real/live samples to train our model but not the fake/spoof ones. The techniques that operate in a semi-supervised mode assume that the training data has labeled instances for only the *live* class. Since they do not require labels for the fake class, they are more widely applicable than supervised techniques and *do not overfit* to the materials in the training set.

An innovative approach in the iris liveness detection context (concerning spoofing attacks using contact lenses) was presented by Bowyer and Doyle [26]. In this work, the authors compare a baseline experiment where the training and test datasets each contained iris images with three lens types. Any of several classifiers could be trained with local binary pattern texture features and achieve 100 percent correct classification of the test set. Then they performed another experiment using the same texture features, classifiers, and images, but the training set contained iris images with two of the three lens types and the test set contained iris images with the third lens type. The results obtained varied according to combinations tested but for the lower results the classification accuracy was no better than 75%. The authors concluded, that these results illustrate how experimental results obtained using the same lens types in both the training and testing data can give a very misleading idea of the accuracy that will result when a new lens type is encountered. Although this approach still fits in the supervised study (considering that in training both live and spoof samples are used) it goes further than the most traditional approaches since in the test step a new type of spoof samples is presented. We hint that the kind of conclusion achieved (in the referred work by Bowyer and Doyle) will hold regarding other biometric traits and particularly in the fingerprint liveness study. Therefore the path to pursue will put us more distant from the supervised classification approach.

8.3 Datasets

The experiments were run in the fingerprint datasets made available for the LivDet2013 [89] competition. The images of these datasets were acquired with four different devices: Biometrika, Crossmatch, Italdata and Swipe. All this four sensors acquire the fingerprint images by contact of the finger with the sensor, however, the first three may be classified as *touch* whilst the fourth is classified as *swipe* due to the fact that this latter requires the subjects to swipe their finger over the sensor surface [251]. More than 4000 images were taken with each of the aforementioned devices. In order to build the fake part of the database, seven different materials were used: Ecoflex

(1), Wood Glue (2), Latex (3), Modasil (4), Gelatin (5), Body Double (6) and Play-Doh (7). The fake images come from approximately 100 fingers of 20 people for the Crossmatch and Swipe datasets and from 100 fingers of 15 people for the Biometrika and Italdata datasets. Also, for the Crossmatch and Swipe datasets cooperative methods were used and for the other two the fingerprints were acquired through non-cooperative ways. The living images come from 440 fingers of 44 people for the Crossmatch dataset, from 250 fingers of 50 subjects for Swipe and from 300 fingers of 30 subjects for Biometrika and Italdata datasets. In Table 8.1 summarized the most relevant characteristics of the images obtained with each sensor.

Table 8.1: Some characteristics of the images from the LivDet2013 datasets.

Dataset	Sensors	Images resolution (dpi)	Images sizes	Gray scale levels
#1	Biometrika	569	315×372	256
#2	Italdata	500	640×480	256
#3	Crossmatch	500	800×750	256
#4	Swipe	96	208×1500	256

8.4 Segmentation method

The segmentation of the fingerprints was performed using the method presented by Ferreira et al. [72]. The method follows the morphological fingerprint segmentation algorithm presented in [68] suggesting a more robust binarization method instead of the simple adaptive thresholding binarization used. In this method, first, a block-wise range filter is applied to the grey-scale fingerprint image in order to enhance the ridges. Afterwards, the resulting range image is binarized using the Fuzzy c-means (FCM) algorithm along with a clusters' merging procedure. Finally, a set of morphological operations is applied to the binary image in order to compute the final foreground mask. The high-level operations that compose the algorithm are presented in Figure 8.3.

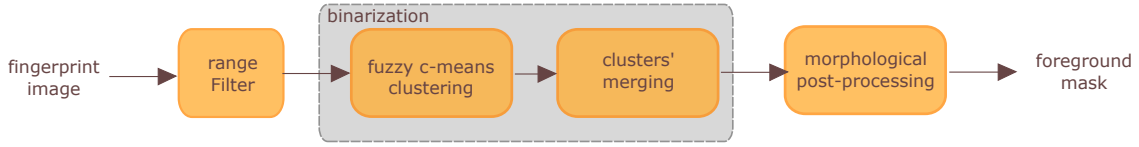


Figure 8.3: Main steps of the segmentation method (From [72]).

In our work we used the region of the fingerprint comprised by the bounding box obtained after the foreground mask computation. In the case of fingerprints images where the fingerprint occupied the most part of the image, see Figure 8.4(a), the segmentation will not impact so much the feature extraction step. However, when the background area is significant in the fingerprint

image, see Figure 8.4(b), we expect that the segmentation step will lead to an improvement in the discriminant capacity of the feature extraction.

(a) (b)

Figure 8.4: Examples of the result obtained by the segmentation method (images were degraded for privacy purposes). (**a**) Pair of original and segmented fingerprint images (Biometrika dataset); (**b**) Pair of original and segmented fingerprint images (Crossmatch dataset).

8.5 Feature Extraction

8.5.1 Weighted Local Binary Patterns (wLBP)

This method [304] combines Local Binary Patterns (LBP) [190,191] with a Scale Invariant Feature Transform (SIFT) [148]. The process starts with the generation of a Gaussian scale space. The output of this operation is a smoothed image in six scales. For each scale, the gradient orientation of each pixel is calculated and an histogram of gradient orientations is created. Every histogram is then converted into a descending rank, from 7 to 0. The LBP method labels the pixels of an image by comparing them with their neighborhood. Combining the LBP method with SIFT results in weighted maps. Three simple maps are constructed using the 3 first scales and a fourth map results from the combination of the last scales. Each map is divided in 8 by 8 blocks and three statistical features are extracted from each block (standard deviation of the wLBP histogram, mean of wLBP map and standard deviation of the wLBP map). That results in a 768 dimensional feature. The only adaptation made in our implementation is the use of all the 8×8 blocks instead of discarding the first and last rows of the image as in the original method. For more details on the method see the work of Zhang *et al.* [304].

8.5.2 Gray Level Co-occurence Matrices (GLCM)

This method [286] is based on GLCM [97] which characterize the relationship between neighboring pixels. Fourteen features are extracted from each GLCM matrix: angular second moment,

contrast, correlation, variance, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy, two information measures of correlation and the maximal correlation coefficient. These features are orientation dependent so four values can be obtained for each feature based on the four orientations (0°, 45°, 90°, 135°). The mean and standard deviation of the four values (four orientations) of each 14 measures, compose a set of 28 features. For more details on the method see the work of Wei et al. [286].

8.6 Classifiers and parameter optimization

Our experiments comprise different classification approaches. We start with the traditional approach using a binary classifier and training and testing within each dataset, then we introduce modifications in the training and testing datasets and end up with a study where only the information about the live samples is used for training our models. For the **classification** task we SVM [46], one-class SVM (OCSVM) [230, 261] and a Mixture of Gaussians Model (GMM).

In the first three studies (1, 2 and 3.1), we used SVM with a polynomial kernel. For optimizing the parameters a "grid-search" was performed on C and d parameters: the exponential growth of $C = 2^N$ was tested, with N varying from -1 to 15 and the polynomial degree (d) was tested with the following values $\{1,2,3,4,5\}$. In study 3.2, we used a OCSVM and a GMM. Concerning the SVM, we used the RBF kernel and for optimizing the parameters a "grid-search" was performed on γ and n parameters: with γ varying from 1×10^{-06} to 1×10^{-04}, for wLBP, and γ varying from 1×10^{-03} to 5×10^{-02}, for GLCM; and the n was tested from 1×10^{-02} to 2×10^{-02}. The optimization of the parameters was performed by nested cross-validation therefore the evaluation is done in the test set which was not used for that purpose. More details will be given in the following sections. Concerning the GMM, the number of components varied from 2^3 to 2^7 and the covariance matrix was not conditioned to be diagonal. For the OCSVM we used an implementation available in the Matlab-based (Mathworks, version R2012b) toolbox "LibSVM" [33]) and concerning the GMM, we used the Matlab-based (Mathworks, version R2012b) toolbox "Netlab" [177].

8.7 Study 1: Impact of fingerprint segmentation

As in the traditional approaches, the classification is performed for each material individually. Therefore, the fake samples used for training and testing are made of the same type of material. The live samples were divided in equal parts according to the number of materials and each part was used combined with the spoof samples of each material. In this study we consider the results obtained with and without segmenting the images. Also, we present the results obtained for each material and the median result of all materials in order to compare it with some state-of-the-art results. In this study, the classification error was obtained by averaging the misclassification error rate in 50 runs and, in each run, the data was divided randomly in 62.5% of the samples for training and 37.5% for testing. The training test was by its turn divided in the same way and the optimization of the parameters was performed by nested cross-validation in this dataset.

8.8 Studies 2 and 3: From supervised classification to semi-supervised classification

As already stated before, we consider the segmentation step to be crucial to the fingerprint liveness detection problem as it is in the fingerprint recognition field. Examples were showed that illustrate well the differences observed in the background areas in fingerprint images from distinct datasets. Naturally this factor will impact the liveness detection performance considering that the background will a have similar nature and on the other hand, ideally for the liveness detector, a live sample will differ from a spoof one. The fact that in some fingerprint images the background may constitute nine tens of the image this will surely impact on the liveness detection method. Taking account of these observations, we restrict the following study to the segmented images.

8.8.1 Study 2

In the "mixed sets" study, the classification is performed within a set composed by a mix of samples made by all types of materials. Considering that, in practice, the system is not aware a priori what is the kind of fake sample to be used in an attempt to fool the system we note that the classification performed in study 1 is not very realistic. The study 2 consists in applying the same methodology used within each material in study 1 but now in a set comprising all the types of materials. In this study, we assume that we have information about all the possible materials (since the training and test sets comprise all the different materials used for making the fake samples) but we do not know which one will be selected by the intruder. In this study the classification error was obtained by averaging the misclassification error rate in 50 runs and in each run the data was divided randomly in 62.5% of the samples for training and 37.5% for testing. The training test was by its turn divided in the same way and the optimization of the parameters was performed by nested cross-validation in this dataset.

8.8.2 Study 3

In the third study we aim to go further and assume that there is no complete knowledge about the kind of material will be used in a spoof attack to be performed. In this specific situation, this means that the system is not aware of all possible types of materials used to build the fake samples or none at all. Two different situations will be consider: in study 3.1 we use one material to test and the remaining materials to train; in study 3.2 we do not assume any knowledge about the fake samples, we use only the real samples train our model and perform a "one class" classification study.

Study 3.1 In this approach, the "leave one material out" study, we assume that we know some possible materials but we do not have knowledge about the one used in the spoofing attack that is going to be perpetrated. So, to perform the classification we use one material to test and the remaining materials to train. We used a k-fold cross-validation in which, for k sets of materials, 1 was used for testing and $k-1$ were used for training. The classification error was the mean of the

k misclassification rate (one for each fold). Also, in this last study, we used the optimization of parameters as described previously. The real samples are divided in k equal parts and each one is combined with one dataset of fake samples.

Study 3.2 Then we moved a step further aiming to perform the semi-supervised approach to our problem. We intended to extend the lack of knowledge about the fake samples to its limit. Therefore, we use for training only the real/live samples and then we perform the testing step on a set composed proportionally by real/live and fake/spoof samples (the latter of only one specific type of material at each time). In this study, for k sets of materials, we consider k folds in which 1 material was used for testing and the classification error was the mean of the k misclassification rate (one for each fold). The training step is done using only the information of the real samples. Considering that the number of features for wLBP was considerably high, a PCA feature selection method was applied using the training data. Then we trained and tested our model using only the best features chosen. The set of features and its cardinality is variable according to the datasets we studied. Experiments were run using two classifiers: a OCSVM and a GMM. Concerning the OCSVM, we chose the RBF kernel (which showed to perform better than the linear or the polynomial). For optimizing the parameters a "grid-search" was performed on γ and n parameters as described previously. Concerning the GMM, the number of components varied from 2^3 to 2^7 and we tested the two types of covariance, the one that obeys the condition that the covariance matrix has to be diagonal and the other that do not impose that condition, obtaining better results for the second one, which are the results presented.

8.8.2.1 Time Efficiency

Since the methods were implemented in MATLAB without any efficiency concerns, a straightforward assessment of the time efficiency is not fair. Nevertheless, some comments on the running time are in order. The most demanding operation is clearly the feature extraction, being the time duration of the remaining operations negligible when using the anti-spoofing system. The feature calculation time, especially concerning the wLBP method, grows with the size of the images. We followed the option of the authors of the method [304], normalizing the iris images into the same size (400×400). For the GLCM method, the feature extraction is significantly faster than for the wLBP method so we kept the original size of the images (that could already have been reduced in the case of segmented images). While the GLCM does not raises any concern for real-time applications, the wLBP extraction took several seconds per image in our implementation. However, as shown by [25], LBP implementation can be brought much closer to real time implementations. Concerning the SIFT part, an option includes replacing SIFT by a fast approximation (such as SURF) to weight the LBP.

8.9 Discussion of results obtained in study 1

The two feature extraction methods were first applied testing the different materials separately (study 1), like presented in section 8.7. Tables 8.2, 8.3, 8.4 and 8.5 present the results obtained for

the Biometrika, Crossmatch, Italdata and Swipe subsets, respectively, using the whole image and the segmented fingerprints.

Table 8.2: Average Classification Error Rate (ACER) for each material of the Biometrika dataset, study 1 (ACER in %).

	Biometrika									
	Ecoflex		Gelatin		Latex		Modasil		Wood Glue	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
wLBP	**1.48**	0.69	8.13	1.41	3.04	0.99	2.43	0.77	2.66	0.91
GLCM	**1.22**	0.59	4.39	1.46	6.63	1.47	7.86	1.54	10.21	2.35
wLBP (segmented images)	**0.39**	0.30	4.25	1.22	1.85	0.72	1.63	0.71	1.90	0.79
GLCM (segmented images)	1.48	0.76	4.90	1.29	9.97	1.75	6.66	1.47	15.63	2.52

Table 8.3: Average Classification Error Rate (ACER) for each material of the Crossmatch dataset, study 1 (ACER in %).

	Crossmatch							
	Latex		Wood Glue		Body Double		Play-Doh	
	mean	σ	mean	σ	mean	σ	mean	σ
wLBP	16.79	1.056	16.71	0.81	16.84	0.74	**1.93**	0.48
GLCM	3.99	0.65	**1.64**	0.45	1.74	0.36	4.72	0.93
wLBP (segmented images)	**0.09**	0.12	0.11	0.12	0.11	0.12	0.10	0.12
GLCM (segmented images)	7.02	1.78	3.30	1.09	**2.35**	0.67	8.60	1.86

Table 8.4: Average Classification Error Rate (ACER) for each material of the Italdata dataset, study 1 (ACER in %).

	Italdata									
	Ecoflex		Gelatin		Latex		Modasil		Wood Glue	
	mean	σ	mean	σ	mean	σ	mean	σ	mean	σ
wLBP	2.47	0.92	3.47	1.46	3.05	1.09	**2.29**	0.92	3.86	1.03
GLCM	**1.21**	0.52	6.26	1.56	7.15	1.71	3.30	1.48	6.97	1.38
wLBP (segmented images)	**1.92**	0.65	4.69	1.13	3.22	0.93	2.36	0.90	3.67	1.12
GLCM (segmented images)	**1.35**	0.80	4.64	1.05	4.73	1.24	2.65	1.14	6.44	1.65

For the Biometrika dataset, the best result is obtained for the Ecoflex material for both wLBP and GLCM methods, and also for both non-segmented and segmented images, respectively 1.48% and 0.39%; 1.22% and 1.48%; see Table 8.2.

For the Crossmatch dataset, the best result, for wLBP, for non-segmented images, 1.93%, is obtained for the Play-Doh material, and for segmented images, 0.09%, is obtained for the Latex material; and for the GLCM method, for non-segmented images,1.64%, is obtained for the Wood Glue material, and for segmented images, 2.35%, is obtained for the Body Double material; see Table 8.3.

Table 8.5: Average Classification Error Rate (ACER) for each material of the Swipe dataset, study 1 (ACER in %).

	Swipe							
	Latex		Wood Glue		Body Double		Play-Doh	
	μ	σ	μ	σ	μ	σ	μ	σ
wLBP	13.49	1.64	6.85	1.01	9.83	1.25	**6.77**	1.26
GLCM	8.71	1.87	6.06	1.14	7.46	1.71	**5.87**	1.32
wLBP (segmented images)	10.92	1.61	8.65	1.42	**7.10**	1.38	11.43	1.52
GLCM (segmented images)	8.62	1.70	7.54	1.55	**6.94**	1.54	7.43	1.17

For the Italdata dataset, the best result, for wLBP, for non-segmented images, 2.29%, is obtained for the Modasil material, and for segmented images, 1.92%, is obtained for the Ecoflex material; and for the GLCM method, for non-segmented images, 1.21%, is obtained for the Ecoflex material, and for segmented images, 1.35%, is obtained for the Ecoflex material; see Table 8.4.

For the Swipe dataset, the best result, for wLBP, for non-segmented images, 6.77%, is obtained for the Play-Doh material, and for segmented images, 7.10%, is obtained for the Body Double material; and for the GLCM method, for non-segmented images, 5.87%, is obtained for the Play-Doh material and for segmented images, 6.94%, is obtained for the Body Double material; see Table 8.5.

Regarding the type of material we can notice that, for both methods, the material that leads to the best error rate is different for each of the four sensors. Nevertheless, we should notice that not all materials are used for all sensors. But for the sensors that use the same materials this observation holds: for Biometrika and Italdata the best result corresponds to different materials however the Ecoflex predominates, and for the Crossmatch and Swipe there are no material that seems to lead to better results than all the others.

Comparing the results of non-segmented and segmented images, we observe that for wLBP the best result improve with the segmentation (with only one exception) however individually for each material the results with the segmentation does not always improve. Regarding the GLCM, we observe that even the best result does not always improve. Nevertheless, we sustain the necessity of performing the segmentation step with the fact that the background of the images vary among the different types and we believe that the feature extraction is highly biased by the background when we use the whole image. The discrepancies occur due to the fact that the descriptors are not equally sensitive to features observed.

In an overall analysis, the best result for wLBP without segmentation, 1.48%, is obtained for the Ecoflex material regarding the Biometrika sensor; and with segmentation, 0.09%, is obtained for the Latex material regarding the CrossMatch sensor. In an overall analysis, the best result for GLCM, without segmentation, 1.21%, is obtained for the Ecoflex material regarding the Italdata sensor; and with segmentation, 1.35%, is obtained for the Ecoflex material regarding the Italdata sensor.

In Table 8.6, a comparison between types of fake fingerprint samples is presented, where it can

be observed which combination of mold and sensor results in higher misclassification rate. We can thereby observe that, concerning the wLBP method, the Play-Doh material may be considered "more difficult" to detect than the others and concerning GLCM is the Wood Glue material that presents a higher error rate.

Table 8.6: Average Classification Error Rate (ACER) for the each type of fake mold and the respective sensor (ACER in %).

