UNIVERSIDADE DA CORUÑA

Facultade de Informática

Departamento de Computación

PhD Thesis

Automatic system for personal authentication using the retinal vessel tree as biometric pattern

Author: Marcos Ortega Hortas

Supervisor: Manuel Fco. González Penedo

A Coruña, June 2009

UNIVERSIDADE DA CORUÑA

Facultade de Informática

Departamento de Computación

PhD Thesis

Automatic system for personal authentication using the retinal vessel tree as biometric pattern

Author: Marcos Ortega Hortas

Supervisor: Manuel Fco. González Penedo

A Coruña, June 2009

Dr. Manuel Francisco González Penedo
Profesor Titular de Universidad
Dpto. de Computación
Universidade da Coruña

CERTIFICA

Que la memoria titulada "*Automatic system for personal authentication using the retinal vessel tree as biometric pattern*" ha sido realizada por D. Marcos Ortega Hortas bajo mi dirección en el Departamento de Computación de la Universidade da Coruña y concluye la Tesis Doctoral que presenta para optar al grado de Doctor en Informática.

Y para que así conste firmo la presente en

A Coruña, 16 de Junio de 2009

Fdo.: Manuel Fco. González Penedo
Director de la Tesis Doctoral

Acknowledgements

This work is the result of a lot of work, experiences and dedication not only of my own. Therefore, I would like to express my gratitude and acknowledgements to all the people who helped me during this process.

I would like to start with my Ph.D. advisor, Manuel González Penedo, who made this possible with his strong advises, support and, specially, the enormous patience and effort this work demanded all this time.

I would like to express my most sincere gratitude to the University of A Coruña, and in particular to the Department of Computer Science for their support in the development of this thesis. Also, thanks to the Xunta de Galicia and Ministry of Science and Innovation for funding my research.

Also, enormous gratitude and appreciation to Prof. Massimo Tistarelli for sharing with me his knowledge and the outstanding research performed in his group during a research stay in Sardegna (Italy). I'd like to show my gratitude to all the people I met there: Manuele, Linda, Massimo G., Elif, Andrea, Gavin, Enrico, etc. I sincerely thank you all for a really unforgettable time.

Thanks to all the members of the VARPA team, past and present for taking part in this sensational group, specially Cas, Noelia and Carmen for allowing me to steal a lot of their time. Also, special thanks to José and Gabriel for their unparalleled ability to share their brilliant minds for every problem. They make the everyday life more enjoyable and, more important, bearing with me

in every situation in a remarkable way. Also thanks to the "Lambda team", Jose, Mene and Marta for their patience in the whining days.

Last but definitely not least, thanks to my parents and grandparents for being always there any time I need them. This also certainly applies to Verónica, one of the most valuable human beings I have ever met and the one person I could not finish thanking all the good things she brings into my life. Of course, all of my friends who supported me and shared the good and not-so-good moments along the road deserve a lot of credit: Yago, Xabi, Laura, Álex, María José and all the people who supports me out there in real life.

THANKS!

Marcos Ortega

To my parents, Juan and María Luz
To Verónica

"If we knew what it was we were doing, it would not be called research, would it?"

– Albert Einstein

Resumen de la tesis

Introducción

La autenticación fiable de personas es un servicio cuya demanda aumenta en muchos campos, no sólo en entornos policiales o militares sino también en aplicaciones civiles tales como el control de acceso a zonas restringidas o la gestión de transacciones financieras. Los sistemas de autenticación tradicionales están basados en el conocimiento (una palabra clave o un *PIN*) o en la posesión (una tarjeta, o una llave). Dichos sistemas no son suficientemente fiables en numerosos entornos, debido a su incapacidad común para diferenciar entre un usuario verdaderamente autorizado y otro que fraudulentamente haya adquirido el privilegio. Una solución para estos problemas se encuentra en las tecnologías de autenticación basadas en biometría. Un sistema biométrico es un sistema de reconocimiento de patrones que establece la autenticidad de los individuos caracterizándolos por medio de alguna característica física o de comportamiento.

Existen muchas tecnologías de autenticación, algunas de ellas ya implementadas en paquetes comerciales. Las técnicas biométricas más comunes son la huella digital, probablemente la característica más antigua usada en biometría, iris, cara, geometría de la mano y, en cuanto a las características de comportamiento, reconocimiento de voz y firma. Hoy en día, la mayoría de los esfuerzos en los sistemas biométricos van encaminados al diseño de entornos más

seguros donde sea más difícil, o virtualmente imposible, crear una copia de las propiedades utilizadas en el sistema para discriminar entre usuarios autorizados y no autorizados. En este contexto, el patrón de vasos sanguíneos en la retina se presenta como una característica biométrica relativamente joven pero muy interesante debido a sus propiedades inherentes. La más importante es que se trata de un patrón único para cada individuo. Además, al ser una característica interna es casi imposible crear una copia falsa. Por último, otra propiedad interesante es que el patrón no cambia significativamente a lo largo del tiempo excepto en casos de algunas patologías serias y no muy comunes. Por todo ello, el patrón de retina puede ser considerado un rasgo biométrico válido para la autenticación personal ya que es único, invariante en el tiempo y casi imposible de imitar.

Por otra parte, el mayor incoveniente en el uso del patrón de vasos de la retina como característica biométrica radica en la etapa de adquisición todavía percibida por el usuario como invasiva e incómoda. Hoy en día, existen mecanismos para obtener imágenes digitales de manera instantánea a través de cámaras no invasivas pero estos avances requieren a su vez una mayor tolerancia a variaciones en la calidad de la imagen adquirida y, por tanto, métodos computacionales más elaborados que sean capaces de procesar la información en entornos más heterogéneos.