Material	wLBP	Dataset	GLCM	Dataset
Ecoflex	1.92	Italdata	1.48	Biometrika
Gelatin	4.69	Italdata	4.90	Biometrika
Latex	10.92	Swipe	9.97	Biometrika
Modasil	2.36	Italdata	6.66	Biometrika
Wood Glue	8.65	Swipe	**15.63**	Biometrika
Body Double	7.10	Swipe	6.94	Swipe
Play-Doh	**11.43**	Swipe	7.43	Swipe

8.10 Comparison of results obtained in study 1 with some state-of-the-art methods

To compare the results obtained with some results of state-of-the-art we present in Table 8.7 the average MER obtained in study 1 (using segmented images) and compare them with the three best methods of the *Fingerprint LivDet2013 Competition*: UniNap1 [89], Anonym2 [89] and Dermalog [89]; and other four state-of-the-art methods evaluated in the LivDet2013 dataset: AugLBP [76]; AugCN [76]; HIG_best [91] and Pore Analysis [124]. We only present the average errors for each dataset because the results are not available by material for the other methods. The precision presented is coherent with the minimum presented in the literature. It has to be stressed that the evaluation protocol in LivDet2013 and in the other works are different. In the competition, the algorithm was trained with one training set and then evaluated with the test set. In the other methods the frameworks for the classification step are variable. Nevertheless, is still pertinent to compare the results in order to show that the feature extraction methods used in the present work are not out of the range of the state-of-the-art results evaluated in the same dataset.

Observing Table 8.7, it can be concluded that the results obtained in our experiments may not outperform the other results in all datasets but in average they are positioned second and third best. In the case of the CrossMatch dataset, the results achieved by both our tested feature extraction methods outperform considerably the others. This may be due to the fact that we segmented the images and, although we do not have information about all the methods, we believe that several of the methods do not perform the segmentation. We recall that in this dataset the images are significantly dominated by the background area. The only exception is the method based in convolutional networks [76] which has lower mean error rate than the presented GLCM method, however we

note the fact that this method determine a region of interest before extracting the liveness detection features.

Table 8.7: Mean of ACER for all materials for each dataset and their average, study 1 (using segmentation) and other state-of-the-art methods (ACER and average in %).

	Study 1 and State-of-the-art methods					
	Biometrika	Crossmatch	Italdata	Swipe	**Average**	
	μ	μ	μ	μ	μ	σ
wLBP (Study 1)	2.0	**0.1**	3.2	9.5	3.7	4.1
GLCM (Study 1)	7.7	5.3	4.0	7.6	6.2	1.8
UniNap1 [89]	4.7	31.2	3.5	13.8	13.3	12.8
Anonym2 [89]	1.8	49.4	0.6	6.1	14.5	23.4
Dermalog [89]	1.7	49.9	0.8	3.5	14.0	24.0
Aug LPB [76]	1.7	49.5	**2.3**	**3.3**	14.2	23.5
Aug CN [76]	**0.8**	3.3	2.5	7.7	**3.6**	2.9
HIG_*best* [91]	3.9	28.8	1.7	14.4	12.2	12.7
Pore Analysis [124]	2.2	34.9	1.0	-	12.7	19.2

8.11 Discussion of results obtained in study 2

Considering that, in practice, the system is not aware a priori what is the kind of fake sample to be used in an attempt to fool the system the classification performed in study 1 is not very realistic. Then different studies were applied aiming to overcome this limitation, presented in section 8.8.1. In study 2 it is assumed that we have information about all the possible materials used in spoofing attacks to be performed since the training and test sets comprise all the different materials used for fake samples. The results presented in Table 8.8 show that there is not a significant discrepancy from the results of study 2. We can observe some differences in specific datasets but if we look to the average error rate the difference is almost unnoticeable.

Table 8.8: Average Classification Error Rate (ACER) for each dataset and their average, study 2 (ACER and average in %).

	Study 2									
	Biometrika		Crossmatch		Italdata		Swipe		**Average**	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
wLBP (segmented images)	1.56	0.30	**0.09**	0.05	2.18	0.52	9.22	0.81	3.26	4.07
GLCM (segmented images)	7.94	0.69	5.94	0.66	**2.88**	0.42	8.14	0.81	6.23	2.44

8.12 Discussion of results obtained in studies 3.1 and 3.2

Our aim to broad the classification approach in order to perform a semi-supervised classification instead of a supervised classification conducted to the study 3, presented in section 8.8.2, in which

we limit the knowledge of fake/spoof samples that is used in the training step. In a first attempt, we still use some sets of fake/spoof samples in the training set but the novelty is that in the training set we include a different type of fake/spoof samples, this is the denominated study 3.1. In study 3.2, we then perform a "full" semi-supervised classification by performing the training of our model using only the real/live samples and then using fake/spoof samples in the test step.

In tables 8.9, 8.10, 8.11 and 8.12 we present the results of studies 3.1 and 3.2 (for the OCSVM and the GMM). These are the minimum MER obtained using PCA for feature selection, the number of features were optimized for all subsets and differ among them. When using the RBF kernel for the SVM, the parameters γ and n, were optimized like described in section 8.8.2 and the optimal parameters are different for each material as we are testing with only one material at a time and ignoring the other types of materials. Also, for the GMM the number of components was optimized and again the optimal value vary according to each subset.

Table 8.9: Average Classification Error Rate (ACER) for the Biometrika dataset, study 3 (ACER in %).

		Biometrika						
		Ecoflex	Gelatin	Latex	Modasil	Wood Glue	μ	σ
Study 3.1	wLBP	4.63	6.25	**1.63**	3.88	3.63	**4.00**	1.68
	GLCM	**19.63**	40.25	48.75	30.75	47.13	37.30	12.16
OCSVM	wLBP	**9.50**	20.88	17.00	11.75	17.25	15.28	4.58
	GLCM	**15.88**	18.13	43.50	28.75	48.00	30.85	14.53
GMM	wLBP	**7.75**	19.00	18.00	12.13	16.25	14.63	4.66
	GLCM	**10.88**	11.88	28.75	20.13	33.63	**21.05**	10.07

Table 8.10: Average Classification Error Rate (ACER) for the CrossMatch dataset, study 3 (ACER in %).

		Crossmatch					
		Latex	WoodGlue	BodyDouble	Play-Doh	μ	σ
Study 3.1	wLBP	8.53	0.08	**0.18**	2.31	2.78	3.79
	GLCM	55.64	43.11	27.56	**17.51**	35.96	16.83
OCSVM	wLBP	34.93	20.00	14.58	**11.02**	20.13	10.53
	GLCM	45.24	24.89	33.78	**23.02**	31.73	10.16
GMM	wLBP	2.22	**0.98**	1.33	1.24	**1.44**	0.54
	GLCM	47.82	**16.00**	24.89	17.69	**26.60**	14.66

Observing the results of tables 8.9, 8.10, 8.11 and 8.12, we may conclude that, in terms of average results, for wLBP, the GMM approach outperforms for only one sensor (Crossmatch), but for GLCM, the GMM approach outperforms for all sensors. These results show that if there is no knowledge about the material to be presented in a spoofing attack then the one-class classifier approach is more suitable than the binary classifier approach used in study 3.1, where the model is trained with several materials but then a new material is encountered in the test phase. One important aspect is how much we are being optimistic when evaluating our models. A capable

Table 8.11: Average Classification Error Rate (ACER) for the Italdata dataset, study 3 (ACER in %).

		Italdata						
		Ecoflex	Gelatin	Latex	Modasil	Wood Glue	μ	σ
Study 3.1	wLBP	**0.88**	4.50	3.13	7.50	13.63	**5.93**	4.93
	GLCM	37.25	**32.63**	49.63	42.25	44.25	41.20	6.53
OCSVM	wLBP	**16.88**	26.00	25.88	23.00	30.63	24.48	5.05
	GLCM	**33.88**	39.38	35.13	37.00	43.50	37.78	3.81
GMM	wLBP	**11.25**	16.50	17.25	17.63	21.00	16.73	3.51
	GLCM	10.63	**10.25**	13.00	10.75	20.13	**12.95**	4.16

Table 8.12: Average Classification Error Rate (ACER) for the Swipe dataset, study 3 (ACER in %).

		Swipe					
		Latex	WoodGlue	BodyDouble	Play-Doh	μ	σ
Study 3.1	wLBP	**15.28**	17.34	16.19	33.98	**20.70**	8.90
	GLCM	54.16	54.13	49.31	**33.43**	47.76	9.82
OCSVM	wLBP	44.24	46.19	46.57	**33.70**	42.68	6.07
	GLCM	45.89	44.54	37.24	**35.38**	40.76	5.23
GMM	wLBP	37.29	48.76	43.37	**26.28**	38.93	9.64
	GLCM	36.14	33.76	20.86	**20.52**	**27.82**	8.29

intruder will try to develop new materials for spoofing different from the known ones. So, even though these results are worse than the results obtained in studies 1 and 2, we consider the semi-supervised approach more realistic than the supervised one where the same material appears in the train and test phase assuming that the system developer will know all the spoofing materials.

8.13 Conclusions and future work

In this study we applied different classification studies in fingerprint liveness detection in order to broad the traditional approaches that use the knowledge about the fake/spoof samples for training the models. Therefore, in a first approach we mixed the fake materials in train and test sets instead of training and testing with only one specific material. However, this approach is still not very realistic since we assumed to have knowledge about all the possible spoof attacks. So, the next approach was to test with one material and train with the rest of the materials. As expected, this approach lead to worse results since we are using a complete unknown material in the test step. Finally, the last approach, which we consider the most worth following, consisted on using only the information of the real samples when training our model and then test it with real and fake samples. In fact, what we are performing is a semi-supervised classification characterizing the real samples and expecting our model to classify correctly as fake the spoof samples in the test set. Two different methods were used, a one-class SVM and a mixture of gaussians, being the

best results produced by this latter one. Although the results of the semi-supervised approach are worse than the supervised classification, we still consider the first to be more realistic. In our opinion, is more adequate to evaluate the robustness of a liveness method to unknown spoof attacks not assuming complete or even partial knowledge about the fake/spoof samples to be used by an intruder. We consider the results obtained in our experiments encouraging to pursue this approach in future works broadening the study to other databases and also other biometric traits. This approach has raised interest recently in the liveness detection field as some referred works show, but, to the best of our knowledge has not been yet fully studied and in our opinion can be further explored. In this work, we also illustrate the variability in the background of fingerprint images among the different datasets used and we argue about the necessity of a segmentation step.

Chapter 9

Robust Fingerprint Binarization

In this chapter we present a method for estimate the orientation map of a fingerprint based on the use of Markov Random Fields. Also, an enhancement method is proposed in order to perform a robust binarization of the fingerprint image allowing a best performance in minutiae extraction.

9.1 System Overview

The adopted fingerprint recognition system, represented in Figure 9.1, includes the following modules:

- Region mask estimation: consists on the identification of the region in the image containing the ridge structures and serves two purposes, on the one hand allows to restrict the processing to the foreground region, speeding up the following tasks, on the other hand allows to remove from the matching process minutiae close to the foreground boundaries and therefore more likely to be unreliable or "false minutiae".

- Ridge orientation map estimation: this represents the local direction of the ridge-valley structure. Typically, the ridge orientation is estimated following one of two approaches: a) correlation with a bank of templates (sine waves) of different frequencies and phases. The orientation for which the correlation with one of the templates is maximized is chosen as the local orientation; b) computation of the orientation of the gradient or related information.

- Frequency estimation: this step is used to estimate the inter-ridge separation in different regions of the fingerprint image.

- Directional enhancement: it should be noted that the isotropic smoothing of fingerprints would destroy the ridge-valley structure. Therefore, images are typically low-pass filtered along the orientation of the ridges/valleys to remove noise and bandpass filtered in the transversal direction to enhance the ridge/valley structure. The two filtering stages can be combined in a single 2D filter such as a Gabor filter. This block may use the estimated inter-ridge distance to set the bandwidth of the bandpass filter.

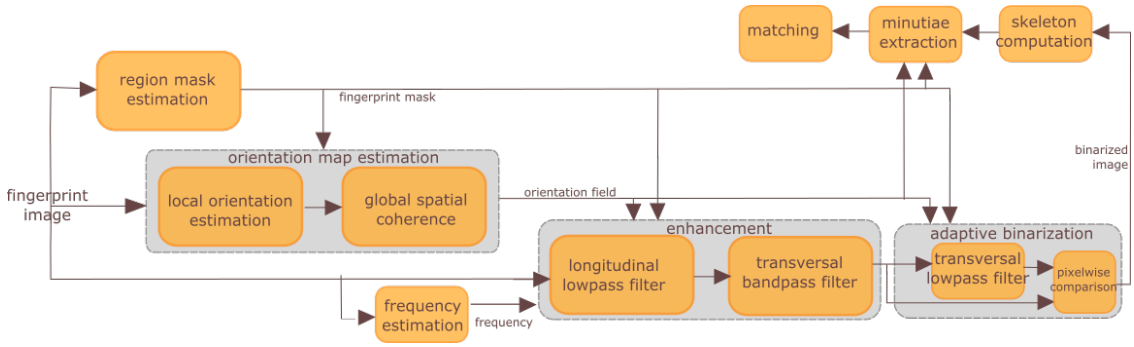


Figure 9.1: Block diagram for adopted fingerprint recognition system.

- Binarization: this step takes in account that the contrast is not uniform in the image, varying with the pressure made during the acquisition and near the border of the fingerprint the contrast is typically lower. Therefore, the binarization may include an adaptive method to set the thresholding decision.

- Skeletonization, minutiae extraction, and matching: the binarized images are then thinned, minutiae are extracted and matched.

We detail our contributions for ridge orientation estimation and directional enhancement in the following two sections.

9.2 Robust Orientation Map Estimation

In the adopted fingerprint recognition system one of the first blocks is the ridge orientation map estimation as can be seen in Figure 9.1. The methodology proposed is based on the use of Markov Random Fields.

9.2.1 Multi-label Markov Random Fields

Many early vision problems require estimating some spatially varying quantity (such as intensity, disparity or orientation). These quantities tend to be smooth; they vary smoothly at most points (but change dramatically at object boundaries). The only information available for the estimation task are noisy measurements in each pixel. These values are considered to be the “priors” for the true quantities. Typically they are obtained using only local information around the pixel and are therefore very noisy, lacking important properties such as smoothness.

The Markov Random Fields problem is to assign labels to the pixels so that the total penalty is minimized. For motion or stereo, the labels are disparities, while for image restoration they represent intensities; in our application they will represent the orientation. The penalty consists of two terms. One is the deviation penalty, or fidelity term, and the second is the separation penalty, or smoothing term. The fidelity term penalizes deviations from the assigned labeling and

the observed data (noisy measurements); the smoothing term measures the extent to which the labeling is not smooth [28, 102].

These vision problems can be naturally formulated in terms of energy minimization. The goal is to find a labeling x that assigns each pixel a label x_i, where x is both smooth and consistent with the observed data:

$$\min_{x_i} \sum_{i \in \mathscr{P}} G_i(x_i) + k \sum_{i \in \mathscr{P}} \frac{1}{\#N(i)} \sum_{j \in N(i)} F_{ij}(x_i - x_j), \tag{9.1}$$

where

- $\mathscr{P}$ is the set of pixels
- $x_i \in \mathscr{L}$ and $\mathscr{L} = \{\ell_1, \cdots, \ell_n\}$ is the set of labels
- $N(i)$ is the set of neighbors of pixel i, $\#N(i)$ is the number of elements in that set
- $G_i()$ is the deviation function
- F_{ij} is the separation function
- k is a constant, balancing the importance of the deviation and separation terms

That is, MRF is a multi-label assignment. The complexity of MRF depends on the form of the penalty functions. This complexity of MRF was fully resolved and classified according to the properties of the penalty functions for sets of consecutive integers. For convex penalty functions MRF is polynomial time solvable, and for non-convex the problem is NP-hard [28, 102].

Nevertheless, there are many methods in the literature that provide approximate solutions to the problem. One example is the Graph Cut algorithm used in many segmentation tasks. Another family of methods used to perform approximate inference in graphical models are the sum-product message passing algorithms. Within this type we find the Belief Propagation (BP) algorithm, and the version for loops Loopy Belief Propagation (LBP), that is able to perform inference in the case of grid-structured models. There are also other algorithms that follow different approaches. One example is the Iterated Conditional Models (ICM), an algorithm for optimization that follows a search paradigm.

9.2.2 Proposed method for Orientation Map Estimation

The orientation at a given position is typically estimated based on the information on a window centered in the position. Small windows are preferred to avoid over smoothing the estimation; large windows have the advantage of being less sensitive to noise and missing information. Implicit in this window based approach is that the orientation is constant inside the window. As such, large windows will tend to fail to properly estimate the orientation in zones of high curvature. In zones of constant orientation large windows tend to be more robust since they 'integrate' more information in the estimation process. Typically the size of the window is chosen to balance these conflicting goals. In here we propose two improvements to the standard scheme:

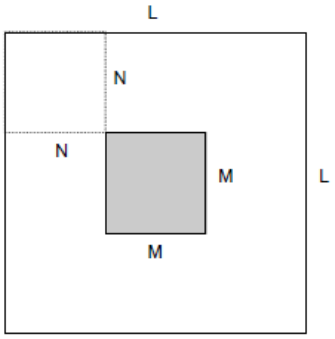


Figure 9.2: Block and window illustration for orientation map estimation method.

1. first, the local estimation of the orientation at a certain position is performed at multiple scales (equivalent to windows sizes). The choice is made for the scale with stronger evidence in favor of a given orientation. Let $Q(\theta,s)$ represent a measure of quality of the orientation $\theta \in [0,\pi[$ at scale s (we also define $C(\theta,s) \equiv 1/Q(\theta,s)$ as the cost function and we will use the concept that is more intuitive for each situation). The evidence of the scale s is defined as $E(s) = \frac{\max_\theta Q(\theta,s)}{\min_\theta Q(\theta,s)}$. We choose the scale s that maximizes the evidence, $s^* = \arg\max_s E(s)$ and the orientation that maximizes $Q(\theta,s^*)$, $\theta^* = \arg\max_\theta Q(\theta,s^*)$.

2. since the previous estimation is still noisy, we perform a regularization step using a MRF. The fidelity is defined as $G_i = C(\theta,s^*)$; the smoothing term is defined as $F_{ij}(\theta_i - \theta_j) = 1 - |\cos(\theta_i - \theta_j)|^\gamma$. The MRF problem results in solving

$$\min_{\theta_i} \sum_{i\in\mathscr{P}} C(\theta_i,s^*) + k \sum_{i\in\mathscr{P}} \frac{1}{\#N(i)} \sum_{j\in N(i)} (1 - |\cos(\theta_i - \theta_j)|^\gamma), \tag{9.2}$$

   where θ_i belongs to a finite set of orientations sampled in $[0,\pi]$ and $\gamma > 0$.

For efficiency purposes, and as typically done [284], the image is divided into a grid of blocks, each with $M \times M$ pixels. All the pixels within a block are assigned the same ridge flow direction. In the pixel at the center of the block, a window $L_s \times L_s$ pixels is used to evaluate $C(\theta,s)$. We set $L_s > M, \quad \forall s$; in this way some of the image that contributed to one block's results is included in the neighboring block's results as well. This helps minimize the discontinuity in block values as you cross the boundary from one block to its neighbor. This is illustrated in Figure 9.2.