En esta tesis se presenta un nuevo sistema de autenticación automático usando el árbol retiniano como característica biométrica. El objetivo es diseñar y desarrollar un patrón biométrico robusto y compacto que sea fácilmente manejable y almacenable en dispositivos móviles de hoy en día como tarjetas con chip. La plantilla biométrica desarrollada a partir del árbol retiniano consiste en sus puntos característicos (bifurcaciones y cruces entre vasos) de forma que no sea necesario el almacenamiento y procesado de todo el árbol para realizar la autenticación.

Metodología de trabajo

El primer paso en el desarrollo de esta tesis ha sido la realización de un estudio bibliográfico en el campo de la biometría de forma que se han analizado y revisado las técnicas más comunes de autenticación de individuos mediante diversos paradigmas. Se ha hecho un mayor hincapié en el estudio de los paradigmas biométricos correspondientes a características físicas sobre las de comportamiento por ser las más implantadas y las más ajustadas al tema de esta tesis.

Una vez hecho el análisis de técnicas presentes en el campo, se ha realizado un estudio inicial que avale la adecuación del uso del patrón retiniano como característica biométrica. Para ello y tras el repaso a las justificaciones médicas y fisiológicas, se realiza un pequeño sistema de autenticación que utilice el árbol retiniano en su conjunto como método de autenticación y también unos primeros experimentos que justifiquen el uso de la retina para tal fin.

El siguiente paso consiste en el desarrollo de una metodología de extracción del patrón biométrico basado en el árbol retiniano. En este caso, el patrón consistirá en los puntos característicos de dicho árbol. Se presentan dos metodologías que serán analizadas y comparadas, una basada en el método de extracción de crestas en imágenes y otra basada en la segmentación inicial del árbol retiniano. Una vez extraído el patrón, se diseña un sistema de registro y emparejamiento basado en puntos cracaterísticos. Para ello, se modelan las transformaciones geométricas que puedan producirse en la fase de adquisición y se diseña un método de búsqueda en el espacio de transformaciones para encontrar aquélla que sea capaz de emparejar el mayor número de puntos.

Para completar la parte metodológica se realiza un extenso estudio sobre métricas de similaridad que permita discriminar entre comparaciones de patrones del mismo o diferentes individuos. Todas las metodologías propuestas

en cada etapa del proceso, son revisadas y estudiadas mediante evaluación experimental empleando para ellos una base de datos de imágenes digitales de retina específicamente diseñada para evaluar sistemas de autenticación en entornos heterogéneos y de calidad variable.

Contribuciones y conclusiones

En esta tesis se presenta un sistema automático de autenticación basado en el árbol retiniano. A modo de enumeración las principales contribuciones de este trabajo se pueden resumir en:

- Estudio de la viabilidad del patrón retiniano como sistema biométrico de autenticación
- Desarrollo de metodologías de obtención y clasificación de puntos característicos en el árbol retiniano. Esta contribución es también relevante en el ámbito médico a la hora de diagnosticar patologías mediante el análisis del árbol retiniano.
- Desarrollo de técnicas de registro y emparejamiento para puntos característicos
- Estudio y creación de métricas de similaridad en el ámbito biométrico que permitan establecer márgenes de confianza suficientes a la hora de reconocer individuos.

Los resultados experimentales realizados a lo largo del trabajo muestran que el sistema resulta perfectamente válido para autenticación incluso en entornos con altos requerimientos de seguridad, como era inicialmente la intención. Así mismo, el uso de puntos característicos como patrón biométrico permite la codificación de dicho patrón en pequeños dispositivos móviles en contrapartida

a la alternativa de utilizar una imagen entera como patrón. En resumen, los individuos pueden ser fielmente caracterizados por un conjunto de puntos de interés en la retina.

Se desarrollaron dos metodologías para la obtención del patrón retiniano, una basada en extracción de crestas usada en el sistema inicial de autenticación. La otra aproximación parte de una segmentacón completa del árbol y, por tanto, supone una aproximación más precisa a expensas de un coste computacional mayor. En general, el método de crestas es suficientemente bueno y ofrece un conjunto de puntos caraterísticos fiable (en torno a un 90 % de tasa de detección) en un tiempo de computación mucho más limitado.

Se ha diseñado también un algoritmo de emparejamiento de puntos característicos que contemple las posibles alteraciones externas que las imágenes pueden sufrir en la fase de adquisición (transformaciones geométricas, alteraciones en iluminación, etc.). Para reducir el coste computacional en esta fase, se ha desarrollado una metodología precisa de clasificación de puntos característicos en cruce y bifurcación que, a su vez, supone una mejora considerable de precisión respecto a las técnicas previas en el ámbito. La reducción total por comparación es de un 20 %, al limitar el emparejamiento entre puntos a aquéllos clasificados de la misma manera. Este aporte es especialmente importante en tareas como la identificación donde un patrón de un usuario es comparado con todos los patrones almacenados en una base de datos de referencia para encontrar la identidad de dicho individuo. En este caso, la tarea que se repite es la de emparejamiento y por tanto es crucial su optimización en términos de tiempo de computación.

Finalmente, en el estudio de métricas de similaridad se realizó un estudio sobre normalizaciones de dichas métricas y, como principal aporte, se introdujo un término de control sobre las métricas que sea capaz de ponderar la influencia absoluta del número de puntos emparejados independientemente del número total de puntos disponibles en los patrones. Este término se introduce

como resultado de observar que a partir de un cierto número absoluto de puntos emparejados entre dos patrones, la probabilidad de que ambos patrones pertenezcan al mismo individuo crece exponencialmente.