Different cost functions $C(\theta,s) \equiv 1/Q(\theta,s)$ have been used in the literature. Since at this stage the ridge frequency is still unknown, template based method resort to a filter bank with different frequencies, with the inevitable trade off between the number of filters and efficiency. Gradient based methods are very susceptible to noise since at this stage the image has not yet been filtered. We opted therefore for a simple yet robust alternative. When the window is aligned with the true direction of the ridges, we expect to observe constant values along the longitudinal direction (see 9.3); when the window is not aligned, we expect jumps in the intensities. As such, we just sum the

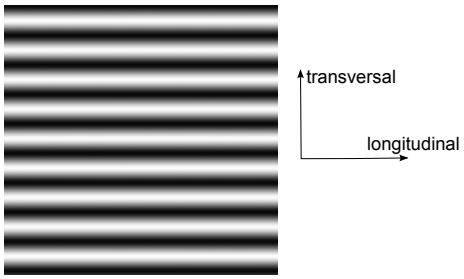


Figure 9.3: Pattern of horizontal stripes of a fingerprint.

jumps in intensity along the longitudinal direction:

$$C(\theta,s) = \frac{1}{L_s(L_s-1)} \sum_{i=1}^{L_s} \sum_{j=1}^{L_s-1} |x(i,j+1) - x(i,j)|, \tag{9.3}$$

where i runs over the transversal direction, j runs over the longitudinal direction, and $x(i,j)$ gives the pixel intensity in position (i,j).

9.3 Robust Fingerprint Enhancement

9.3.1 Recursive Gaussian filters

Gaussian filtering has established itself as one of the most widely used image processing operations in computer vision as well as biological vision. It is often a pre-processing step in many feature extraction techniques. The Gaussian function is defined by

$$g(t;\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}},$$

where the parameter σ denotes the Gaussian half-width. The Fourier transform of the Gaussian function is also a Gaussian:

$$G(w;\sigma) = e^{-w^2\sigma^2/2}$$

The Gaussian filter can be implemented using a finite impulse response (FIR) filter. It is reasonable to assume that the Gaussian function itself has finite support. It is typically assumed that the Gaussian function (and its derivatives) are approximately zero for $|t| > 4\sigma$. Gaussian filtering can then be performed via direct convolution. Exploiting symmetry in the filter, this FIR approximation to Gaussian filtering requires roughly $1+4\sigma$ multiplications and 8σ additions per output sample. The computational cost of applying an FIR Gaussian or Gaussian derivative filter grows linearly with the Gaussian half-width σ.

In contrast, the cost of an infinite-impulse-response (IIR) approximation to Gaussian filtering is independent of the half-width σ. IIR filters solve recursively a sequence of difference equations such as

$$\begin{aligned} y[k] = \quad & n_0 x[k] + n_1 x[k-1] + n_2 x[k-2] + n_3 x[k-3] \\ & - d_1 y[k-1] - d_2 y[k-2] - d_3 y[k-3] - d_4 y[k-4] \end{aligned} \tag{9.4}$$

or

$$\begin{aligned} y[k] = \quad & m_1 x[k+1] + m_2 x[k+2] + m_3 x[k+3] + m_4 x[k+4] \\ & -d_1 y[k+1] - d_2 y[k+2] - d_3 y[k+3] - d_4 y[k-4] \end{aligned} \tag{9.5}$$

Eq. (9.4) represents a causal system while Eq. (9.5) represents an anti-causal system. However, the recursive solution of a single difference equation cannot well approximate Gaussian filtering. The impulse response of the system corresponding to the previous equations is one-sided, not symmetric like the Gaussian function. Deriche [62, 63, 69, 96] proposed to split the symmetric filtering in the sum of subsystems, one *causal* and one *anti-causal* system. This splits the impulse response into two halves such that $h[n] = h_+[n] + h_-[n]$ with:

$$h_+[k] = \begin{cases} h[k] & k \geq 0 \\ 0 & k < 0 \end{cases}$$

$$h_-[k] = \begin{cases} 0 & k \geq 0 \\ h[k] & k < 0 \end{cases}$$

The corresponding coefficients of the two differential equations can be obtained by minimizing a suitable goal function [62, 63, 69]. Except for small values of σ, the IIR implementation is more efficient that the FIR alternative.

To compute the filtered output sequence $y[n]$, we apply the causal recursive filter to the finite-length sequence $x[n]$ to obtain an intermediate sequence $y_+[n]$. For a finite-length sequence, this estimation is efficiently performed from 'left to right' (or from the 'past' to the 'future'). Similarly, for the anti-causal system, we obtain the intermediate sequence $y_-[n]$ by operating from 'right to left' (or from the 'future' to the 'past'). The desired output sequence $y[n]$ is the simple sum of both.

9.3.2 Proposed method for Directional Enhancement

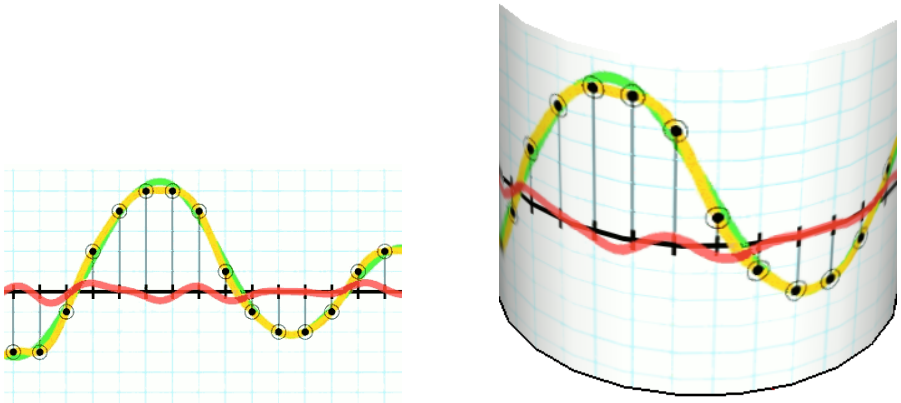
Consider the horizontal stripes represented in Figure 9.3. If the image was contaminated with noise, the 'optimal' way of recovering the original stripes would make use of an horizontal low-pass filter and a vertical bandpass filter centered at the frequency of the stripes.

Although the pattern in Figure 9.3 is far from being a good model for the whole typical fingerprint, it may be considered a good local approximation, after rotating the fingerprint according to the estimated orientation angle. As such, the directional smoothing has been again typically performed by defining a window oriented according to the orientation map at the current position and applying a longitudinal low pass filter, most often a Gaussian filter, and a transversal bandpass filter. Once again, it is implicitly assumed that the orientation is constant inside the windows and therefore it is sensible to filter the values according to the information along the orientation direction. This assumption puts a limit in the window size and tends to perform poorly in zones of high curvature.

Let's analyze first our proposal for the longitudinal lowpass filter.

Longitudinal Gaussian filter

It seems preferable to use the information along the complete ridge or valley curve to better filter the values. Conceptually, for every position, one could define a one-dimensional signal by following the complete ridge or valley curve according to the orientation field and then perform a one-dimensional filtering over that signal, as illustrated in Figure 9.4. The challenge now is on how to efficiently define the curved 1D signal.



(a) (b)

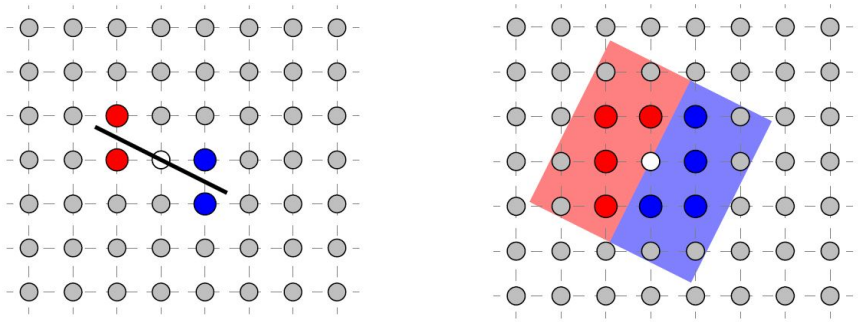
Figure 9.4: Signal Processing. a) Signal processing over a straight line; b) Signal processing over a curve. The curve corresponds to the ridge or valley structure.

Consider Figure 9.7(a). Although the image is a 2D signal, we want to perform a one-dimensional filtering dictated by the direction of the orientation field. Using the orientation field values $\theta(x,y)$ and a first order approximation, the previous and next values of the curved signal through position (x,y) is readily estimated as a combination of the neighboring values. In the example of 9.5(a) the line segment represents the angle as given by the orientation map (and provides a linear approximation to the ridge curve in the white pixel). The weighted average of the two red pixels and the weighted average of the two blue pixels provide the previous and the next values for the 1D smoothing of the white pixel. The weights in the averaging are a function of $\theta(x,y)$. Repeating for all the pixels, one obtain the previous and next values in each position. Iterating the process, one can easily accumulate the values of the curved signal through every position (x,y).

In the preceding explanation, one needs to define what's 'past' and what's 'future'. Any range of π radians would be suitable to set as the 'past' (and the complementary as the future). We set the range as illustrated in 9.5(b) to facilitate the implementation of the recursive Gaussian filter.

To apply Deriche's method to estimate $y[n]$ at position (x,y), and focusing only on the causal subsystem, one needs to know the previous element $y[n-1]$ and its position in the 2D grid. Assume a left to right, top to bottom scanning of the image. When visiting pixel p, half of its neighbors have already been visited (causal neighbors, see 9.5(b)).

Using the orientation field values $\theta(x,y)$ (see Figure) and a first order approximation, the previous position is readily estimated as a linear combination of the causal neighboring positions.



(a) (b)

Figure 9.5: Estimation of past and next values in the curve as a combination of neighboring pixels values. a) Estimating the previous and next values in the curve through a given pixel. b) Causal (red) and anti-causal (blue) neighbors for the 8-neighborhood. We assume a left to right, top to bottom scanning of the image.

The previous input value $x[n-1]$ and the output at the previous position, $y[n-1]$, can then be estimated using the same linear combination of the corresponding values observed in those causal neighboring positions. Proceeding left to right, one can then estimate the causal output $y_{+,h}$; proceeding similarly from right to left for the anti-causal subsystem one gets $y_{-,h}$; finally by summing, one obtains the desired filtered output $y_h = y_{+,h} + y_{-,h}$.

The just described scanning left to right works well propagating the filtered values when the orientation field is 'horizontally biased'; when the ridge/valleys structures are mainly vertical, the filtering in not so effective. In those cases, a top to bottom scanning would be preferable. We propose to conduct an 'horizontal scanning', left to right obtaining y_h and a 'vertical scanning' obtaining y_v. Finally, at each position the smoothed value will be estimated as

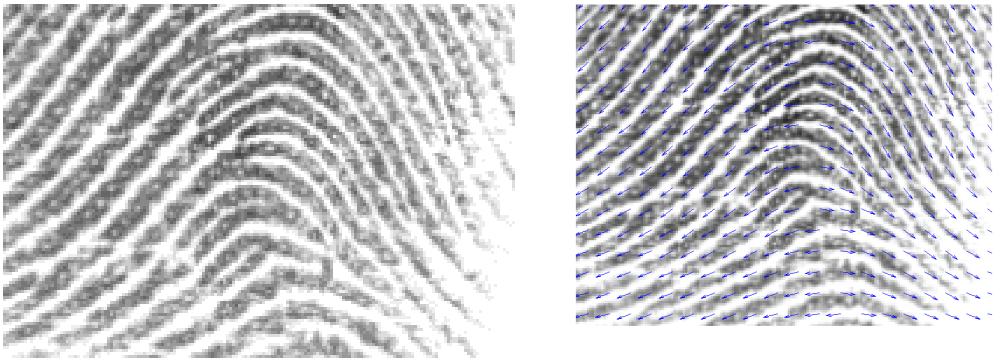
$$y = \sin^2(\theta_i - \theta_0)\, y_h + (1 - \sin^2(\theta_i - \theta_0))\, y_v. \tag{9.6}$$

If the orientation is horizontal, the left to right scanning will dominate the estimation; otherwise the top to bottom scanning prevails.

Figure 9.7 show an example of longitudinal enhancement: first the original image, then figures 9.7(b) and 9.7(c) illustrate the horizontal and the vertical scanning for a fingerprint image. At last 9.7(d) illustrates the final enhanced image resulting of the combination of the previous two smoothed images.

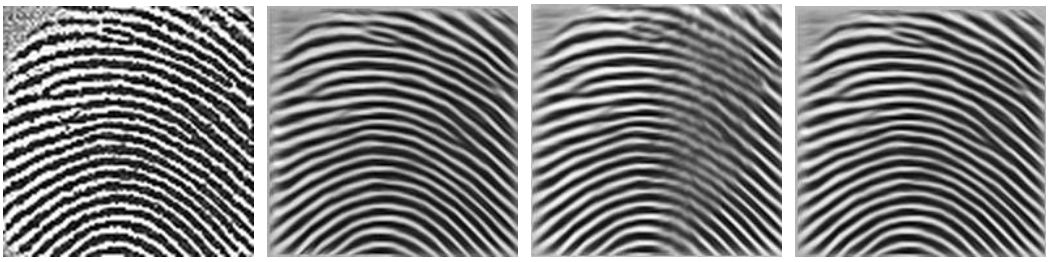
Transversal bandpass filter

Regarding the transversal enhancement, a 1D cosine filter is applied to perform the transversal bandpass filtering. Since no efficient recursive implementation of the bandpass filter is known, this method was implemented using a FIR approach.



(a) (b)

Figure 9.6: Longitudinal Directional Smoothing based on orientation map. a) Original image; b) Original image with orientation field superimposed.



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

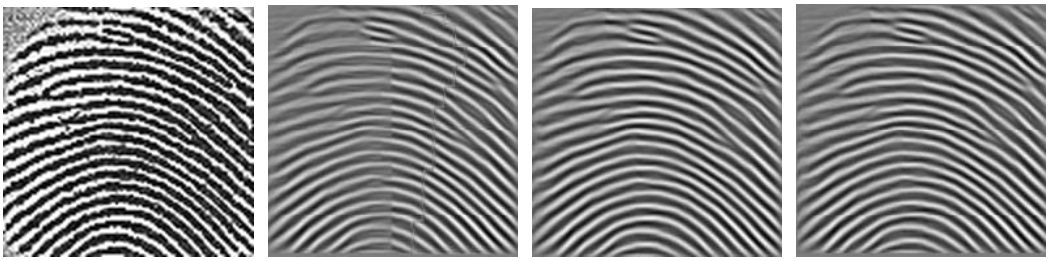
Figure 9.7: Longitudinal Directional Smoothing. a) Original image; b) Smoothed image image after "horizontal" scanning; c) Smoothed image image after "vertical" scanning. d) Final smoothed image.

A curved bandpass filter is computed by mapping a curved region to a two-dimensional array, followed by a point wise multiplication with an unrotated cosine filter, g. The curved region centered in pixel (i, j) consists of $2p+1$ parallel lines and $2q+1$ points along each line. The corresponding array $A_{i,j}$ contains the interpolated gray values. Therefore for a pixel given by coordinates (i, j) then its enhanced value is given by:

$$E(i,j,A_{i,j},f_{(i,j)}) = \sum_{k=0}^{2p+1}\sum_{l=0}^{2q+1} A_{i,j}(k,l)\cdot g(k-p,l-q,0,f_{(i,j)},\sigma_x,\sigma_y).$$

In the case of the tranversal enhancement, this is performed in the orthogonal direction relatively to the orientation of the ridge given by the orientation direction map. The left to right scanning works well propagating the filtered values when the orientation field is 'horizontally biased'; when the ridge/valleys structures are mainly vertical, the filtering in not so effective. In those cases, a top to bottom scanning would be preferable. Similarly to the longitudinal enhancement, we perform both an "horizontal" and a "vertical" scanning and then these are combined as shown before (see Equation 9.6).

An example of the tranversal enhancement is depicted in Figure 9.8: the result of the enhancement in the vertical direction (Figure 9.8(b)) and in the horizontal direction (Figure 9.8(c)) and the enhanced imaged (Figure 9.8(d)).



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

Figure 9.8: Transversal Directional Smoothing. a) Original image; b) Smothed image after "horizontal" scanning; c) Smoothed image after "vertical" scanning; d) Final smoothed image.

9.4 Datasets

The datasets used to evaluate the proposed method were FVC2002 [157], FVC2004 database set A [31] and FVC2006 [29].

9.5 Experimental results and discussion

To evaluate the proposed methods we used a a goal-directed performance evaluation. A goal-directed performance evaluation assesses the overall improvement in the system performance that incorporates the proposed modules as components. Therefore, it is capable of providing a more reliable assessment of the performance benchmark and is directly associated with the ultimate goal of the system.

We compared the following four fingerprint verification systems:

1. The first one is the public-domain NIST Biometric Image Software (NBIS) [284] combined with the matching Minutia Cylinder-Code (MCC) algorithm [30].

2. The second system is a built in-house robust fingerprint verification software (RFV), combining state of the art algorithms for each of the individual blocks:

   - the region mask was obtained by the method proposed by Ferreira *et al.* [72];
   - the orientation map was extracted with the MINDTCT module from NBIS [284], this one was preferred over other approaches tested that provided worst results;
   - the enhancement (simultaneously in the longitudinal and transversal direction) was performed with a Gabor filter using support windows of Hong *et al.* [104];

Table 9.1: Comparison of the 3 first fingerprint verification systems in terms of EER (%). Figures in bold face are the best results.

	FVC2002				FVC2004				FVC2006			
	DB1	DB2	DB3	DB4	DB1	DB2	DB3	DB4	DB1	DB2	DB3	DB4
NBIS	1.72	1.01	8.91	3.80	8.55	10.08	6.55	6.04	17.44	0.73	7.16	5.32
RFV_C	1.46	**0.75**	**3.94**	1.68	4.93	**4.39**	4.87	3.61	15.39	0.60	5.24	4.98
RFV_{MRF}	**1.25**	1.04	4.12	**1.46**	**4.68**	4.44	**4.07**	**2.96**	16.42	1.15	**4.93**	**4.20**

- the adaptive binarization was done by comparing the pixel value with the average along a strip in the transversal direction.
- the skeletonization step was performed by applying a standard thining process based on morphological operations to the binarized image;
- for the minutiae extraction step was applied the method of Ratha *et al.* [217];
- for the matching we resorted to the Minutia Cylinder-Code (MCC) algorithm [30].

3. The third system, denominated as RFV_{MRF}. In this system we replaced the NBIS algorithm by the proposed approach for estimating the orientation field, in the RFV system.

4. Finally, the fourth system: $RFV_{(MRF+CSP)}$. To obtain this one, we replaced the orientation map module by our proposed method, MRF, and the Gabor filter enhancement module by the Curved Signal Processing (CSP) in the longitudinal and transversal directions, in the RFV system.

Proposed orientation map estimation method

In Table 9.1, we present the results of the goal-directed performance evaluation of the first three systems.

Observing the results presented in Table 9.1, we observe that the proposed method, RFV_{MRF}, leads to results that outperforms the inhouse system in 7 out of 12 datasets evaluated. From these experimental results, we observe that the performance of the fingerprint verification systems is significantly improved when our orientation field is integrated in the system.