Los tiempos de ejecución medios en un PC de sobremesa con una arquitectura Pentium IV a 2.4Ghz fueron de 0.542s para la etapa de extracción del patrón biométrico y de 0.112s para el emparejamiento de patrones.

Las contribuciones expuestas en esta tesis se encuentran en las siguientes publicaciones:

- M. Ortega, C. Marino, M. G. Penedo, M. Blanco, F. González, "Biometric Authentication using digital retinal images", *International Conference on Applied Computer Science*, pp. 422-427, China, April 2006.

- M. Ortega, C. Marino, M. G. Penedo, M. Blanco, F. González, "Personal Authentication Based on Feature Extraction and Optica Nerve Location in Digital Retinal Images", *WSEAS Transactions on Computers*, 5(6), pp. 1169-1176, 2006.

- M. Ortega, M. G. Penedo, L. Espona, F. González, "Biometric Authentication Based on Feature Extraction in Digital Retinal Images.", *1st Spanish Workshop on Biometrics (SWB 07)*, Girona (Spain), CD proceedings, June 2007.

- L. Espona, M. J. Carreira, M. Ortega, M. G. Penedo, .A Snake for Retinal Vessel Segmentation", *Lecture Notes in Computer Science: Pattern Recognition and Image Analysis*, 4478, pp. 178-185, 2007.

- M. Ortega, M. G. Penedo, C. Marino, M. J. Carreira, "Similarity Metrics Analysis for Feature Point Based Retinal Authentication", *Lecture Notes in Computer Science: Image Analysis and Recognition*, 5112, pp. 1023-1032, 2008.

- L. Espona, M. J. Carreira, M. G. Penedo, M. Ortega, Çomparison of Pixel and Subpixel Retinal Vessel Tree Segmentation using a Deformable Contour Model", *Lecture Notes in Computer Science: Pattern Recognition*, 5197, pp. 683-690, 2008.

- L. Espona, M. J. Carreira, M. G. Penedo, M. Ortega, Retinal Vessel Tree Segmentation using a Deformable Contour Model", *19th International Conference on Pattern Recognition (ICPR 2008)*,
DOI 10.1109/ICPR.2008.4761762 , 4 pp., Tampa Bay, USA, December 2008.

- C. Alonso-Montes, M. Ortega, M. G. Penedo, D. L. Vilarino, "Pixel Parallel Vessel Tree Extraction for a Personal Authentication System", *IEEE International Symposium on Circuits and Systems ISCAS*, pp. 1596-1599, Seattle (USA), May 2008.

- M. Ortega, M. G. Penedo, C. Marino, M. J. Carreira, .A Novel Similarity Metric for Retinal Images Based Authentication", *International Conference on Bio-inspired Systems and Signal Processing (Biosignals 2009)*, pp. 249-253, Porto, Portugal, January 2009.

- M. Ortega, J. Rouco, J. Novo, M. G. Penedo, "Vascular landmark detection in retinal images", *Eurocast 2009*, pp. 71-72, Las Palmas, Spain, February 2009.

- M. Ortega, M. G. Penedo, J. Rouco, N. Barreira, M. J. Carreira, Retinal verification using a feature points based biometric pattern", *EURASIP Journal on Advances in Signal Processing. Special Issue on Recent Advances in Biometrics Systems*, Article ID 235746, 13 pp.,
DOI 10.1155/2009/235746, 2009.

- M. Ortega, M. G. Penedo, J. Rouco, N. Barreira, M. J. Carreira, "Personal verification based on extraction and characterization of retinal feature points",*Journal of Visual Languages and Computing*, 20 (2), 80-90, 2009.

- D. Calvo, M. Ortega, M. G. Penedo, J. Rouco, B. Remeseiro, Çharacterisation of retinal feature points applied to a biometric system", International Conference on Image Analysis and Processing (ICIAP), Salerno, Italy, September 2009. (Aceptado, pendiente de publicación)

Preface

Reliable authentication of persons is a growing demanding service in many fields, not only in police or military environments, but also in civilian applications, such as access control or financial transactions. Traditional authentication systems are based on knowledge (a password, a pin) or possession (a card, a key). But these systems are not reliable enough for many environments, due to their common inability to differentiate between a true authorized user and an user who fraudulently acquired the privilege of the authorized user. A solution to these problems has been found in the biometric based authentication technologies. A biometric system is a pattern recognition system that establishes the authenticity of a specific physiological or behavioral characteristic. Authentication is usually used in the form of verification (checking the validity of a claimed identity) or identification (determination of an identity from a database of known people, this is, determining who a person is without knowledge of his/her name).

There exist many authentication technologies, some of them already implemented in commercial authentication packages. Most common biometrics characteristics are the fingerprint, perhaps the oldest of all the biometric techniques, hand geometry, face, iris or speech recognition. Nowadays the most of the efforts in authentication systems tend to develop more secure environments, where it is harder, or ideally impossible, to create a copy of the properties used by the system to discriminate between authorized and unauthorized individu-

als. In this context, retina blood vessel pattern appears as a very interesting biometric characteristic. This is an unique pattern in each individual and it is almost impossible to forge it in a false individual. Of course, the pattern does not change through the individual's life, unless a serious pathology appears in the eye. Most common diseases like diabetes do not change the pattern in a way that its topology is affected. Some lesions (points or small regions) can appear but they are easily avoided in the vessels extraction method that will be discussed later. Thus, retinal vessel tree pattern can be considered a valid biometric trait for personal authentication as it is unique, time invariant and very hard to forge.