In order to incorporate the proposed algorithm into an on-line fingerprint verification system, the whole process should not take more than a few (3) seconds. Currently RFV_{MRF}, implemented in Matlab without time efficiency concerns, is taking around 20 seconds. We aim for a 10 fold improvement in future work.

Proposed enhancement method

The purpose of a fingerprint enhancement algorithm is to make the images more suitable for the minutiae extraction algorithm. A direct criterion for evaluating such binarization is to assess the quality of the binarized images, either comparing with manually binarized images or by using some goodness index. The subjective assessment by visual inspection is useful but is confined to

small number of cases. Illustrative examples are shown in Figure 9.9 and show that the proposed method did binarize properly the original images.

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

Figure 9.9: Examples of Binarized Images. a) Original image; b) Binarized version of image "a)"; c) Original image; d) Binarized version of image "c)".

As already mentioned, we choose to evaluate the proposed modules through an assessment of the overall performance of the system since this allow to assess the improvement of the performance benchmark in a more reliable way.

In Table 9.2, we present the results of the goal-directed performance evaluation of the four systems.

Table 9.2: The comparison of the 4 fingerprint verification systems in terms of EER (%). Figures in bold face are the best results.

	FVC2002				FVC2004				FVC2006			
	DB1	DB2	DB3	DB4	DB1	DB2	DB3	DB4	DB1	DB2	DB3	DB4
NBIS	1.72	1.01	8.91	3.80	8.55	10.08	6.55	6.04	17.44	0.73	7.16	5.32
RFV_C	1.46	**0.75**	**3.94**	1.68	4.93	**4.39**	4.87	3.61	**15.39**	**0.60**	5.24	4.98
RFV_{MRF}	**1.25**	1.04	4.12	**1.46**	**4.68**	4.44	**4.07**	**2.96**	16.42	1.15	**4.93**	**4.20**
$RFV_{MRF+CSP}$	**1.11**	1.36	4.91	2.89	5.86	6.21	5.46	3.79	18.71	1.83	6.11	6.39

From the experimental results of Table 9.2, we observe that the performance of the fingerprint verification systems is improved in some subsets when our enhancement method is integrated in the system. Furthermore, it can be noted that the performance of the fingerprint verification systems is improved in almost all subsets when our enhancement method is integrated in the system relatively to the existing NBIS system results.

Improvements can still be made in the enhancement step focusing on aspects which are still susceptible of further analysis. We stress the fact that the results obtained are still an improvement regarding the current state-of-the-art for most databases evaluated.

9.6 Conclusions and future work

We found it hard to compare directly our proposed methods with recent methods due to the fact that most state-of-the-art methods found in literature are only evaluated in some databases or even in some subsets of some of the databases or, in another perspective, different evaluation metrics

are used (for example evaluating the performance in minutiae detection [302] or other indirect ways). Nevertheless, we observed that the proposed method for estimating the orientation map improved the overall performance of our inhouse fingerprint verification system and beyond doubt outperformed the NBIS algorithm. For the enhancement proposed method, it was observed a number of improvements relatively to this latter. We expect that improvements can be done in future works such as applying loopy believe propagation in the markov random field for choosing the penalty functions; using a quadratic approximation for the orientation map estimation and the enhancement step; or using recursive Gabor filter for the enhancement step; among other possible ones.

The work developed in fingerprint robust recognition can be seen as a first approach to the research with this biometric trait. The work developed in this thesis was tested in existing databases composed by images captured with traditional sensors, nevertheless a new acquiring scenario has to be taken in account since the inclusion of fingerprint authentication systems in mobile devices has been observed recently. We believe that soon the research will broad and new databases will appear allowing to evaluate the methods developed in more suitable images. Although the rising of new applications, it is also our conviction that fingerprint recognition systems will remain to be one of the first choice systems in more traditional approaches so the fundamental research will never loose its applicability.

Part IV

Conclusions and Future work

Chapter 10

Conclusions and Future Work

10.1 Conclusion

Biometrics is a research field with some decades of existence and a significant amount of research. Nevertheless, the interest it raises not only has not decreased but in fact has been increasing and gaining more and more relevance in daily life.

The importance of personal recognition is of utmost importance in our most common activities. The fact that we live in a society increasingly networked in which our devices contain not only personal information but also professional and financial just like all sorts of details of the daily life of the user instigates the search to reinforce the security of all this data.

Biometrics play its role as a way of identification which relies on behavioral and anatomical properties of the individual and not on knowledge which he can forget or possessions that he can loose or can be stolen. The spread of biometric applications is impossible to overlook as we see this systems appearing in airport control, bank management, military applications and access control of our handheld devices, among several others.

Therefore, new scenarios appeared in which the acquisition conditions are much less controlled and consequently the data collected will be much more noisy. The characteristics observed in the biometric data captured in these new and challenging scenarios demand new biometric tools adequate to the more defying conditions.

Not only in this scenario but mainly in the recent field of Mobile Biometrics a effort to develop new methods as to be performed in order to guarantee a satisfactory minimum efficacy in the biometric applications deployed.

In spite of the enormous variety of biometric traits suitable for biometric recognition, some are specially well shaped for the mobile scenario. Due to the spread of high quality cameras in these devices as well as voice recorders face, iris, periocular zone and voice have been used in this contexts. Nevertheless, we observe recently a growing investment in fingerprint sensors in mobile devices probably due to the comproved efficacy of this biometric trait.

This work had as a first focus the unconstrained scenario in iris biometric recognition which motivated a collaboration in a the development of a segmentation method for noisy iris images.

Pursuing the new imaging scenarios, a multimodal database was constructed acquired exclusively with a handheld device. This database was used as a benchmark in a Multimodal Biometric Competition and motivated the construction of a database for iris liveness detection. This problem became then the main focus of the work developed. This iris liveness detection database was used as a benchmark in an iris liveness detection competition we organized and was then used to evaluated new methodologies for iris liveness detection. Then, the focus of attention shifted to a new biometric trait, the fingerprint. The work developed in fingerprint liveness detection constituted a new approach to this problem in a more realistic way than it has been performed in traditional approaches. When considering spoofing attacks, usually the evaluation of liveness methods was done regarding one specific type of material used to build fake samples although in a scenario more close to a real attack we should not assume to be able to predict which type of fake materials was to be used. Observing this optimism in traditional approaches a new one was proposed in which we use the information of the real samples and try to model the fake samples as the ones less likely to be classified as real by our trained model. Another aspect of the work developed was the proposition of a robust fingerprint binarization method including the estimation of the direction map using Markov Random Fields. We note, the available databases regarding fingerprint do not contemplate yet the new imaging scenarios but it is our conviction that in a near future new databases will become available for evaluate methods more adapted to the new trend observed. In a way the work developed intends to be a first approach to the fingerprint liveness detection and recognition problem with the objective of broadening the scenario to the mobile fingerprint biometrics scenario as done regarding iris.

It can be considered that contributions were made in these different steps of liveness and recognition, regarding iris or fingerprint. Also the construction of datasets and organization of biometric competitions may be considered as great value contributions to the scientific community. On one hand, providing benchmarks that allow fair comparisons for the methods developed and on the other hand, stimulating the development of new methods through a healthy competition.

10.2 Future Work

Several questions remain open after the work developed in this thesis that could be pursued in a future work.

The iris liveness detection problem was considerably explored and nowadays excellent methods with excellent results are found in the literature. The following step, in our opinion, is to broad even more the acquisition scenarios and perform the classification assuming less knowledge about fake samples in the train phase.

This latter approach is, in our opinion, the path to pursue as well in the fingerprint scenario. The proposed methodology showed to have some limitations when compared to the traditional approach and doubtless new classification tools may be considered to achieve better results with the more realistic approach.

Regarding the fingerprint recognition methods, even though the proposed method for estimating the orientation map improved the results of our in-house recognition system there are still some limitations to overcome. Besides outperforming the inhouse method in more subsets, also the computational time can still be improved. Concerning the enhancement step, this proposed methodology showed some promise but we are convinced that more can be done to achieve more significant improvements.

Appendix A

Robust Iris Segmentation in Unconstrained Settings*

Here is presented with more details the method for iris segmentation used in the work developed. In this work we focus on *iris segmentation*, as proposed by Daugman in his 1993 pioneer work [52]. Iris segmentation consists on the detection of the two defining contours of the iris region: the *limbic contour* separates the iris from the sclera, and the *pupillary contour*, the iris from the pupil. The detection of these contours is the main goal of segmentation and an essential step in the development of high accuracy recognition systems.

We argue that iris segmentation can benefit from the simultaneous detection of the iris centre and iris external contour. When performed independently, both tasks are nontrivial since many other parts of the image may be falsely detected. However, the two tasks can benefit greatly from serving as context for each other. Central to our method to detect iris centre candidates is the use of gradient flow information with a specific gradient vector field template; the detection of the limbic contour relies on the search of strong closed contours around the centre candidates. Further context information can be used to localise the pupil region in the areas adjacent to the centroid of the segmented limbic contour.

This work extended an initial work [174] in which a method for the detection of external iris contour was proposed. The main contributions of the present method are the detection and evaluation of the pupillary contour and the evaluation of the method in an iris database with images captured with an handheld device [234].

A.1 Iris Segmentation methods - Literature Review

When analysing most of the methods cited in the literature, it is possible to detect some main drawbacks. In almost all of these methods, inner and outer boundaries, eyelashes and eyelids are detected in different steps, causing a considerable increase in processing time of the system. Usually, the inner and outer boundaries are detected by circle fitting techniques. This is a source

*Some portions of this Chapter appeared in [174, 175]

of error, since the iris boundaries are not exactly circles and in noisy situations, the outer boundary of iris may not present sharp edges [20].

In some of the aforementioned algorithms, there are a lot of implicit or explicit assumptions about the acquisition process, which are no longer valid in unconstrained acquisition scenarios. Therefore, some of the promising results reported in the literature must be taken with caution and reassessed under these new, more challenging, conditions.

In recent years it has been recognized that the path forward, regarding iris recognition, is the development of algorithms that can work independently of subject collaboration and proper NIR illumination conditions, in order to achieve robust (i.e. accurate even with noisy images) and unconstrained (i.e. accurate for several sets of acquisition conditions: distance, movement, illumination) iris recognition and, in this way, become a real-world applicable method [221]. This paradigm shift led to the rise of new trends in the research of iris recognition, for example, exploring VL illumination instead of NIR.

A.2 Joint detection of iris centre and limbic contour

Researchers are now paying more attention to the context to aid visual recognition processes. Context plays an important role in recognition by the human visual system, with many important visual recognition tasks critically relying on it.

The proposed work aimed to accomplish accurate iris segmentation by using simultaneously acquired information from two main sources: *iris centre* and *limbic contour*. Both sources contribute to discriminate between a series of iris segmentation candidates. Context information regarding typical iris characteristics in eye images, namely colour and shape, represented the basis of the developed algorithm. By using more than a single source of information, we aimed to lower the misdetection of areas likely to be wrongly segmented, such as eyebrows and glass frames.

A.2.1 Algorithm overview

A simplification is adopted in relation to the main rationale outlined above. The simultaneous detection of the iris centre and limbic contour will be addressed by first over-detecting centre candidates, followed by a contour detection around each of them.

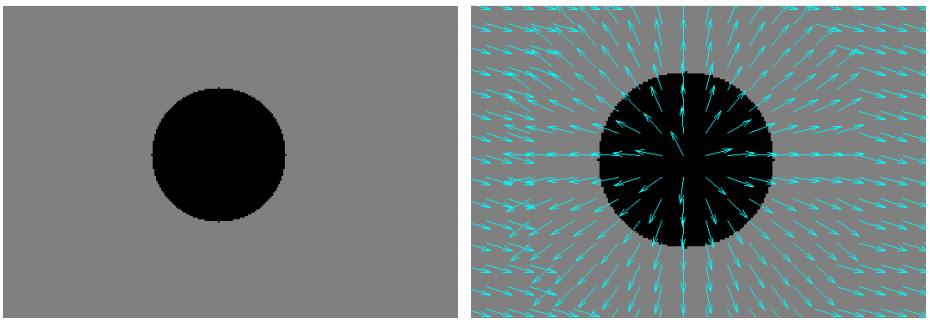
The centre candidates are estimated using a convergence index filter methodology [131]. Next, a window centred in each candidate is converted into the polar domain followed by shortest path algorithm to determine good closed paths around the centre. Using combined data from the centre and respective contour, the best pair centre/contour is selected. Finally, the pupillary segmentation is performed using a new polar image around the centroid of the detected limbic contour.

Typical iris images present two very distinct regions: a high intensity region corresponding to the eye and the skin, and the iris region, at least *partially circular* and *lower in intensity*. These two sources of knowledge can be presented separately but are intrinsically connected. The fact that the iris is a darker region against a brighter background translates into a specific divergent

gradient orientation from its centre. At the same time the limbic contour (iris outer edge) will present a high gradient magnitude as well as a closed shape. The approach taken in this work was that of detecting pairs of iris centre and limbic contour candidates that maximise a quality factor weighted by the aforementioned combined knowledge. The segmentation of the pupillary contour is then performed in a limited region of interest, concentric with the previously segmented limbic contour.

A.2.2 Iris centre detection

Iris centre candidates are detected using a template matching algorithm based on gradient vector field orientation. Theoretically the gradient is a vector field that points in the direction of the greatest rate of increase of a scalar field. Considering an image as a scalar field, it is easy to perceive the gradient as a vector field that points from darker regions (of lower intensity) towards brighter regions (of higher intensity). Fig. A.1(b) depicts a simple example of gradient vector field orientation on a synthetic image.



(a) (b)

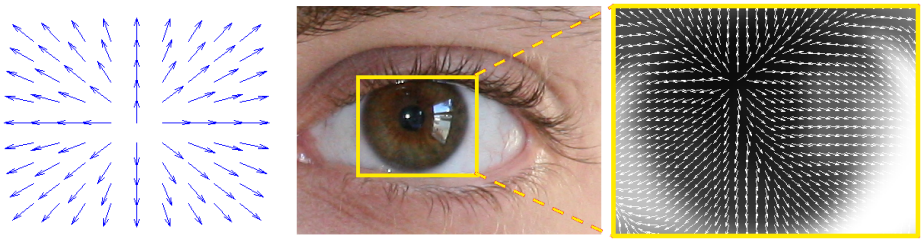
Figure A.1: Gradient orientation vector field in synthetic images. Notice how the vector field diverges from darker regions and converges to brighter regions.

The iris is surrounded by two distinct higher intensity regions: the sclera and the skin. With this in mind a divergent gradient orientation is expected from the center of the iris towards the aforementioned brighter regions, as observed in Fig. A.2(b).

The centre candidates are, thus, detected by computing the cross-correlation, c_{corr}, between the gradient vector field orientation and the divergent template vector field depicted in Fig. A.2(a). The c_{corr} values are calculated as:

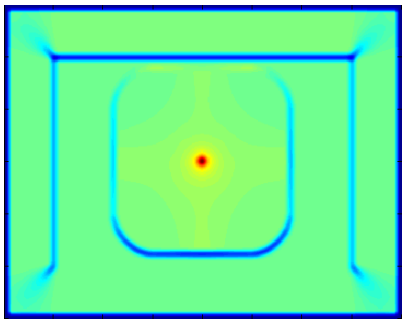
$$c_{corr} = (f * g)[\mathbf{n}] \stackrel{def}{=} \sum_{\mathbf{m}} f^*[\mathbf{m}] g[\mathbf{n}+\mathbf{m}] \tag{A.1}$$

where f and g represent the gradient orientation vector field and the template vector field respectively. The resulting c_{corr} matrix can be graphically represented as exemplified in Fig. A.3(a), where the values range from -1 to 1, with -1 being represented in blue and 1 in red. The centre candidates are detected as the N local maxima with the highest c_{corr} values.



(a) (b)

Figure A.2: The iris centre detection is based on two vector fields: a) Template vector field and b) Gradient orientation vector field.



(a)

Figure A.3: Cross-correlation results on the synthetic image from Fig. A.1(a).

A.2.3 Limbic contour detection

In the proposed method for limbic boundary detection we consider the image grid as a graph, with pixels as nodes and edges connecting neighbouring pixels. With this in mind the proposed algorithm defines a limbic contour candidate as the best closed contour around a given centre candidate.

The computation of this best contour is simplified by working in polar coordinates (relative to each iris centre candidate). In this domain a closed contour around a given point becomes a curve from the left side of the polar image ($\theta = 0^\circ$) to the right side of the same image ($\theta = 360^\circ$). With the aforementioned consideration of the image as a graph, computation of the best closed contour becomes computation of the shortest left-to-right path in polar domain. To better understand the proposed limbic contour detection algorithm, we start by introducing some graph concepts [193].

A.2.3.1 Graph Concepts

A *graph* $G = (V,A)$ is composed of two sets V and A. V is the set of nodes, and A the set of arcs (p,q), $p,q \in V$. The graph is *weighted* if a weight $w(p,q)$ is associated to each arc. The weight of each arc, $w(p,q)$, is a function of pixels values and pixels relative positions. A path from vertex (pixel) v_1 to vertex (pixel) v_n is a list of unique vertices $v_1, v_2, \ldots, v_n$, with v_i and v_{i+1}

corresponding to neighbour pixels. The total cost of a path is the sum of each arc weight in the path $\sum_{i=2}^{n} w(v_{i-1}, v_i)$.

A path from a source vertex v to a target vertex u is said to be the *shortest path* if its total cost is minimum among all v-to-u paths. The distance between a source vertex v and a target vertex u on a graph, $d(v,u)$, is the total cost of the shortest path between v and u.

A path from a source vertex v to a sub-graph Ω is said to be the shortest path between v and Ω if its total cost is minimum among all v-to-$u \in \Omega$ paths. The distance from a node v to a sub-graph Ω, $d(v,\Omega)$, is the total cost of the shortest path between v and Ω:

$$d(v,\Omega) = \min_{u \in \Omega} d(v,u). \tag{A.2}$$

A path from a sub-graph Ω_1 to a sub-graph Ω_2 is said to be the shortest path between Ω_1 and Ω_2 if its total cost is minimum among all $v \in \Omega_1$-to-$u \in \Omega_2$ paths. The distance from a sub-graph Ω_1 to a sub-graph Ω_2, $d(\Omega_1,\Omega_2)$, is the total cost of the shortest path between Ω_1 and Ω_2:

$$d(\Omega_1,\Omega_2) = \min_{v \in \Omega_1, u \in \Omega_2} d(v,u). \tag{A.3}$$

A.2.3.2 Algorithm for limbic contour detection

Intuitively, the limbic boundary appears as a closed contour in the image, enclosing the iris centre, and over pixels with a strong transition in the grey-level values. Assuming that paths through pixels with high directional derivative are preferred over paths through low directional derivative pixels, the limbic contour can then be found among the shortest closed paths enclosing the iris centre candidate.

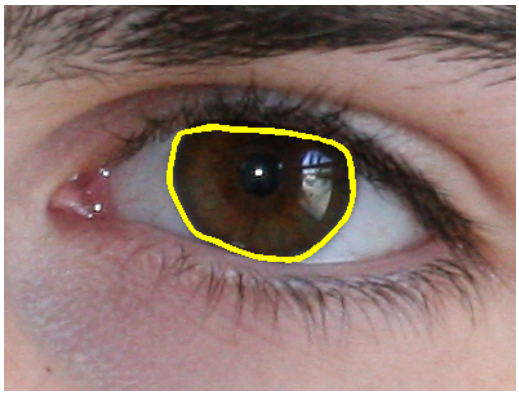
A difficulty with searching for the shortest closed path enclosing a given point C is that small paths, collapsing in the point C, are naturally favoured. We overcome that difficulty by working on polar coordinates.