On the other hand, the main drawback of the retinal vessel tree pattern as biometric trait is the acquisition stage, still perceived as not user-friendly. Nowadays, there are new mechanisms to obtain instant digital images of the retina through non-invasive cameras but these advances also require higher tolerance to image variations during the image processing stage.

In this thesis, a novel fully-automatic authentication system is proposed using the retinal vessel tree pattern as biometric characteristic. The goal is to develop a robust biometric template extraction and matching methodology capable of dealing with a more heterogeneous scenario than before. Also, the template should be as compact as possible to allow for fast computation and, above all, easy storage in mobile devices (such as ID cards, chips, etc.). A template is introduced where a set of landmarks (bifurcations and crossovers of retinal vessel tree) are extracted and used as feature points. A registration and matching algorithm for feature points is presented along with a study on similarity metrics for that process.

The experimental results show that the authentication method is capable of achieving a confidence band between scores of authorized and unauthorized individuals in a complex database while using a small but reliable template in a very short computation time.

This document is organized as follows. First, an introduction to different biometric concepts and techniques is presented along with an overview of the retinal verification background. Second chapter presents a study on the retinal vessel tree as biometric pattern, first from a medical point of view and also introducing the research and results obtained in an early authentication system. *Chapter 3* presents the methodology developed to build the authentication system including biometric template construction and template matching algorithms. *Chapter 4* discusses the experiments aimed to test the proposed methodologies and their results. Finally, *chapter 5* offers some conclusions and a final discussion.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In current society, reliable authentication and authorization of individuals are becoming more and more necessary tasks for everyday activities or applications. Just for instance, common situations such as accessing to a building restricted to authorized people (members, workers,...), taking a flight or performing a money transfer require the verification of the identity of the individual trying to perform these tasks. When considering automation of the identity verification, the most challenging aspect is the need of high accuracy, in terms of avoiding incorrect authorizations or rejections. While the user should not be denied to perform a task if authorized, he/she should be also ideally inconvenienced to a minimum which further complicates the whole verification process [77].

With this scope in mind, the term *biometrics* refers to identifying an individual based on his/her distinguished intrinsic characteristics. Particularly, this characteristics usually consist of physiological or behavioral features. Physiological features, such as fingerprints, are physical characteristics usually measured at a particular point of time. Behavioral characteristics, such as speech or handwriting, make reference to the way some action is performed by every individual. As they characterize a particular activity, behavioral biomet-

rics are usually measured over time and are more dependant on the individual's state of mind or deliberated alteration. To reinforce the active versus passive idea of both paradigms, physiological biometrics are also usually referred to as *static biometrics* while behavioral ones are referred to as *dynamic biometrics*.

The term *a biometric* is also used to refer to a specific way of recognizing persons. Therefore, face and iris recognition are considered two different *biometrics*. The term also can refer to samples of the features to recognize such as a fingerprint image or handwritten text samples as shown in Figure 1.1.

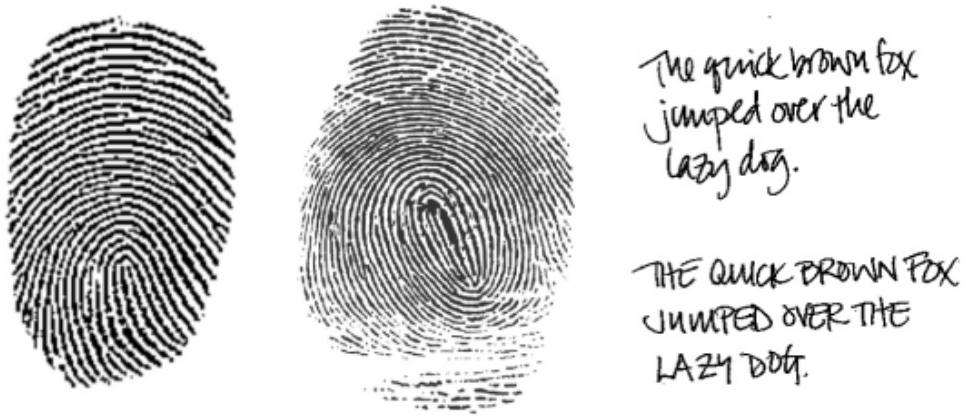


Figure 1.1: Samples of distinct biometrics: physiological (fingerprints) and behavioral (handwriting)

Prior to the arousal of biometrics, personal authentication has been a problem in many fields as described above and many solutions were adopted to overcome it. Traditionally speaking, there are three modes of authentication [63] (Figure 1.2):

1. *Possessions:* Physical objects owned by the individual such as ID cards or keys. It represents something the user *has*.

2. *Knowledge:* Some information that only the individual is supposed to know and, therefore, it must be kept secret. The most typical example

are passwords. It represents something the user *knows*.

3. *Biometrics:* As discussed above, they are physical or behavioral characteristics inherently related to any individual. These characteristics typically have to be unique for every person in order to differentiate between them. It represents something the user *is/does*.

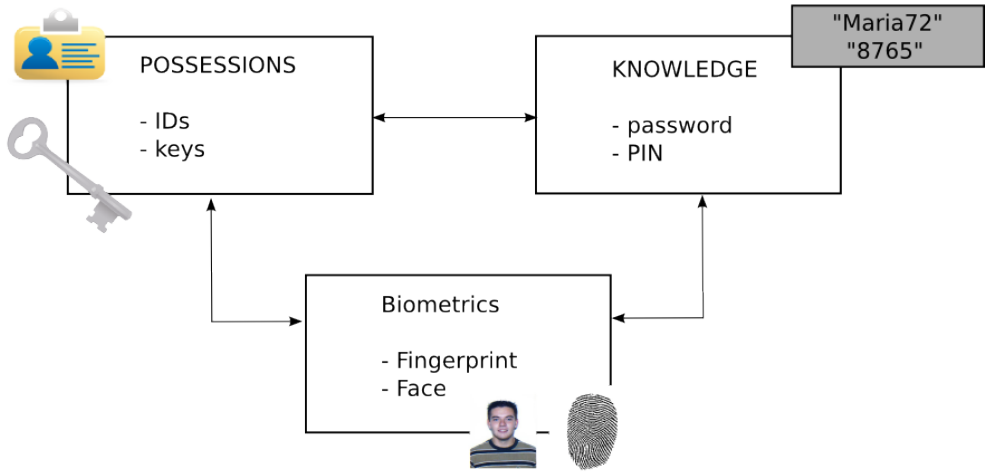


Figure 1.2: Illustration of the three typical modes of authentication.

These authentication modes can be used in combination, especially in automated authentication. For example, a credit card combines possession and knowledge while a passport uses fingerprint and face information.

The traditional authentication systems based on possessions or knowledge are widely spread in the society but they have many drawbacks that biometrics try to overcome. For instance, in the scope of the knowledge-based authentication, it is well known that password systems are vulnerable mainly due to the wrong use of users and administrators. It is not rare to find some administrators sharing the same password, or users giving away their own to other people. One of the most common problems is the use of easily discovered passwords (child names, birth dates, car plate,...). On the other hand, the use of

sophisticated passwords consisting of numbers, upper and lower case letters and even punctuation marks makes it harder to remember them for an user.

Nevertheless, the password systems are easily broken by the use of *brute force* where powerful computers generate all the possible combinations and test it against the authentication system.

In the scope of the possession-based authentication, it is obvious that the main concerns are related to the loss of the identification token. If the token was stolen or found by another individual, the privacy and/or security would be compromised.

Biometrics overcome most of these concerns while they also allow an easy entry to computer systems to non expert users with no need to recall complex passwords. Additionally, commercial webs on the Internet are favored not only by the increasing trust being transmitted to the user but also by the possibility of offering a customizable environment for every individual along with the valuable information on personal preferences for each of them.

According to different authors [17, 92] there are some properties that a particular biometric needs to possess in order to be considered a valid biometric for the authentication task:

- *Universality*: All individuals should have the characteristic.
- *Uniqueness*: Two different individuals should not be the same in terms of the characteristic.
- *Permanence*: The characteristic should be invariant over time.
- *Collectability*: The characteristic should be measurable with some device.
- *Acceptability*: The public in general should not have strong objections to the measurement of the characteristic.

- *Performance*: The characteristic should ideally be measured fast and with a high degree of accuracy.
- *Circumvention*: The biometric characteristic should be hard, or ideally impossible, to forge or simulate.

These properties define the value for a biometric characteristic with the goal of developing useful tools to perform the authentication task. Of course, these properties are not only dependant on the biometric characteristic but also on the particular application of it for a given authentication system [40].

The authentication based on biometrics can be classified into two main categories according to the process of recognizing the individual biometric pattern:

1. *Identification.* The biometric system answers the question *Who is this person?*. We have the biometric pattern for that person available and we need to compare it to the patterns of all the known individuals to determine who he/she is (if any).
2. *Verification.* The biometric system answers the question *Is this person the one he/she claims to be?*. The biometric data of the individual is compared to the previously stored data from the individual whose identity is claimed.

1.1. Biometric system structure

One of the key aspects for a biometric system is its ability to perform the authentication task in a fully automated way. While it is sometimes easy to perform manually a recognition task based on an ID document by comparing the photograph with the individual carrying it, it is a tedious and hard task in

many other situations such as handwriting verification or fingerprint matching. This is why biometric systems are expected to run automatically using computational resources. Figure 1.3 illustrates a generic automated biometric system.

The way to handle the automation in a biometric system can be briefly discussed for every stage in the authentication process:

- The final users need to be initially enrolled in the system. The stored data for any of them will be compared to the acquired data during every access. It is obvious that there must be devices able to obtain the required biometric data (fingerprint scanner, voice recorder, etc.). The acquisition of these data requires some tolerance in the device in order to allow the reproducibility of the captured information at different times.

- It is also needed to consider the quality of the data during the enrollment stage. This is, of course, related to the first stage. The quality will depend, among any other factors, on the number of samples taken from each user. The reason is that the conditions around the acquisition may vary over time. Additionally, several samples enable to proceed using average data. Due to all of these factors, the enrollment procedure usually takes longer than the subsequent authentication accesses. Therefore, sometimes the procedure is optional or controlled by dynamic thresholds. This allows to associate any threshold to a particular access level to the system. In general, there exists a measure to evaluate quantitatively the enrollment stage, the *Failure To Enroll Rate (FTER)*. This parameter is calculated as follows:

$$FTER = \frac{Insatisfactory\ number\ of\ enrollments}{Total\ number\ of\ enrollments} \tag{1.1}$$

  This parameter is conditioned by several factors such as the job of the users. Consider for instance a person using his/her hands in hard tasks.