A circular window centred in each candidate is transformed to polar coordinates. A closed path in the original Cartesian coordinates (Figure A.4(a)) is transformed into a path from left to right margins in the window in polar coordinates, starting and ending in the same row of the transformed window (Figure A.4(b)).

Note that the main assumptions are a) the candidate centre lies within the true limbic contour; b) the limbic contour constitutes a closed path over pixels of strong directional derivative. The limbic contour is not necessarily circular and the candidate centre does not need to match the true iris centre for a correct contour detection.

A.2.3.3 Computation of the Shortest Closed Path

In spite of the efficiency of the computation of the shortest path between the whole left and right margins, or between two pre-defined points in the margins, or between one of the margins and a pre-defined point in the other margin, the search for the shortest path between the left and right margins with the constraint that the path should start and end in the same row seems to increase



(a)

(b)

Figure A.4: a) Original limbic contour in Cartesian coordinates; b) corresponding left-to-right path in the polar domain.

the complexity of the procedure. As typical, optimization with constraints is more difficult than without.

Had one been interested in the simple shortest path between the left and right margin and the computation would be very efficiently performed using dynamic programming. Assuming the simplifying assumption that the vertical paths do not zigzag back and forth, up and down, in the transformed image, the search may be restricted among connected paths containing one, and only one, pixel in each column between the two end-columns.

Formally, let I be an $N_1 \times N_2$ window (after polar coordinate transform) with N_1 columns and N_2 rows; define an admissible path to be

$$\mathbf{s} = \{(x, y(x))\}_{x=1}^{N_1} \text{ , s.t. } \forall x \; |y(x) - y(x-1)| \leq 1,$$

where y is a mapping $y : [1, \cdots, N_1] \rightarrow [1, \cdots, N_2]$. That is, an admissible path is an 8-connected path of pixels in the image from left to right, containing one, and only one, pixel in each column of the image.

The first step is to traverse the image from the second column to the last column and compute the cumulative minimum cost C for each entry (i, j):

$$C(i,j) = \min \begin{cases} C(i-1,j-1) & + & w(p_{i-1,j-1}; p_{i,j}) \\ C(i-1,j) & + & w(p_{i-1,j}; p_{i,j}) \\ C(i-1,j+1) & + & w(p_{i-1,j+1}; p_{i,j}) \end{cases},$$

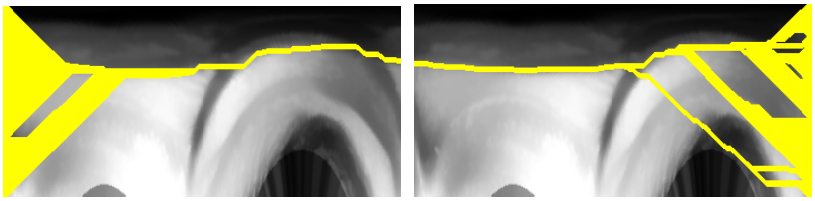
where $w(p_{i,j}; p_{l,m})$ represents the weight of the edge incident with pixels at positions (i, j) and

(l,m). At the end of this process,

$$\min_{j\in\{1,\cdots,N_2\}} C(N_1,j)$$

indicates the end of the minimal connected path. Hence, in the second step, one backtrack from this minimum entry on C to find the optimal path.

Note that this procedure gives not only the shortest path between the left and right margins but also yields the shortest path between any point in the right margin and the whole left margin: for any point (N_1,j) in the right margin, $C(N_1,j)$ indicates the cost of the shortest path between (N_1,j) and the whole left margin, see Figure A.5. Finally, it should be clear how to change the initial conditions of the above procedure to yield the shortest path between two pre-defined points in the opposite margins.



(a) (b)

Figure A.5: Example of shortest path starting point detection. (a) shows all paths from the left margin to the right margin and (b) all the paths from the right margin to the left margin. As is easily deductable, at least one closed contour will result from this process.

Noting that if j and ℓ are two distinct points in the right margin, then the shortest paths between each of these points and the whole left margin do not intersect, it is trivial to conclude that there is at least one point m in the right margin for which the shortest path between m and the whole left margin starts also at row m. Note that the paths correspond to closed paths in the original window in Cartesian coordinates (not necessarily including the shortest one). Similarly, interchanging the role of the left and right margin, it is possible to obtain at least one point n in the left margin for which the shortest path to the whole right margin is closed. By computing all the paths from the left to the right margin (and vice-versa), a set of k closed contours is obtained for each centre candidate. The procedure is illustrated in Figure A.5.

A.2.3.4 Design of the Weight Function

The weight of an edge in the graph is a function of the values of the incident nodes (pixels). We start by computing the derivative in the radial direction (centred in the iris candidate position) in the original space, using a 3-point numerical differentiation, as defined in Eq. (A.4).

$$G_\theta(r) = \frac{I(r+h) - I(r-h)}{2h} \tag{A.4}$$

In the graph, to each edge incident with 4-neighbouring pixels correspond a weight determined by the derivative value of the two incident pixels, expressed as an exponential law, presented in Eq. (A.5).

$$f(g) = f_\ell + (f_h - f_\ell)\,\frac{\exp(\beta\,(255-g)) - 1}{\exp(\beta\,255) - 1} \tag{A.5}$$

with $f_\ell = 2, f_h = 32, \beta = 0.0208$ and g is the minimum of the derivative computed on the two incident pixels. For 8-neighbour pixels the weight was set to $\sqrt{2}$ times that value. The parameter β was experimentally tuned using a grid search method. The remaining parameters were manually optimised in some of our previous works [193].

A.2.4 Best pair centre/contour

From the previously described steps a set of centre/contour candidate pairs (Cp) is built. An example of such candidate pairs is depicted in Figure A.6, where the yellow circles represent the centres and the purple curves the limbic contour candidates.

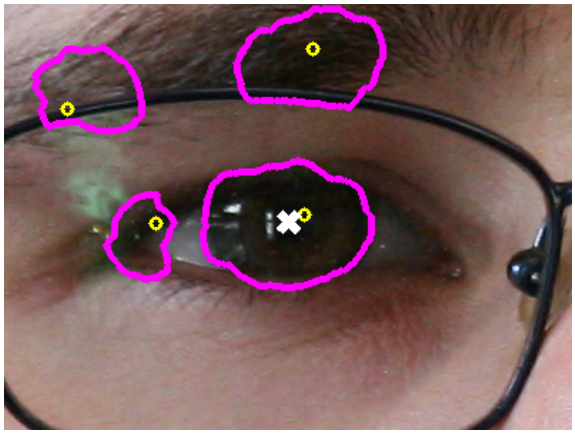


Figure A.6: Example of the centre/contour set of candidates. The centre candidates are represented by yellow circles, the detected contours by purple curves and the ground-truth iris centre by a white cross.

The joint decision for the centre and contour is taken to maximise the joint probability of the individual parts. In here, we assume that the joint probability is a monotonous function of the product of individual measures of quality, combined in an overall quality factor, Q. The discrimination between candidates is performed by choosing the pair with the highest Q. The quality factor is given by:

$$Q(Cp) = \frac{\mu(\Delta C)\cdot\rho_p}{|1 - S(C)|} \tag{A.6}$$

where $\mu(\Delta C)$ is the mean directional derivative alongside the contour, ρ_p is the cross-correlation value of the centre candidate, and S is the shape factor of the contour (with perimeter P and area

A), given by:

$$S(C) = \frac{4\pi \cdot A}{P^2} \tag{A.7}$$

This way the best centre/contour pair, Cp_Q, is selected based on mutual information from both iris centre and limbic contour quality.

A.2.5 Pupillary contour detection

The detection of the pupillary contour was performed by taking into consideration the context knowledge concerning its relation with the limbic contour. Even though the iris and the pupil are not always exactly concentric structures [56], one can approximate the center of the pupil as the centroid of the previously segmented limbic contour. From this point the detection of the pupillary contour becomes a simple repetition of the methodology presented in the former sections. The only novel constraint to this new problem concerns the size of the structure of interest. As the pupil is a contractile structure, whose size is controlled by the lihgtning conditions of the acquisition environemnt, we set the range of possible pupil sizes as $[\frac{1}{4}, \frac{2}{3}]$ of the size of the iris [142]. We then proceed to compute the best closed path around the new centre considering only a limited region of interest in the polar domain.

A.3 Experimental Setup and Results

A.3.1 Tested dataset

The proposed algorithm was tested on the iris subset of the MobBIO multimodal database [234]. The proposed algorithm was already tested on the UBIRIS.v2 iris image database [207] on a previous work [174]. Images in UBIRIS.v2 were captured under non-constrained conditions (at-a-distance, on-the-move and on the visible wavelength), with corresponding realistic noise factors.

A.3.2 Iris centre candidate detection

The accuracy of the centre candidate detection step was computed as the minimum Euclidean distance between each center candidate and the manually annotated ground-truth centre. A mean distance of 5.04 ± 3.36 pixels was obtained for the tested dataset. Considering that the mean iris radius was 33.34 ± 6.90 pixels this result might seem not that promising. The observed deviations of the center candidates from the real iris center arise mainly from two causes: a) the partial occlusion of the iris by the eyelids results in a deviation from an ideal circular shape and b) the extent to which specular reflections contaminate the iris region causes the gradient flow to diverge towards those regions instead of the sclera.

However, given how the limbic contour detection algorithm is designed, there is no need to achieve perfect accuracy on the real iris centre with any of the detected candidates. As long as one of the candidates lies inside the iris/pupil region, the detection of a closed contour around it (not

necessarily centred on it) is guaranteed. For the puppilary contour it is assumed that an accurate limbic contour detection was performed *a priori*, so that the new centroid lies inside the pupil, allowing the best closed contour detection as explained above.

A.3.3 Best centre/contour pair discrimination

The discriminative performance of the proposed quality factor, $Q(Cp)$, was analysed by computing the misdetection ratio, M_r. This value corresponds to the ratio between the number of images where the best centre/contour pair was not correctly discriminated and the total number of tested images. We have shown in our previous work [174] that mutual context information improves results obtained by singular sources of information. Working with the MobBIO dataset a M_r value of 0.88% was obtained, confirming the powerful discriminatory ability of the designed quality factor.

A.3.4 Limbic contour segmentation errors

To evaluate the segmentation accuracy of both the previously discriminated best limbic contour candidates and its respective puppilary contour, a series of metrics were computed. Table A.1 summarises the most relevant results:

- Mean, median and maximum (Hausdorff) distance, in pixels, between the detected limbic contour and the manually annotated ground-truth;
- The accuracy of segmentation, which corresponds to the ratio of images where the mean deviation between the ground truth and the detected contour did not exceed 10% of the radius of the iris;
- Mean percentage of false iris/pupil (FPR) and false non-iris/non-pupil (FNR) segmented pixels.

The first three measurements refer to point-to-point distances between the two referred contours and their respective ground-truths. Concerning solely the limbic contour, Fig. A.7 presents the histogram describing the distribution of these three metrics for all tested images. The information presented in the histograms show that the segmentation errors are relatively low. The larger Hausdorff distances indicate, however, that contours tend to present a localized behaviour of higher deviation from the ground truth. This can be readily explained if the effect of eyelashes in the upper region of the eye is taken into account. As the eyelashes often present an higher contrast with the skin than the iris with the eyelashes, it is only safe to assume that a directional derivative weighted shortest path algorithm will tend to prefer the eyelash-skin boundary to the iris-eyelash boundary. Such effect can be more easily perceived through the visual analysis of the results presented in Fig. A.8.

The obtained FPR and FNR value leads to some interesting conclusions. A 0.00325 FNR is an excellent indicator that very few iris pixels are classified as non-iris. This means that almost no

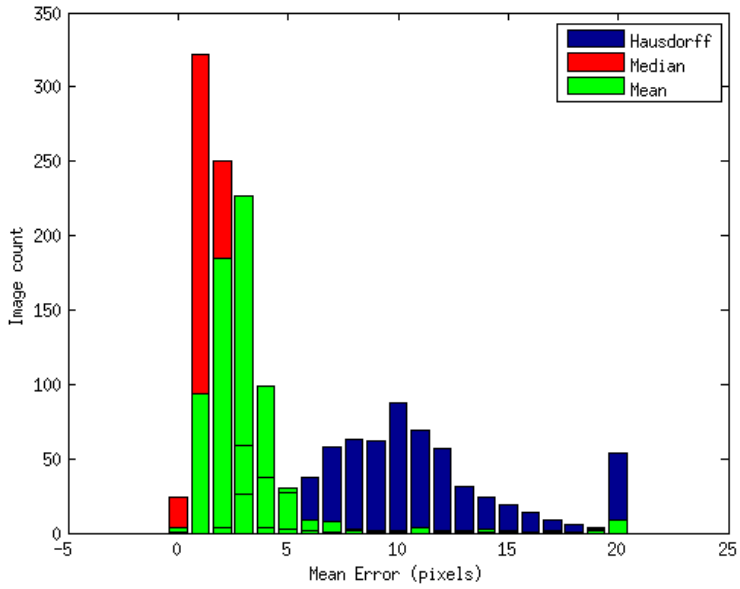


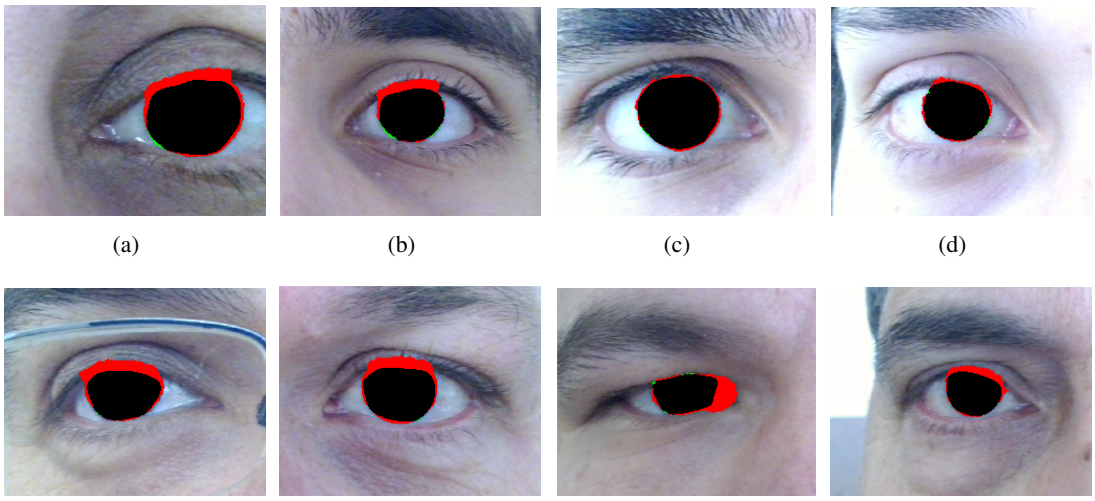
Figure A.7: Histogram of evaluation metric for the limbic contour detection.

Contour	Mean	Median	Hausdorff	Accuracy	*FPR*	*FNR*
Limbic	3.02 ± 1.55	2.08 ± 1.05	10.63 ± 5.46	0.96	0.0191	0.00325
Pupillary	3.90 ± 6.84	3.37 ± 6.86	8.47 ± 8.04	0.044	0.434	0.0085
	Pixels			$[0-1]$		

Table A.1: Summary of the most relevant segmentation evaluation metrics.

useful recognition information is lost *a priori* due to faulty segmentation. The FPR also accounts for what was already referred in the discussion of the Hausdorff distance and observed in Fig. A.8: the high contrast observed in the eyelash region causes a tendency of the closed paths to stick to the eyelash/skin boundary. This yields a strip of false iris pixels above the eye in the eyelash region that justifies the higher FPR values, in comparison with the FNR.

The analysis of the pupillary contour results corroborates some of the expectations regarding the difficulties of its segmentation in VW images. As stated on a previous work [174], the contrast between the pupil and the iris is extremely dependent on a set of hardly controlable factors (illumination, iris pigmentation, obstructions, etc.), thus creating a serious challenges as far as the development of robust segmentation algorithms is concerned. In the present work the obtained results seem to support the aforementioned claim. The 4.40% segmentation accuracy and the 43.4% false pupil rate results are good indicators of how difficult and non-trivial the process of pupillary segmentation is. One could argue that the obtained distance metrics in pixels are not very dissimilar to the values presented for the limbic contour. However, the considerably larger standard deviations, alongside the fact that the average pupil size is approximately one third of the average iris size dismiss such conclusions and only act as to further show how limited the proposed



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

(e) (f) (g) (h)

Figure A.8: Visual representation of the obtained limbic contour segmentation results. The true iris pixels are marked as black, while the false iris are marked red and the false non-iris green.

algorithm is, as far as pupil segmentation is concerned.

A.4 Conclusion

The use of mutual information from gradient orientation for centre detection and directional derivative magnitude for contour detection presented good results for future works. Using the extracted iris regions as inputs for a feature extraction and matching module is the obvious step to carry on after the segmentation algorithm.

However some improvements can be easily suggested to the proposed algorithm. First, improvements on the best centre/contour pair discrimination, so as to improve its robustness, by taking advantage of more powerful machine learning techniques, for example. So as to allow a fairer comparison with other state-of-the-art methods, a noise detection module is necessary. With this new module the number of points that could induce misleading results will be reduced, leading to improved recognition performance. When the noisy conditions of an image limit its usability for recognition tasks, further investigation into quality assessment metrics could improve the global recognition rates of the system. Finally, we argue that a recognition algorithm with no need of pupillary segmentation is probably the way forwarded in unconstrained acquisition settings. However, the same problem that concerns noise detection is applicable to pupil localisation: if the pixels corresponding to this region are not removed from the segmented iris mask, misleading information will be introduced in the recognition module, resulting in loss of accuracy. As accurate segmentation is rendered difficult by the intrinsic characteristics of the MobBIO images, estimating a probability of each pixel belonging to the pupil seems a more robust way of approaching the problem. Future works will certainly focus on these four points of interest.

References

[1] About NEXUS. http://cbsa-asfc.gc.ca/prog/nexus/about-sujet-eng.html.

[2] Ensuring uniqueness: Collecting iris biometrics for the unique ID mission. http://uidai.gov.in/UID_PDF/Working_Papers/UID_and_iris_paper_final.pdf.

[3] Iris recognition - CairoAmman bank. http://www.cab.jo/service-details/61.

[4] Iris recognition finds favor. http://www.bankersonline.com/articles/bhv10n02/bhv10n02a2.html.

[5] Nexus pass application, US immigration visa & travel. http://usa.immigrationvisaforms.com/travel/nexus-pass.

[6] Using iris to enter the uk. http://www.ukba.homeoffice.gov.uk/customs-travel/Enteringtheuk/usingiris/.

[7] Don't blink: Iris recognition for biometric identification. http://www.sans.org/reading-room/whitepapers/authentication/dont-blink-iris-recognition-biometric-identification-1341, 2004.