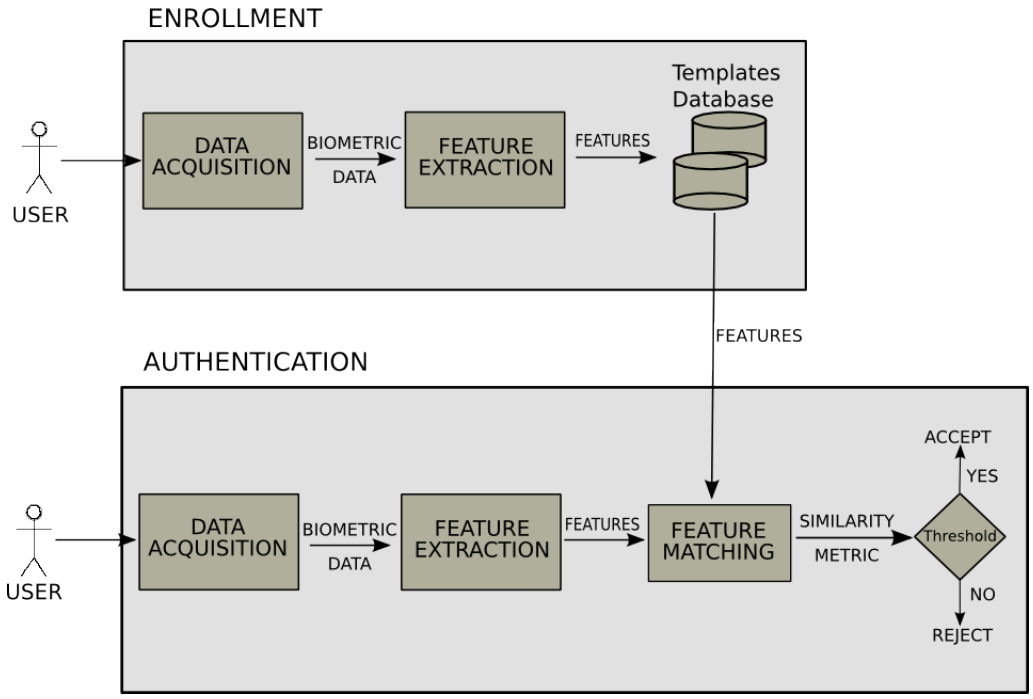


Figure 1.3: Illustration of the typical fully automated biometric system schema. The main stages of the authentication process are acquisition, biometric pattern extraction, comparison and decision.

He/she may have more difficulties to get his/her fingerprint recognized.

- The biometric data acquired are usually raw information without any further processing: images, sound recordings, time between keystrokes, etc. These biometric samples have to be processed in order to be used for recognition. This is required mainly because of their size and complexity. To reduce the size and complexity of the samples, the main characteristics (or features) must be obtained. Some example of feature extraction may be: noise removal from images, filtering of particular voice frequencies, extraction of maxima and minima values in a signal over time, etc. This processed information is usually called biometric template or pattern and it will consist of the data to be compared and analyzed in the recognition process. This is the reason why they must be a compact and reliable representation [8].

- The extracted patterns in the enrollment and authentication stages could be different. In the latter, patterns may be less complex or smaller to speed up the process. Nevertheless, in the matching stage both patterns (the originally stored for an user and the newly acquired for the current access try) are compared in order to establish a degree of correlation between them.

- The final stage in the biometric authentication process is the analysis of the patterns correlation to perform the final identification/verification of the user:

    - *Measurement* of similarity (s) between patterns: unlike the traditional systems based on passwords or PINs, which are usually binary (Yes/No, Correct/Incorrect), the biometric systems establish a similarity score generally in a particular interval. The most typical are normalized value ranges [41]: [-1,1], [0, 1], [0,100], etc. The advantage of the score system is the generation of a measurement of the

degree of similarity between patterns instead of using a categorical value.

- Establishing a *threshold* (T). Once the similarity measure has been obtained, it is compared to a previously calculated threshold score which can vary between users.
- Final *decision*. The threshold establishes the minimum similarity (or maximum dissimilarity) to accept the hypothesis of both patterns belonging to the same individual. It may be the case that both thresholds are used and, if a particular similarity between both is obtained, a final decision can not be taken.

1.2. Evaluation of performance in biometric systems

The evaluation of a biometric application performance is a crucial aspect in the development of this kind of systems. The specifications or the domain where a particular biometric authentication system is going to be installed usually require maximum allowable error rates. There are several types of error measures or specifications and any of them might be more or less valuable in an specific case. Generally speaking, in a verification environment, when we are comparing two biometric patterns to conclude if they belong to the same person or not, there are two possible associated errors to the matching process [89]:

- *False Match* (FM): This error occurs when the system concludes that both patterns come from the same identity, when in reality they belong to different individuals. The frequency of occurrence associated to this error is called *False Match Rate* (FMR).

- *False Non-Match* (FNM): This error occurs when the system concludes that both patterns do not come from the same identity, when in reality they belong to the same individual. The frequency associated to this error is called *False Non-Match Rate* (FNMR).

As previously discussed, the match or non-match between two biometric patterns is usually dependant on a similarity score s. A threshold T is established such as two patterns are considered to match if $s > T$. The score could be expressed in terms of distance rather than similarity. In this case the condition to accept the match would be $s < T$. The challenge is to try to keep the reliability of the score as high as possible. The score is affected by several factors, such as external conditions in the acquisition or data processing. Thus, the analysis of the score distributions between biometric samples (both belonging and not belonging to the same individual) is a key factor to minimize the possible error rates of the matcher by setting an appropriate value of T. In Figure1.4 an example of score distribution is shown for Match and Non-Match groups.