[8] A. Abhyankar and S. Schuckers. Fingerprint liveness detection using local ridge frequencies and multiresolution texture analysis techniques. In *IEEE International Conference on Image Processing, 2006*, pages 321–324. IEEE, 2006.

[9] A. Abhyankar and S. Schuckers. Iris quality assessment and bi-orthogonal wavelet based encoding for recognition. *Pattern Recognition*, 42(9):1878 – 1894, 2009.

[10] A. Abutaleb and M. Kamel. A genetic algorithm for the estimation of ridges in fingerprints. *IEEE Transactions on Image Processing*, 8(8):1134–1139, 1999.

[11] Z. Akhtar, C. Michelon, and G. L. Foresti. Liveness detection for biometric authentication in mobile applications. In *International Carnahan Conference on Security Technology (ICCST), 2014*, pages 1–6. IEEE, 2014.

[12] Z. Akhtar, C. Micheloni, C. Piciarelli, and G. Foresti. Mobio_livdet: Mobile biometric liveness detection. In *11th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2014*, pages 187–192, Aug 2014.

[13] A. Almansa and T. Lindeberg. Fingerprint enhancement by shape adaptation of scale-space operators with automatic scale selection. *IEEE Transactions on Image Processing*, 9(12):2027–2042, 2000.

[14] P. Almeida. A knowledge-based approach to the iris segmentation problem. *Image and Vision Computing*, 28(2):238–245, 2010.

[15] F. Alonso-Fernandez and J. Bigun. Exploting periocular and rgb information in fake iris detection. In *37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2014*, pages 1354–1359, May 2014.

[16] F. Alonso-Fernandez and J. Bigun. Eye detection by complex filtering for periocular recognition. In *Proc. IWBF*, 2014.

[17] F. Alonso-Fernandez and J. Bigun. Near-infrared and visible-light periocular recognition with gabor features using frequency-adaptive automatic eye detection. *IET Biometrics*, 4:74–89(15), June 2015.

[18] S. Angle, R. Bhagtani, and H. Chheda. Biometrics: A further echelon of security. In *UAE International Conference on Biological and Medical Physics*, 2005.

[19] E. Bailly-Baillire, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariéthoz, J. Matas, K. Messer, F. Porée, and B. Ruiz. The banca database and evaluation protocol. In *In Proc. Int. Conf. on Audio- and Video-Based Biometric Person Authentication (AVBPA03*, pages 625–638. Springer-Verlag, 2003.

[20] N. Barzegar and M. Moin. A new approach for iris localisation in iris recognition systems. In *Proceedings of the International Conference on Computer Systems and Applications*, pages 516 –523, 2008.

[21] A. M. Bazen and S. H. Gerez. Segmentation of fingerprint images. In *ProRISC 2001 Workshop on Circuits, Systems and Signal Processing*, pages 276–280, 2001.

[22] A. M. Bazen and S. H. Gerez. Systematic methods for the computation of the directional fields and singular points of fingerprints. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):905–919, 2002.

[23] J. Bergstra, D. Yamins, and D. D. Cox. Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures. In *Intl. Conf. on Machine Learning*, pages 115 – 123, 2013.

[24] S. Bernard, N. Boujemaa, D. Vitale, and C. Bricot. Fingerprint segmentation using the phase of multiscale gabor wavelets. In *The 5th Asian Conference on Computer Vision, Melbourne, Australia*. Citeseer, 2002.

[25] M. Bordallo López, A. Nieto, J. Boutellier, J. Hannuksela, and O. Silvén. Evaluation of real-time lbp computing in multiple architectures. *Journal of Real-Time Image Processing*, pages 1–22, 2014.

[26] K. Bowyer and J. Doyle. Cosmetic contact lenses and iris recognition spoofing. *Computer*, 47(5):96–98, May 2014.

[27] K. Bowyer, K. Hollingsworth, and P. Flynn. Image understanding for iris biometrics: A survey. *Computer vision and image understanding*, 110(2):281–307, 2008.

[28] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(11):1222–1239, Nov. 2001.

[29] R. Cappelli, M. Ferrara, A. Franco, and D. Maltoni. Fingerprint verification competition 2006. *Biometric Technology Today*, 15(7):7–9, 2007.

[30] R. Cappelli, M. Ferrara, and D. Maltoni. Minutia cylinder-code: A new representation and matching technique for fingerprint recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(12):2128–2141, Dec 2010.

[31] R. Cappelli, D. Maio, D. Maltoni, J. Wayman, and A. Jain. Performance evaluation of fingerprint verification systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(1):3–18, Jan 2006.

[32] V. Chandola, A. Banerjee, and V. Kumar. Anomaly detection: A survey. *ACM Computing Surveys (CSUR)*, 41(3):15, 2009.

[33] C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[34] J.-H. Chang and K.-C. Fan. Fingerprint ridge allocation in direct gray-scale domain. *Pattern Recognition*, 34(10):1907–1925, 2001.

[35] R. Chen, X. Lin, and T. Ding. Iris segmentation for non-cooperative recognition systems. *Image Processing*, 5(5):448 –456, 2011.

[36] X. Chen, J. Tian, J. Cheng, and X. Yang. Segmentation of fingerprint images using linear classifier. *EURASIP J. Appl. Signal Process.*, pages 480–494, Jan. 2004.

[37] Y. Chen, M. Adjouadi, C. Han, J. Wang, A. Barreto, N. Rishe, and J. Andrian. A highly accurate and computationally efficient approach for unconstrained iris segmentation. *Image and Vision Computing*, 28(2):261 – 269, 2010.

[38] J. Cheng and J. Tian. Fingerprint enhancement with dyadic scale-space. *Pattern Recognition Letters*, 25(11):1273–1284, 2004.

[39] S. Chikkerur, A. N. Cartwright, and V. Govindaraju. Fingerprint enhancement using stft analysis. *Pattern Recognition*, 40(1):198–211, Jan. 2007.

[40] P. Chiranjeevi and S. Sengupta. Detection of moving objects using multi-channel kernel fuzzy correlogram based background subtraction. *IEEE Transactions on Cybernetics*, 44(6):870–881, 2014.

[41] N. D. U. Clarkson University and W. U. of Technology. Liveness Detection-iris competition 2013. IEEE BTAS 2013, 2013. http://people.clarkson.edu/projects/biosal/iris/.

[42] N. D. U. Clarkson University and W. U. of Technology. Liveness Detection-iris competition 2013. IEEE BTAS 2013, 2013. http://people.clarkson.edu/projects/biosal/iris/results.php.

[43] D. A. Clausi. An analysis of co-occurrence texture statistics as a function of grey level quantization. In *Can J Remote Sensing*, volume 28, no. 1, pages 45 – 62, 2002.

[44] A. Coates and A. Ng. The Importance of Encoding Versus Training with Sparse Coding and Vector Quantization. In *Intl. Conf. on Machine Learning*, pages 921 – 928, 2011.

[45] P. Coli, G. Marcialis, and F. Roli. Vitality detection from fingerprint images: A critical survey. In S.-W. Lee and S. Li, editors, *Advances in Biometrics*, volume 4642 of *Lecture Notes in Computer Science*, pages 722–731. Springer Berlin Heidelberg, 2007.

[46] C. Cortes and V. N. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273 – 297, 1995.

[47] S. Crihalmeanu, A. Ross, R. Govindarajan, L. Hornak, and S. Schuckers. A centralized web-enabled multimodal biometric database. In *Biometric Consortium Conference (BCC), Crystal City, Virginia*, 2004.

[48] S. Crihalmeanu, A. Ross, S. Schuckers, and L. Hornak. A protocol for multibiometric data acquisition, storage and dissemination. Technical report, WVU, Lane Department of Computer Science and Electrical Engineering, 2007.

[49] A. Czajka. Database of iris printouts and is application: Development of liveness detection method for iris recognition. In *The 18th International Conference on Methods and Models in Automation and Control (MMAR 2013)*, pages 28–33. IEEE, 2013.

[50] J. P. Da Costa, F. Le Pouliquen, C. Germain, and P. Baylou. New operators for optimized orientation estimation. In *International Conference on Image Processing, 2001. Proceedings.*, volume 3, pages 744–747. IEEE, 2001.

[51] S. C. Dass. Markov random field models for directional field and singularity extraction in fingerprint images. *IEEE Transactions on Image Processing*, 13(10):1358–1367, 2004.

[52] J. Daugman. High confidence visual recognition of persons by a test of statistical independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1148–1161, 1993.

[53] J. Daugman. Biometric personal identification system based on iris analysis. 1994.

[54] J. Daugman. Recognizing people by their iris patterns. *Information Security Technical Report*, 3(1):33–39, 1998.

[55] J. Daugman. Anti-spoofing liveness detection. *University of Cambridge, Computer Laboratory, Cambridge.*, 2001.

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

[57] J. Daugman. Demodulation by complex-valued wavelets for stochastic pattern recognition. *International Journal of Wavelets, Multiresolution and Information Processing*, 1(01):1–17, 2003.

[58] J. Daugman. Iris recognition and anti-spoofing countermeasures. In *7-th International Biometrics Conference*, 2004.

[59] J. Daugman. Iris recognition border-crossing system in the UAE. *International Airport Review*, 8(2), 2004.

[60] J. Daugman. New methods in iris recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics,*, 37(5):1167 –1175, 2007.

[61] M. De Marsico, M. Nappi, and D. Riccio. Noisy iris recognition integrated scheme. *Pattern Recogn. Lett.*, 33(8):1006–1011, June 2012.

[62] R. Deriche. Recursively implementing the gaussian and its derivatives. In *Proceedings of the 2nd International Conference on Image Processing*, pages 263–267, 1992.

[63] R. Deriche. Recursively implementating the Gaussian and its derivatives. Research Report RR-1893, INRIA, 1993.

[64] M. Dobes and L. Machala. *UPOL iris image database*, 2004. http://phoenix.inf.upol.cz/iris/.

[65] M. Donahue and S. Rokhlin. On the use of level curves in image analysis. *CVGIP: Image Understanding*, 57(2):185–203, 1993.

[66] J. Doyle and K. W. Bowyer. Notre dame image dataset for contact lens detection in iris recognition, 2004.

[67] A. Erol, U. Halici, and G. Ongun. Feature selective filtering for ridge extraction. In *Intelligent biometric techniques in fingerprint and face recognition*, pages 193–215. CRC Press, Inc., 1999.

[68] M. Fahmy and M. Thabet. A fingerprint segmentation technique based on morphological processing. In *International Symposium on Signal Processing and Information Technology (ISSPIT)*, pages 000215–000220, December 12 - 15 December 2013.

[69] G. Farnebäck and C.-F. Westin. Improving deriche-style recursive gaussian filters. *Journal of Mathematical Imaging and Vision*, 26(3):293–299, 2006.

[70] A. Fazel and S. Chakrabartty. An overview of statistical pattern recognition techniques for speaker verification. *IEEE Circuits and Systems Magazine*, 11(2):62–81, 2011.

[71] J. Feng and A. Cai. Fingerprint representation and matching in ridge coordinate system. In *18th International Conference on Pattern Recognition, 2006. ICPR 2006.*, volume 4, pages 485–488. IEEE, 2006.

[72] P. M. Ferreira, A. F. Sequeira, and A. Rebelo. A fuzzy c-means algorithm for fingerprint segmentation. In R. Paredes, J. S. Cardoso, and X. M. Pardo, editors, *Pattern Recognition and Image Analysis*, volume 9117 of *Lecture Notes in Computer Science*, pages 245–252. Springer International Publishing, 2015.

[73] J. Fierrez-Aguilar, J. Ortega-Garcia, D. Torre-Toledano, and J. Gonzalez-Rodriguez. Biosec baseline corpus: A multimodal biometric database. *Pattern Recognition*, 40(4):1389–1392, 2007.

[74] J. Fierrez-Aguilar, J. Ortega-Garcia, D. Torre-Ttoledano, and J. Gonzalez-Rodriguez. Mcyt baseline corpus: A bimodal biometric database. *IEE Proc.Vis. Image Signal Process.*, 150:395–401, 2003.

[75] R. A. Fisher. The statistical utilization of multiple measurements. *Annals of eugenics*, 8(4):376–386, 1938.

[76] R. Frassetto Nogueira, R. de Alencar Lotufo, and R. Campos Machado. Evaluating software-based fingerprint liveness detection using convolutional networks and local binary patterns. In *IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS) Proceedings*, pages 22–29, 17 October 17 October 2014.

[77] H. Fronthaler, K. Kollreider, and J. Bigun. Pyramid-based image enhancement of fingerprints. In *IEEE Workshop on Automatic Identification Advanced Technologies, 2007*, pages 45–50. IEEE, 2007.

[78] H. Fronthaler, K. Kollreider, and J. Bigun. Local features for enhancement and minutiae extraction in fingerprints. *IEEE Transactions on image processing*, 17(3):354–363, 2008.

[79] K. Fukunaga. *Introduction to statistical pattern recognition (2nd ed.)*. Academic Press Professional, Inc., San Diego, CA, USA, 1990.

[80] J. Galbally, F. Alonso-Fernandez, J. Fierrez, and J. Ortega-Garcia. A high performance fingerprint liveness detection method based on quality related features. *Future Generation Computer Systems*, 28(1):311–321, 2012.

[81] J. Galbally, J. Fierrez, F. Alonso-Fernandez, and M. Martinez-Diaz. Evaluation of direct attacks to fingerprint verification systems. *Telecommunication Systems*, 47(3-4):243–254, 2011.

[82] J. Galbally, J. Fierrez, and J. Ortega-Garcia. Vulnerabilities in biometric systems: attacks and recent advances in liveness detection. *DATABASE*, 1(3):4, 2007.

[83] J. Galbally, J. Ortiz-Lopez, J. Fierrez, and J. Ortega-Garcia. Iris liveness detection based on quality related features. In *5th IAPR International Conference on Biometrics (ICB)*, pages 271–276. IEEE, 2012.

[84] M. Gamassi, V. Piuri, and F. Scotti. Fingerprint local analysis for high-performance minutiae extraction. In *IEEE International Conference on Image Processing, 2005. ICIP 2005.*, volume 3, pages III–265. IEEE, 2005.

[85] S. Garcia-Salicetti, C. Beumier, G. Chollet, B. Dorizzi, J. L. les Jardins, J. Lunter, Y. Ni, and D. Petrovska-Delacrétaz. Biomet: a multimodal person authentication database including face, voice, fingerprint, hand and signature modalities. In *Audio-and Video-Based Biometric Person Authentication*, pages 845–853. Springer, 2003.

[86] M. Garris and M. McCabe. NIST special database 27: Fingerprint minutiae from latent and matching tenprint images. 2000.

[87] L. Ghiani, A. Hadid, G. L. Marcialis, and F. Roli. Fingerprint liveness detection using binarized statistical image features. In *IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–6. IEEE, 29 September - 2 October 2013.

[88] L. Ghiani, G. L. Marcialis, and F. Roli. Fingerprint liveness detection by local phase quantization. In *21st International Conference on Pattern Recognition (ICPR)*, pages 537 – 540. IEEE, 11-15 November 2012.

[89] L. Ghiani, D. Yambay, V. Mura, S. Tocco, G. L. Marcialis, F. Roli, and S. Schuckers. LivDet 2013 Fingerprint liveness detection competition. In *International Conference on Biometrics (ICB)*, pages 1–6. IEEE, 4 - 7 June 2013.

[90] G. GIMP. Image manipulation program. *User Manual, Edge-Detect Filters, Sobel, The GIMP Documentation Team*, 2008.

[91] C. Gottschlich, E. Marasco, A. Y. Yang, and B. Cukic. Fingerprint liveness detection based on histograms of invariant gradients. In *International Joint Conference on Biometrics (IJCB)*, pages 1–7. IEEE, 29 September - 2 October 2014.

[92] D. Gragnaniello, G. Poggi, C. Sansone, and L. Verdoliva. Fingerprint liveness detection based on weber local image descriptor. In *IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS)*, pages 46–50. IEEE, 9 September 2013.

[93] S. Greenberg, M. Aladjem, D. Kogan, and I. Dimitrov. Fingerprint image enhancement using filtering techniques. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 3, pages 322–325. IEEE, 2000.

[94] G. Guo and M. Jones. Iris extraction based on intensity gradient and texture difference. In *IEEE Workshop on Applications of Computer Vision (WACV 2008)*, pages 1 –6, 2008.

[95] P. Gupta, S. Behera, M. Vatsa, and R. Singh. On iris spoofing using print attack. In *22nd International Conference on Pattern Recognition (ICPR), 2014*, pages 1681–1686, Aug 2014.

[96] D. Hale. Recursive Gaussian filters. *CWP-546*, 2006.

[97] R. M. Haralick, K. Shanmugam, and I. H. Dinstein. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions*, (6):610–621, 1973.

[98] S. Haykin. *Neural Networks: A Comprehensive Foundation (2nd Edition)*. Prentice Hall, July 1998.

[99] X. He, S. An, and P. Shi. Statistical texture analysis-based approach for fake iris detection using support vector machines. In *Advances in Biometrics*, pages 540–546. Springer, 2007.

[100] X. He, Y. Lu, and P. Shi. A new fake iris detection method. In *Advances in Biometrics*, pages 1132–1139. Springer, 2009.

[101] Z. He, T. Tan, Z. Sun, and X. Qiu. Toward accurate and fast iris segmentation for iris biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(9):1670 –1684, 2009.

[102] D. S. Hochbaum. Multi-label markov random fields as an efficient and effective tool for image segmentation, total variations and regularization. *Numerical Mathematics: Theory, Methods & Applications*, 6(1):169–198, 2013.

[103] L. Hong, A. K. Jain, S. Pankanti, and R. Bolle. Fingerprint enhancement. pages 202–207, 1996.

[104] L. Hong, Y. Wan, and A. Jain. Fingerprint image enhancement: algorithm and performance evaluation. *Transactions on Pattern Analysis and Machine Intelligence*, 20(8):777–789, 1998.

[105] N. Houhou, A. Lemkaddem, V. Duay, A. Alla, and J.-P. Thiran. Shape prior based on statistical map for active contour segmentation. In *15th IEEE International Conference on Image Processing*, pages 2284 –2287, 2008.

[106] C.-T. Hsieh, E. Lai, and Y.-C. Wang. An effective algorithm for fingerprint image enhancement based on wavelet transform. *Pattern Recognition*, 36(2):303–312, 2003.

[107] S. Huckemann, T. Hotz, and A. Munk. Global models for the orientation field of fingerprints: an approach based on quadratic differentials. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(9):1507–1519, 2008.

[108] International Organization for Standardization. Information Technology - Biometrics - Presentation Attack Detection - Part 3: Testing, Reporting and Classification of Attacks, 2014.

[109] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, and G. E. Hinton. Adaptive mixtures of local experts. *Neural computation*, 3(1):79–87, 1991.

[110] A. Jain, R. Bolle, and S. Pankanti. Introduction to biometrics. In *Biometrics*, pages 1–41. 2002.

[111] A. Jain, L. Hong, and Y. Kulkarni. A multimodal biometric system using fingerprint, face and speech. In *Proceedings of 2nd International Conference on Audio-and Video-based Biometric Person Authentication, Washington DC*, pages 182–187, 1999.