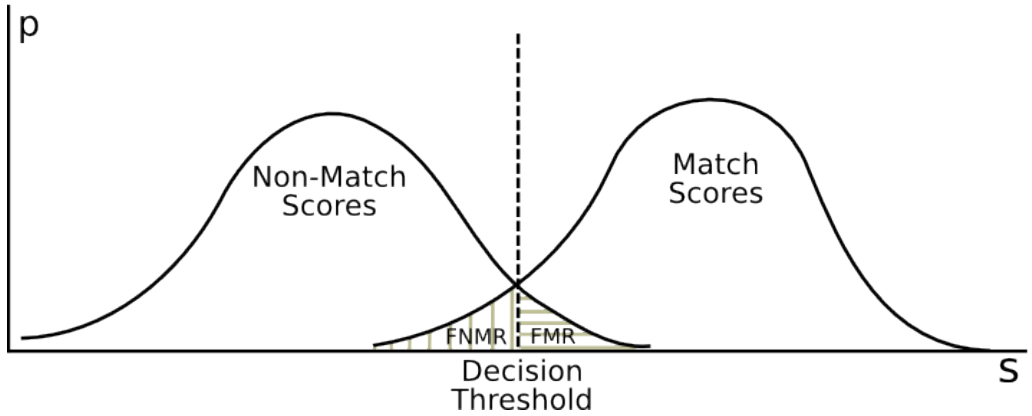


Figure 1.4: Score distribution for match and non-match comparisons. p indicates the probability of obtaining any particular score for every distribution. On average, the non-match scores are lower than the match ones. The value of the decision threshold T determines the False Matching and Non-Matching Rates.

In biometrics, the terms *Accept* and *Reject* are commonly used in contrast to *Match* and *Non-Match*. The acceptance/rejection terminology refers to the acceptance or rejection of a subject to access the system protected by the biometric application. The differentiation is useful in environments where several matching process take place to finally accept or reject the individual. Like in the match terminology, there are two errors associated to acceptance and rejection: False Accept (when the user is granted the access when he should have been denied) and False Reject (when the user is denied to access and he should have been authorized). The frequencies of both errors are denominated *False Acceptance Rate* (FAR) and *False Rejection Rate* (FRR).

As the error rates are tuned by the threshold T the typical error representations are computed varying the threshold. Two typical representations of error rates are the FAR- FRR curves (Figure 1.5) and the ROC curve (Figure 1.6) which expresses the trade-off between FAR and FRR.

From this representations, numerical values are usually extracted to summarize the performance of the biometric application in a single value. The most typical value is the *Equal Error Rate* (EER) which is the error value where $FAR = FRR$.

1.3. Biometric recognition review

Although biometric technology seems to belong in the 21st century, the history of biometrics goes back thousands of years. There are records of use of biometric characteristics as a means of identification since the 8th century, the date of documents and clay sculptures found in China and containing fingerprints in them. Fingerprints have been also used and studied along the centuries but not proficiently. Around 1890, fingerprints started to be treated as a systematic way of identification by Sir Richard Edward Henry of Scotland

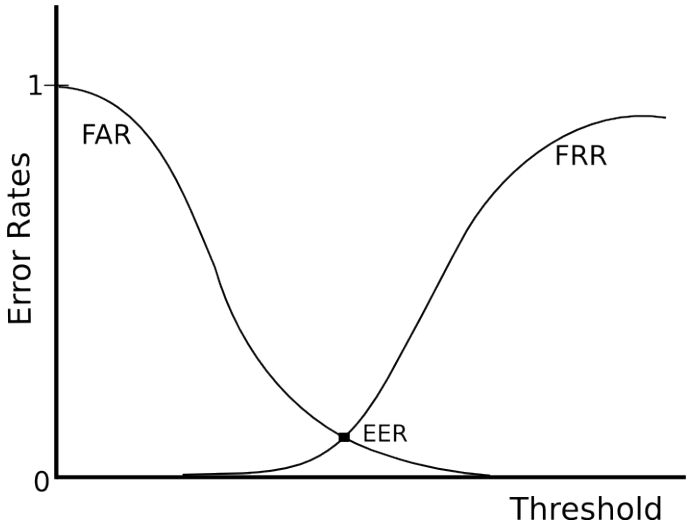


Figure 1.5: Example of FAR curve and FRR curve versus decision threshold. As the threshold is raised, the rate of false acceptances is lowered but the false rejections are increased.

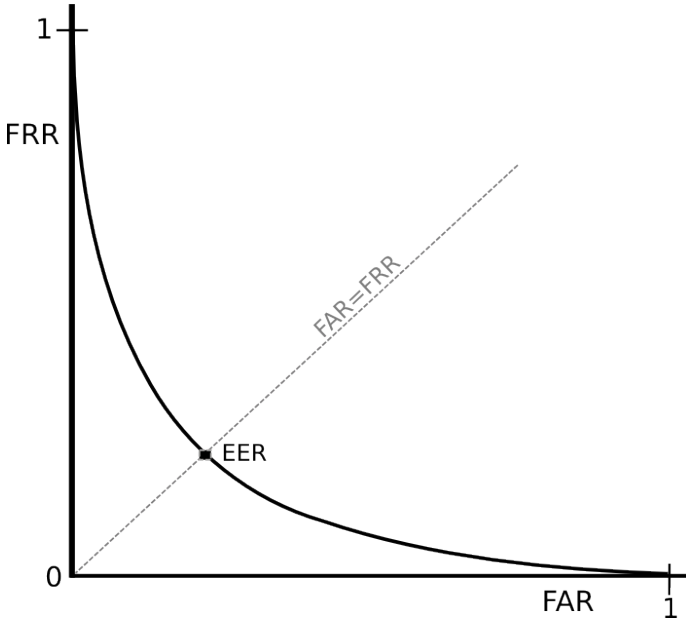


Figure 1.6: ROC curve expresses graphically the trade-off between both error rates.