[112] A. Jain, L. Hong, and S. Pankanti. Biometric identification. *Communications of the ACM*, 43(2):90–98, 2000.

[113] A. Jain and D. Zongker. Feature selection: Evaluation, application, and small sample performance. *Pattern Analysis and Machine Intelligence, IEEE Transactions*, 19(2):153–158, 1997.

[114] A. K. Jain, S. Prabhakar, L. Hong, and S. Pankanti. Filterbank-based fingerprint matching. *IEEE Transactions on Image Processing*, 9(5):846–859, 2000.

[115] A. K. Jain and A. Ross. Multibiometric systems. *Communications of the ACM*, 47(1):34–40, 2004.

[116] K. Jarrett, K. Kavukcuoglu, M. Ranzato, and Y. LeCun. What is the Best Multi-Stage Architecture for Object Recognition? In *IEEE Intl. Conf. on Computer Vision*, pages 2146 – 2153, 2009.

[117] D. S. Jeong, J. W. Hwang, B. J. Kang, K. R. Park, C. S. Won, D.-K. Park, and J. Kim. A new iris segmentation method for non-ideal iris images. *Image Vision Comput.*, 28(2):254–260, Feb. 2010.

[118] L. Ji and Z. Yi. Fingerprint orientation field estimation using ridge projection. *Pattern Recognition*, 41(5):1491–1503, 2008.

[119] X. Jiang. Fingerprint image ridge frequency estimation by higher order spectrum. In *International Conference on Image Processing, 2000. Proceedings. 2000*, volume 1, pages 462–465. IEEE, 2000.

[120] X. Jiang, M. Liu, and A. C. Kot. Reference point detection for fingerprint recognition. In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, volume 1, pages 540–543. IEEE, 2004.

[121] X. Jiang, W. Y. Yau, and W. Ser. Minutiae extraction by adaptive tracing the gray level ridge of the fingerprint image. In *International Conference on Image Processing. ICIP 99.*, volume 2, pages 852–856. IEEE, 1999.

[122] X. Jiang, W.-Y. Yau, and W. Ser. Detecting the fingerprint minutiae by adaptive tracing the gray-level ridge. *Pattern Recognition*, 34(5):999–1013, 2001.

[123] S. Jirachaweng and V. Areekul. Fingerprint enhancement based on discrete cosine transform. In *Advances in Biometrics*, pages 96–105. Springer, 2007.

[124] P. Johnson and S. Schuckers. Fingerprint pore characteristics for liveness detection. In *International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–8. IEEE, 10 - 12 September 2014.

[125] I. Jolliffe. Introduction. In *Principal Component Analysis*, pages 1–7. Springer, 1986.

[126] T. Kamei and M. Mizoguchi. Image filter design for fingerprint enhancement. In *International Symposium on Computer Vision, 1995.*, pages 109–114. IEEE, 1995.

[127] M. Kanematsu, H. Takano, and K. Nakamura. Highly reliable liveness detection method for iris recognition. In *SICE, 2007 Annual Conference*, pages 361–364. IEEE, 2007.

[128] M. Kass and A. Witkin. Analyzing oriented patterns. *Computer vision, graphics, and image processing*, 37(3):362–385, 1987.

[129] K. H. Kevin W. Bowyer and P. J. Flynn. A survey of iris biometrics research: 2008-2010. *Handbook of Iris Biometrics*, 2012.

[130] M. Khitrov. Talking passwords: voice biometrics for data access and security. *Biometric Technology Today*, 2013(2):9 – 11, 2013.

[131] H. Kobatake and S. Hashimoto. Convergence index filter for vector fields. *IEEE Transactions on Image Processing*, 8(8):1029 –1038, 1999.

[132] N. Kohli, D. Yadav, M. Vatsa, and R. Singh. Revisiting iris recognition with color cosmetic contact lenses. In *International Conference on Biometrics (ICB), 2013*, pages 1–7, June 2013.

[133] E. Krichen, S. Garcia-Salicetti, and B. Dorizzi. A new phase-correlation-based iris matching for degraded images. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 39(4):924 –934, 2009.

[134] K. Larkin. Uniform estimation of orientation using local and nonlocal 2-d energy operators. *Optics Express*, 13(20):8097–8121, 2005.

[135] S. Lazebnik, C. Schmid, and J. Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2169 – 2178, 2006.

[136] C.-J. Lee, T.-N. Yang, C.-J. Chen, and K.-L. Lin. Direct minutiae matching in gray-level fingerprint images. In *First International Conference on Innovative Computing, Information and Control, 2006. ICICIC'06.*, volume 3, pages 577–580. IEEE, 2006.

[137] E. Lee, K. Park, and J. Kim. Fake iris detection by using purkinje image. In *Advances in Biometrics*, volume 3832 of *Lecture Notes in Computer Science*, pages 397–403. Springer Berlin / Heidelberg, 2005.

[138] K.-c. Lee and S. Prabhakar. Probabilistic orientation field estimation for fingerprint enhancement and verification. In *Biometrics Symposium, 2008. BSYM'08*, pages 41–46. IEEE, 2008.

[139] W. Leung, S. Leung, W. Lau, and A. Luk. Fingerprint recognition using neural network. In *Neural Networks for Signal Processing [1991]., Proceedings of the 1991 IEEE Workshop*, pages 226–235. IEEE, 1991.

[140] J. Li, Y. Wang, T. Tan, and A. K. Jain. Live face detection based on the analysis of fourier spectra. In *Defense and Security*, pages 296–303. International Society for Optics and Photonics, 2004.

[141] J. Li, W.-Y. Yau, and H. Wang. Constrained nonlinear models of fingerprint orientations with prediction. *Pattern Recognition*, 39(1):102–114, 2006.

[142] P. Li, X. Liu, L. Xiao, and Q. Song. Robust and accurate iris segmentation in very noisy iris images. *Image and Vision Computing*, 28(2):246 – 253, 2010.

[143] P. Li, X. Liu, and N. Zhao. Weighted co-occurrence phase histogram for iris recognition. *Pattern Recognition Letters*, 33(8):1000 – 1005, 2012.

[144] P. Li and H. Ma. Iris recognition in non-ideal imaging conditions. *Pattern Recognition Letters*, 33(8):1012 – 1018, 2012.

[145] A. Lindoso, L. Entrena, C. López-Ongil, and J. Liu. Correlation-based fingerprint matching using fpgas. In *IEEE International Conference on Field-Programmable Technology, 2005.*, pages 87–94. IEEE, 2005.

[146] E. Liu, H. Zhao, F. Guo, J. Liang, and J. Tian. Fingerprint segmentation based on an adaboost classifier. *Frontiers of Computer Science in China*, 5(2):148–157, 2011.

[147] J. Liu, Z. Huang, and K. L. Chan. Direct minutiae extraction from gray-level fingerprint image by relationship examination. In *International Conference on Image Processing, 2000.*, volume 2, pages 427–430. IEEE, 2000.

[148] D. G. Lowe. Object recognition from local scale-invariant features. In *7th IEEE International Conference on Computer Vision*, volume 2, pages 1150–1157, 20 - 27 September 1999.

[149] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[150] C. Lu and Z. Lu. Local feature extraction for iris recognition with automatic scale selection. *Image and Vision Computing*, 26(7):935 – 940, 2008.

[151] H. Lu, C. R. Chatwin, and R. C. Young. http://www.irisguard.com/eyebank/. accessed on 14th of march of 2013.

[152] M. A. Luengo-Oroz, E. Faure, and J. Angulo. Robust iris segmentation on uncalibrated noisy images using mathematical morphology. *Image Vision Comput.*, 28(2):278–284, Feb. 2010.

[153] L. Ma, T. Tan, Y. Wang, and D. Zhang. Personal identification based on iris texture analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions*, 25(12):1519–1533, 2003.

[154] L. Ma, T. Tan, Y. Wang, and D. Zhang. Local intensity variation analysis for iris recognition. *Pattern Recognition*, 37(6):1287 – 1298, 2004.

[155] D. Maio and D. Maltoni. Direct gray-scale minutiae detection in fingerprints. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(1):27–40, 1997.

[156] D. Maio and D. Maltoni. Ridge-line density estimation in digital images. In *International Conference on Pattern Recognition*, volume 1, pages 534–534. IEEE Computer Society, 1998.

[157] D. Maio, D. Maltoni, R. Cappelli, J. Wayman, and A. Jain. Fvc2002: Second fingerprint verification competition. In *16th International Conference on Pattern Recognition.*, volume 3, pages 811–814 vol.3, 2002.

[158] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain. Fvc2000: Fingerprint verification competition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3):402–412, 2002.

[159] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain. Fvc2004: Third fingerprint verification competition. In *Biometric Authentication*, pages 1–7. Springer, 2004.

[160] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar. *Handbook of fingerprint recognition*. Springer, 2009.

[161] N. Manivannan, S. Memon, and W. Balachandran. Automatic detection of active sweat pores of fingerprint using highpass and correlation filtering. *Electronics Letters*, 46(18):1268–1269, September 2010.

[162] E. Marasco and A. Ross. A survey on antispoofing schemes for fingerprint recognition systems. *ACM Computing Surveys (CSUR)*, 47(2):28, 2014.

[163] E. Marasco and C. Sansone. On the robustness of fingerprint liveness detection algorithms against new materials used for spoofing. In *BIOSIGNALS*, pages 553–55 – 8, 2011.

[164] G. L. Marcialis, A. Lewicke, B. Tan, P. Coli, D. Grimberg, A. Congiu, A. Tidu, F. Roli, and S. Schuckers. First international fingerprint liveness detection competition - livDet 2009. In *Image Analysis and Processing - ICIAP 2009*, pages 12–23. Springer, 8 - 11 September 2009.

[165] L. Masek. *Recognition of Human Iris Patterns for Biometric Identification. Towards Noncooperative Biometric Iris Recognition*. PhD thesis, 2003.

[166] C. McCool, S. Marcel, A. Hadid, M. Pietikainen, P. Matejka, N. Poh, J. Kittler, A. Larcher, C. Levy, D. Matrouf, et al. Bi-modal person recognition on a mobile phone: using mobile phone data. In *IEEE International Conference on Multimedia and Expo Workshops*, pages 635–640. IEEE, 2012.

[167] B. M. Mehtre and B. Chatterjee. Segmentation of fingerprint images—a composite method. *Pattern recognition*, 22(4):381–385, 1989.

[168] B. M. Mehtre, N. Murthy, S. Kapoor, and B. Chatterjee. Segmentation of fingerprint images using the directional image. *Pattern Recognition*, 20(4):429–435, 1987.

[169] S. Memon, N. Manivannan, A. Boulgouris, and W. Balachandran. Fingerprint sensors: Liveness detection and hardware solutions. In S. Yurish, editor, *Sensors and Biosensors, MEMS Technologies and its Applications*, volume 2 of *Advances in Sensors: Reviews*, pages 121–148. International Frequency Sensor Association Publishing, 2013.

[170] D. Menotti, G. Chiachia, A. Pinto, W. Robson Schwartz, H. Pedrini, A. Xavier Falcao, and A. Rocha. Deep representations for iris, face, and fingerprint spoofing detection. *IEEE Transactions on Information Forensics and Security*, 10(4):864–879, April 2015.

[171] A. A. Moenssens. *Fingerprint techniques*. Chilton Book Company London, 1971.

[172] J. Monteiro. Master's thesis.

[173] J. C. Monteiro and J. S. Cardoso. Periocular recognition under unconstrained settings with universal background models.

[174] J. C. Monteiro, H. P. Oliveira, A. F. Sequeira, and J. S. Cardoso. Robust iris segmentation under unconstrained settings. In *Proceedings of International Conference on Computer Vision Theory and Applications (VISAPP)*, pages 180–190, 2013.

[175] J. C. Monteiro, A. F. Sequeira, H. P. Oliveira, and J. S. Cardoso. Robust iris localisation in challenging scenarios. In *CCIS – Communications in Computer and Information Science*, pages 146–162. Springer-Verlag, 2014.

[176] J. Murari. Detecção de vivacidade em sistemas de reconhecimento de íris. Master's thesis, Faculdade de Engenharia da Universidade do Porto, Portugal, 2013.

[177] I. Nabney. *NETLAB: algorithms for pattern recognition*. Springer, 2002.

[178] M. Nabti and A. Bouridane. An effective and fast iris recognition system based on a combined multiscale feature extraction technique. *Pattern Recognition*, 43(3):868 – 879, 2008.

[179] T. Nakamura, M. Hirooka, H. Fujiwara, and K. Sumi. Fingerprint image enhancement using a parallel ridge filter. In *Proceedings of the 17th International Conference on Pattern Recognition (ICPR 2004).*, volume 1, pages 536–539. IEEE, 2004.

[180] S. Nikam and S. Agarwal. Fingerprint liveness detection using curvelet energy and co-occurrence signatures. In *Fifth International Conference on Computer Graphics, Imaging and Visualisation (CGIV'08)*, pages 217–222. IEEE, 26-28 August 2008.

[181] S. B. Nikam and S. Agarwal. Texture and wavelet-based spoof fingerprint detection for fingerprint biometric systems. In *First International Conference on Emerging Trends in Engineering and Technology (ICETET'08)*, pages 675–680. IEEE, 16-18 July 2008.

[182] S. B. Nikam and S. Agarwal. Gabor filter-based fingerprint anti-spoofing. In *Advanced Concepts for Intelligent Vision Systems*, pages 1103–1114. Springer, 2008.

[183] S. B. Nikam and S. Agarwal. Wavelet energy signature and glcm features-based fingerprint anti-spoofing. In *International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR'08)*, volume 2, pages 717–723. IEEE, 30 - 31 August 2008.

[184] R. Nimkar and A. Mishra. Fingerprint segmentation algorithms: A literature review. *International Journal of Computer Applications*, 95(5):20–24, 2014.

[185] R. Nosaka, Y. Ohkawa, and K. Fukui. Feature extraction based on co-occurrence of adjacent local binary patterns. In *Advances in image & video technology*, pages 82 – 91, 2012.

[186] I. of Automation Chinese Academy of Sciences. Casia iris image database. http://www.cbsr.ia.ac.cn/china/Iris%20Databases%20CH.asp, 2004.

[187] N. I. of Standards and T. (NIST). Iris challenge evaluation. http://www.nist.gov/itl/iad/ig/ice.cfm, 2006.

[188] L. O'Gorman and J. V. Nickerson. Matched filter design for fingerprint image enhancement. In *International Conference on Acoustics, Speech, and Signal Processing, 1988. ICASSP-88.*, pages 916–919. IEEE, 1988.

[189] L. O'Gorman and J. V. Nickerson. An approach to fingerprint filter design. *Pattern recognition*, 22(1):29–38, 1989.

[190] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1):51–59, 1996.

[191] T. Ojala, M. Pietikäinen, and T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.

[192] V. Ojansivu and J. Heikkilä. Blur insensitive texture classification using local phase quantization. In *Image and Signal Processing, Lecture Notes in Computer Science*, volume 5099, pages 236 – 243, 2008.

[193] H. Oliveira, J. Cardoso, A. Magalhaes, and M. Cardoso. Simultaneous detection of prominent points on breast cancer conservative treatment images. In *Proceedings of the 19th IEEE International Conference on Image Processing*, pages 2841–2844, 2012.

[194] H. P. Oliveira and F. Magalhães. Two unconstrained biometric databases. In *Image Analysis and Recognition*, pages 11–19. Springer, 2012.

[195] M. Oliveira and N. J. Leite. A multiscale directional operator and morphological tools for reconnecting broken ridges in fingerprint images. *Pattern Recognition*, 41(1):367–377, 2008.

[196] J. Ortega-Garcia, J. Fierrez, F. Alonso-Fernandez, J. Galbally, M. R. Freire, J. Gonzalez-Rodriguez, C. Garcia-Mateo, J.-L. Alba-Castro, E. Gonzalez-Agulla, E. Otero-Muras, et al. The multiscenario multienvironment biosecure multimodal database (bmdb). *IEEE Transactions on Pattern Analysis and Machine Intelligence,*, 32(6):1097–1111, 2010.

[197] M. Pawar, S. Lokande, and V. Bapat. Iris segmentation using geodesic active contour for improved texture extraction in recognition. *International Journal of Computer Applications*, 47(16):448–456, 2012.

[198] D. Petrovska-Delacrétaz, S. Lelandais, J. Colineau, L. Chen, B. Dorizzi, E. Krichen, M. Anouar-Mellakh, A. Chaari, S. Guerfi, M. Ardabilian, J. D Hose, and B. Ben Amor. The IV2 Multimodal Biometric Database (Including Iris, 2D, 3D, Stereoscopic and Talking Face Data) and the IV2-2007 Evaluation Campaign. In IEEE, editor, *IEEE Second International Conference on Biometrics: Theory, Applications and Systems (BTAS 08)*, pages 1–7, Sept. 2008.

[199] M. Pietikäinen. Image analysis with local binary patterns. In *SCIA*, pages 115–118, 2005.

[200] N. Pinto and D. D. Cox. Beyond simple features: A large-scale feature search approach to unconstrained face recognition. In *IEEE International Conference on Automatic Face and Gesture Recognition*, pages 8 – 15, 2011.

[201] A. Pocovnicu. Biometric security for cell phones. *Informatica Economica*, 13(1):57–63, 2009.

[202] N. Poh and S. Bengio. Database, protocols and tools for evaluating score-level fusion algorithms in biometric authentication. *Pattern Recognition*, 39(2):223–233, 2006.

[203] S. Prabhakar, S. Pankanti, and A. K. Jain. Biometric recognition: Security and privacy concerns. *Security & Privacy, IEEE*, 1(2):33–42, 2003.

[204] H. Proença. Non-cooperative iris recognition: Issues and trends. In *19th European Signal Processing Conference*, pages 1–5, 2011.

[205] H. Proença and L. Alexandre. Ubiris: A noisy iris image database. In *Image Analysis and Processing–ICIAP 2005*, pages 970–977. Springer, 2005.

[206] H. Proença and L. Alexandre. The NICE. I: noisy iris challenge evaluation - part I. In *First IEEE International Conference on Biometrics: Theory, Applications, and Systems*, pages 1–4. IEEE, 2007.

[207] H. Proenca, S. Filipe, R. Santos, J. Oliveira, and L. Alexandre. The ubiris. v2: A database of visible wavelength iris images captured on-the-move and at-a-distance. *IEE Transactions on Pattern Analysis and Machine Intelligence*, 32(8):1529–1535, 2010.

[208] H. Proença. *Towards Non-Cooperative Biometric Iris Recognition*. PhD thesis, 2007.

[209] P. Pudil, J. Novovičová, and J. Kittler. Floating search methods in feature selection. *Pattern recognition letters*, 15(11):1119–1125, 1994.

[210] J. Quionero-Candela, M. Sugiyama, A. Schwaighofer, and N. D. Lawrence. *Dataset Shift in Machine Learning*. The MIT Press, 2009.

[211] A. Radman, K. Jumari, and N. Zainal. Iris segmentation in visible wavelength environment. *Procedia Engineering*, 41:743–748, 2012.

[212] R. Raghavendra and C. Busch. Presentation attack detection algorithm for face and iris biometrics. In *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European*, pages 1387–1391, Sept 2014.

[213] R. Raghavendra and C. Busch. Presentation attack detection on visible spectrum iris recognition by exploring inherent characteristics of light field camera. In *IEEE International Joint Conference on Biometrics, Clearwater, IJCB 2014, FL, USA, September 29 - October 2, 2014*, pages 1–8, 2014.

[214] R. Raghavendra and C. Busch. Robust scheme for iris presentation attack detection using multiscale binarized statistical image features. *IEEE Transactions on Information Forensics and Security*, 10(4):703–715, April 2015.

[215] K. B. Raja, R. Raghavendra, V. K. Vemuri, and C. Busch. Smartphone based visible iris recognition using deep sparse filtering. *Pattern Recognition Letters*, 57(0):33 – 42, 2015.

[216] T. Rakesh and M. Khogare. Survey of biometric recognition system for iris. *International Journal of Emerging Technology and Advanced Engineering*, pages 2250–2459, 2012.

[217] N. K. Ratha, S. Chen, and A. K. Jain. Adaptive flow orientation-based feature extraction in fingerprint images. *Pattern Recognition*, 28(11):1657–1672, 1995.

[218] N. K. Ratha, J. H. Connell, and R. M. Bolle. An analysis of minutiae matching strength. In *Audio-and Video-Based Biometric Person Authentication*, pages 223–228. Springer, 2001.

[219] A. Rattani and A. Ross. Automatic adaptation of fingerprint liveness detector to new spoof materials. In *IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–8. IEEE, 29 September - 2 October 2014.

[220] D. Reynolds. Gaussian mixture models. In *Encyclopedia of Biometrics*, pages 659–663. Springer, 2009.

[221] A. Ross. Iris recognition: The path forward. *Computer*, 43(2):30–35, 2010.

[222] A. Ross and A. K. Jain. Multimodal biometrics: An overview. In *Proceedings of 12th European Signal Processing Conference*, pages 1221–1224, 2004.

[223] K. Roy, P. Bhattacharya, C. Suen, and J. You. Recognition of unideal iris images using region-based active contour model and game theory. In *17th IEEE International Conference on Image Processing*, pages 1705 –1708, 2010.

[224] V. Ruiz-Albacete, P. Tome-Gonzalez, F. Alonso-Fernandez, J. Galbally, J. Fierrez, and J. Ortega-Garcia. Direct attacks using fake images in iris verification. In *Biometrics and Identity Management*, pages 181–190. Springer, 2008.

[225] B. Sabarigiri and D. Suganyadevi. Counter measures against iris direct attacks using fake images and liveness detection based on electroencephalogram (eeg). *World Applied Sciences Journal 29 (Data Mining and Soft Computing Techniques)*, pages 93 – 98, 2014.

[226] C. Sanchez-Avila, R. Sanchez-Reillo, and D. de Martin-Roche. Iris-based biometric recognition using dyadic wavelet transform. *Aerospace and Electronic Systems Magazine, IEEE*, 17(10):3 – 6, 2002.

[227] M. Sandström. Liveness detection in fingerprint recognition systems. 2004.

[228] W. Sankowski, K. Grabowski, M. Napieralska, M. Zubert, and A. Napieralski. Reliable algorithm for iris segmentation in eye image. *Image and Vision Computing*, 28(2):231–237, 2010.

[229] G. Santos and E. Hoyle. A fusion approach to unconstrained iris recognition. *Pattern Recognition Letters*, 33(8):984 – 990, 2012.

[230] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson. Estimating the support of a high-dimensional distribution. *Neural computation*, 13(7):1443–1471, 2001.

[231] R. Seeley, T. Stephens, and P. Tate. Essentials of anatomy & physiology. St. Louis: Mosby-year book, 1996.

[232] M. Sepasian, C. Mares, and W. Balachandran. Vitality detection in fingerprint identification. *Information science and applications*, (4), 2010.

[233] A. F. Sequeira and J. S. Cardoso. Fingerprint liveness detection in the presence of capable intruders. *Sensors*, (accepted on 25th, May, 2015)(0):0–0, 2015.

[234] A. F. Sequeira, J. C. Monteiro, A. Rebelo, and H. P. Oliveira. MobBIO a multimodal database captured with an handheld device. In *Proceedings of International Conference on Computer Vision Theory and Applications (VISAPP)*, pages 133 – 139, 2014.

[235] A. F. Sequeira, J. Murari, and J. S. Cardoso. Iris liveness detection methods in mobile applications. In *Proceedings of International Conference on Computer Vision Theory and Applications*, pages 22 – 33, 2014.

[236] A. F. Sequeira, J. Murari, and J. S. Cardoso. Iris liveness detection methods in the mobile biometrics scenario. In *Proceedings of International Joint Conference on Neural Networks (IJCNN)*, pages 3002–3008, 2014.

[237] A. F. Sequeira, H. P. Oliveira, J. C. Monteiro, J. P. Monteiro, and J. S. Cardoso. Mobilive 2014 - mobile iris liveness detection competition. In *Proceedings of the International Joint Conference on Biometrics (IJCB)*, 2014.

[238] L. Sha, F. Zhao, and X. Tang. Improved fingercode for filterbank-based fingerprint matching. In *International Conference on Image Processing, 2003. ICIP 2003.*, volume 2, pages II–895. IEEE, 2003.

[239] L. Sha, F. Zhao, and X. Tang. Fingerprint matching using minutiae and interpolation-based square tessellation fingercode. In *IEEE International Conference on Image Processing, 2005. ICIP 2005.*, volume 2, pages II–41. IEEE, 2005.

[240] S. Shah and A. Ross. Iris segmentation using geodesic active contours. *IEEE Transactions on Information Forensics and Security,*, 4(4):824 –836, 2009.

[241] L. Shen, A. Kot, and W. Koo. Quality measures of fingerprint images. In *Audio-and Video-based Biometric Person Authentication*, pages 266–271. Springer, 2001.

[242] B. Sherlock. Computer enhancement and modeling of fingerprint images. In *Automatic Fingerprint Recognition Systems*, pages 87–112. Springer, 2004.

[243] B. Sherlock, D. Monro, and K. Millard. Algorithm for enhancing fingerprint images. *Electronics Letters*, 28(18):1720–1721, 1992.

[244] B. Sherlock, D. Monro, and K. Millard. Fingerprint enhancement by directional fourier filtering. In *Vision, Image and Signal Processing, IEE Proceedings-*, volume 141, pages 87–94. IET, 1994.

[245] B. G. Sherlock and D. M. Monro. A model for interpreting fingerprint topology. *Pattern recognition*, 26(7):1047–1055, 1993.

[246] W. Shi, J. Yang, Y. Jiang, F. Yang, and Y. Xiong. Senguard: Passive user identification on smartphones using multiple sensors. In *IEEE 7th International Conference on Wireless and Mobile Computing, Networking and Communications*, pages 141–148, 2011.

[247] J.-H. Shin, H.-Y. Hwang, and S.-I. Chien. Detecting fingerprint minutiae by run length encoding scheme. *Pattern recognition*, 39(6):1140–1154, 2006.

[248] K. Y. Shin, G. P. Nam, D. S. Jeong, D. H. Cho, B. J. Kang, K. R. Park, and J. Kim. New iris recognition method for noisy iris images. *Pattern Recognition Letters*, 33(8):991 – 999, 2012.

[249] Y. N. Singh and S. K. Singh. Vitality detection from biometrics: state-of-the-art. In *2011 World Congress on Information and Communication Technologies (WICT)*, pages 106–111. IEEE, 2011.

[250] L. Soh and C. Tsatsoulis. Texture analysis of sar sea ice imagery using gray level co-occurrence matrices. In *IEEE TGRS*, volume 37, pages 780 – 795, 1999.

[251] C. Sousedik and C. Busch. Presentation attack detection methods for fingerprint recognition systems: A survey. 3:219 – 233, 2014.

[252] S. D. Stearns. On selecting features for pattern classifiers. In *Proceedings of the 3rd International Joint Conference on Pattern Recognition*, pages 71–75, 1976.

[253] R. Stock and C. Swonger. Development and evaluation of a reader of fingerprint minutiae. *Cornell Aeronautical Laboratory, Technical Report CAL no. XM-2478-X-1*, pages 13–17, 1969.

[254] E. Sung, X. Chen, J. Zhu, and J. Yang. Towards non-cooperative iris recognition systems. In *7th International Conference on Control, Automation, Robotics and Vision (ICARCV2002)*, volume 2, pages 990 – 995, dec. 2002.

[255] R. Szewczyk, K. Grabowski, M. Napieralska, W. Sankowski, M. Zubert, and A. Napieralski. A reliable iris recognition algorithm based on reverse biorthogonal wavelet transform. *Pattern Recognition Letters*, 33(8):1019 – 1026, 2012.

[256] B. Tan and S. Schuckers. Liveness detection for fingerprint scanners based on the statistics of wavelet signal processing. In *Computer Vision and Pattern Recognition Workshop (CVPRW'06)*, pages 26–26. IEEE, 17 - 22 June 2006.

[257] B. Tan and S. Schuckers. Spoofing protection for fingerprint scanner by fusing ridge signal and valley noise. *Pattern Recognition*, 43(8):2845–2857, 2010.

[258] C.-W. Tan and A. Kumar. Integrating ocular and iris descriptors for fake iris image detection. In *Biometrics and Forensics (IWBF), 2014 International Workshop on*, pages 1–4. IEEE, 2014.

[259] T. Tan, Z. He, and Z. Sun. Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition. *Image and Vision Computing*, 28(2):223 – 230, 2010.

[260] T. Tan, X. Zhang, Z. Sun, and H. Zhang. Noisy iris image matching by using multiple cues. *Pattern Recognition Letters*, 33(8):970 – 977, 2012.

[261] D. M. J. Tax and R. P. W. Duin. Support vector domain description. *Pattern Recognition Letters*, 20:1191–1199, 1999.

[262] P. Tufegdzic. iSuppli: Smartphone cameras are getting smarter with computational photography; Last check: 06.06.2013, 2013.

[263] M. Une and Y. Tamura. "liveness detection techniques. *IPSJ Magazine*, 47(6):605–608, 2006.

[264] M. University. MMU iris image database. http://pesona.mmu.edu.my/~ccteo/, 2004.

[265] University of Bath. *University of Bath iris image database*, 2004. http://bath.ac.uk/elec-eng/pages/sipg/.

[266] University of Beira Interior. *MICHE I - Mobile Iris Challenge Evaluation*, 2014. http://http://biplab.unisa.it/MICHE/.

[267] University of Salerno. *BIPLab database*, 2015. http://biplab.unisa.it/MICHE/database/.

[268] T. Van der Putte and J. Keuning. Biometrical fingerprint recognition: don't get your fingers burned. In *Smart Card Research and Advanced Applications*, pages 289–303. Springer, 2000.

[269] V. Vapnik. In *The Nature of Statistical Learning Theory*, 1995.

[270] M. Vatsa, R. Singh, and A. Noore. Improving iris recognition performance using segmentation, quality enhancement, match score fusion, and indexing. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics,*, 38(4):1021 –1035, 2008.

[271] D. S. Vernon. Automatic detection of secondary creases in fingerprints. *Optical Engineering*, 32(10):2616–2623, 1993.

[272] P. R. Vizcaya and L. A. Gerhardt. A nonlinear orientation model for global description of fingerprints. *Pattern Recognition*, 29(7):1221–1231, 1996.

[273] U. von Seelen. Countermeasures against iris spoofing with contact lenses. In *Biometric Consortium Conference BC*, 2005.

[274] N.-S. Vu and A. Caplier. Face recognition with patterns of oriented edge magnitudes. *Computer Vision–ECCV 2010*, pages 313–326, 2010.

[275] Q. Wang, X. Zhang, M. Li, X. Dong, Q. Zhou, and Y. Yin. Adaboost and multi-orientation 2d gabor-based noisy iris recognition. *Pattern Recogn. Lett.*, 33(8):978–983, June 2012.

[276] W. Wang, J. Li, F. Huang, and H. Feng. Design and implementation of log-gabor filter in fingerprint image enhancement. *Pattern Recognition Letters*, 29(3):301–308, 2008.

[277] Y. Wang, J. Hu, and D. Phillips. A fingerprint orientation model based on 2d fourier expansion (fomfe) and its application to singular-point detection and fingerprint indexing. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(4):573–585, 2007.

[278] B. Warwante and M. S. Maske. Wavelet based fingerprint liveness detection. *International Journal of Engineering Research and Applications*, 2(2):1643–1645, 2012.

[279] C. Watson. NIST special database 10: Supplemental fingerprint card data (sfcd) for NIST special database 9. 1993.

[280] C. Watson. NIST special database 24 digital video of live-scan fingerprint data. 1998.

[281] C. Watson and C. Wilson. NIST special database 4. *Fingerprint Database, National Institute of Standards and Technology*, 17, 1992.

[282] C. Watson and C. Wilson. NIST special database 9, fingerprint database. *Special Database*, 1992.

[283] C. I. Watson. Special database 14. In *S. Department of Commerce, NIST, Advanced Systems Division, Gaithersburg, Maryland*. Citeseer, 1993.

[284] C. I. Watson, M. D. Garris, E. Tabassi, C. L. Wilson, R. M. McCabe, S. Janet, and K. Ko. User's guide to nist biometric image software (nbis).

[285] H. Wei, M. Guo, and Z. Ou. Fingerprint verification based on multistage minutiae matching. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 2, pages 1058–1061. IEEE, 2006.

[286] Z. Wei, X. Qiu, Z. Sun, and T. Tan. Counterfeit iris detection based on texture analysis. In *ICPR 2008. 19th International Conference on Pattern Recognition.*, pages 1–4. IEEE, 2008.

[287] M. J. Weinberger, G. Seroussi, and G. Sapiro. Loco-i: a low complexity, context-based, lossless image compression algorithm. In *Data Compression Conf.*, pages 140 – 149, 1996.

[288] A. W. Whitney. A direct method of nonparametric measurement selection. *Computers, IEEE Transactions*, 100(9):1100–1103, 1971.

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

[290] C. L. Wilson, C. I. Watson, and E. G. Paek. Combined optical and neural network fingerprint matching. In *AeroSense'97*, pages 373–382. International Society for Optics and Photonics, 1997.

[291] C. Wu and V. Govindaraju. Singularity preserving fingerprint image adaptive filtering. In *Image Processing, 2006 IEEE International Conference on*, pages 313–316. IEEE, 2006.

[292] C. Wu, S. Tulyakov, and V. Govindaraju. Robust point-based feature fingerprint segmentation algorithm. In S.-W. Lee and S. Li, editors, *Advances in Biometrics*, volume 4642 of *Lecture Notes in Computer Science*, pages 1095–1103. Springer Berlin Heidelberg, 2007.

[293] C. Wu, J. Zhou, Z.-q. Bian, and G. Rong. Robust crease detection in fingerprint images. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 2, pages II–505. IEEE, 2003.

[294] J. Wu and J. M. Rehg. Centrist: A visual descriptor for scene categorization. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(8):1489–1501, 2011.

[295] D. Yadav, N. Kohli, J. Doyle, R. Singh, M. Vatsa, and K. Bowyer. Unraveling the effect of textured contact lenses on iris recognition. *Information Forensics and Security, IEEE Transactions on*, 9(5):851–862, May 2014.

[296] D. Yambay, L. Ghiani, P. Denti, G. L. Marcialis, F. Roli, and S. Schuckers. LivDet 2011 fingerprint liveness detection competition. In *5th IAPR International Conference on Biometrics (ICB)*, pages 208–215. IEEE, 29 March - 1 April 2012.

[297] J. Yang, L. Liu, T. Jiang, and Y. Fan. A modified gabor filter design method for fingerprint image enhancement. *Pattern Recognition Letters*, 24(12):1805–1817, 2003.

[298] J. Yang, K. Yu, Y. Gong, and T. Huang. Linear spatial pyramid matching using sparse coding for image classification. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1794 – 1801, 2009.

[299] Y. Yin, Y. Wang, and X. Yang. Fingerprint image segmentation based on quadric surface model. In *Audio-and Video-Based Biometric Person Authentication*, pages 647–655. Springer, 2005.

[300] Y. W. Yun. The '123'of biometric technology. *Synthesis Journal*, 2002, 2002.

[301] T. T. Z. Sun, H. Zhang and J. Wang. Iris image classification based on hierarchical visual codebook. 36(6):1120 – 1133, June 2014.

[302] M. Zahedi and O. Ghadi. Combining gabor filter and fft for fingerprint enhancement based on a regional adaption method and automatic segmentation. *Signal, Image and Video Processing*, 9(2):267–275, 2015.

[303] X. Zhan, Z. Sun, Y. Yin, and Y. Chu. Fingerprint ridge distance estimation: algorithms and the performance. In *Advances in Biometrics*, pages 294–301. Springer, 2005.

[304] H. Zhang, Z. Sun, and T. Tan. Contact lens detection based on weighted LBP. In *20th International Conference on Pattern Recognition (ICPR)*, pages 4279–4282. IEEE, 23 - 26 August 2010.

[305] L. Zhang, Z. Zhou, and H. Li. Binary gabor pattern: an efficient and robust descriptor for texture classification. In *Proc. 19th Int. Conf. on Image Proc.*, pages 81 – 84, 2012.

[306] W. Zhang and Y. Wang. Core-based structure matching algorithm of fingerprint verification. In *Proceedings of 16th International Conference on Pattern Recognition.*, volume 1, pages 70–74. IEEE, 2002.

[307] Y. Zhang and Q. Xiao. An optimized approach for fingerprint binarization. In *International Joint Conference on Neural Networks. IJCNN'06.*, pages 391–395. IEEE, 2006.

[308] Y. Zhang, X. Yang, Q. Su, and J. Tian. Fingerprint recognition based on combined features. In *Advances in Biometrics*, pages 281–289. Springer, 2007.

[309] J. Zhou and J. Gu. A model-based method for the computation of fingerprints' orientation field. *IEEE Transactions on Image Processing*, 13(6):821–835, 2004.

[310] J. Zhou and J. Gu. Modeling orientation fields of fingerprints with rational complex functions. *Pattern Recognition*, 37(2):389–391, 2004.

[311] J. Zhou, C. Wu, Z. Bian, and D. Zhang. Improving fingerprint recognition based on crease detection. In *Biometric Authentication*, pages 287–293. Springer, 2004.

[312] E. Zhu, J. Yin, C. Hu, and G. Zhang. Quality estimation of fingerprint image based on neural network. In *Advances in Natural Computation*, pages 65–70. Springer, 2005.

[313] E. Zhu, J. Yin, C. Hu, and G. Zhang. A systematic method for fingerprint ridge orientation estimation and image segmentation. *Pattern Recognition*, 39(8):1452–1472, 2006.

[314] E. Zhu, J. Yin, and G. Zhang. Fingerprint enhancement using circular gabor filter. In *Image Analysis and Recognition*, pages 750–758. Springer, 2004.

[315] A. Ziegler, E. Christiansen, D. Kriegman, and S. J. Belongie. Locally uniform comparison image descriptor. In *Advances in Neural Information Processing Systems*, pages 1–9, 2012.

[316] J. Zuo and N. Schmid. On a methodology for robust segmentation of nonideal iris images. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics,*, 40(3):703 – 718, 2010.