Physiological	Behavioral
Common	
Face	Signature
Fingerprint	Voice
Hand geometry	
Iris	
Uncommon	
DNA	Gait
Ear shape	Keystroke
Odor	Lip motion
Retina	
Skin reflectance	
Thermogram	

Table 1.1: The most commonly used biometrics (*Common*) and those biometric identifiers either least used or in early stages of research (*Uncommon*) classified into physiological or behavioral characteristic.

Yard using studies from Sir Francis Galton. In 1941, Murray Hill from Bell Laboratories began to work in the voice recognition field. The employment of biometrics as a modern technology applied to commercial tasks was started around the 70s when the first automatic fingerprint recognition systems were developed. In the last two decades, interest in biometrics has been unrelently growing gathering attention and funds from public and private organizations around the world.

This growing interest in biometrics has led to the arousal of many different biometric characteristics used as identity recognition patterns. There are many characteristics widely studied and established while many others are in the initial stages of development. Table 1.1 summarizes the most known biometrics classifying them attending to two criteria: frequency of use (or degree of development) and nature (physiological/behavioral) [7].

Common physiological biometrics have been extensively developed for many years. The behavioral is a younger paradigm where only signature and voice

can be considered to be in a mature state of development. In this work an authentication system based on retinal fundus is proposed, developed and tested using computer vision and pattern recognition techniques. Both fields have been closely related to biometric recognition. Usually, the recognition strategies applied to well-established biometric characteristics are the starting point to develop newer approaches for the modern or uncommon biometrics. Thus, in this section we introduce a brief review on the techniques developed for some of the most widely researched physiological biometrics and some of the emerging techniques. After the review, retinal biometrics will be briefly introduced to place it in relation to other physiological biometrics.

1.3.1. Fingerprint recognition

The fingerprints are morphological biometric characteristics. Their main features are a lines called *ridges* that are skin regions in the finger above other depressed zones of variable width called *valleys*. These lines appear in the human being due to a random process rather than genetic so identical twins can be distinguished by their fingerprints [14]. They are also time invariant: they grow as the human grows but the orientation and relative position between ridges is not distorted at any time. Generally speaking, there are three broad classes of fingerprint recognition techniques:

- *Image Techniques*: this class of techniques is mainly based on computing image correlation or some transform [75, 86].
- *Feature Techniques*: this class extracts some features from the initial data and builds a new representation of the fingerprint based on them [9, 57].
- *Hybrid Techniques*: techniques in this class combine feature and image processing. [1, 20, 70]

The vast majority of the fingerprint recognition approaches is based on feature extraction, in particular landmarks on the ridge lines called *minutiae* (Figure 1.7) [2]. There are two main kinds of minutiae:

- *Ridge ending*: feature defined as the point where the ridge abruptly ends.
- *Ridge bifurcation*: feature defined as the point where a ridge forks into two or more ridges.

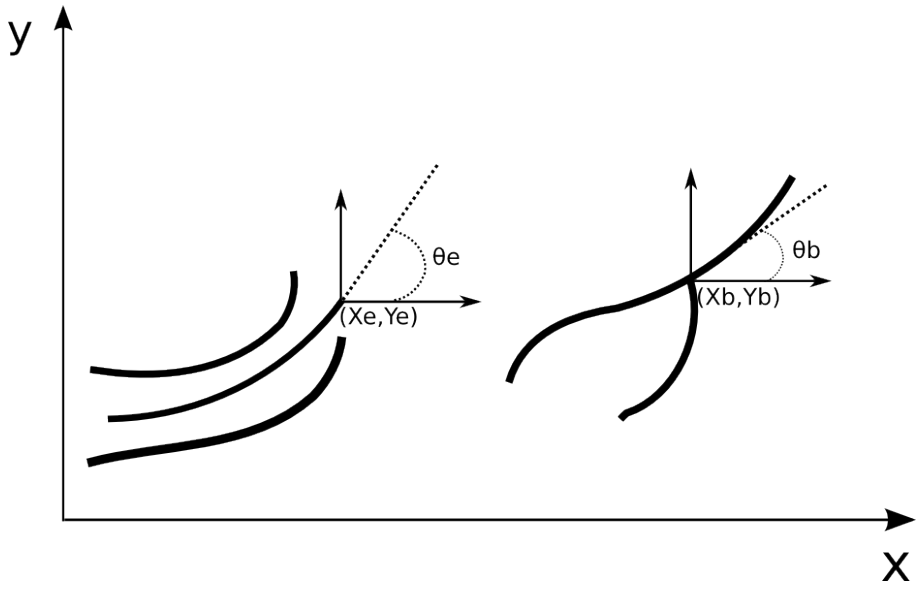


Figure 1.7: Schema of the ending and bifurcation minutiae.

Both features are unequivocally defined by their location (spatial coordinates in a reference coordinate system of the image) and orientation (forming an angle with some reference axis). Roughly, every three minutiae found on a fingerprint, two of them will be ridge endings and one will belong to the bifurcation category. The total number of minutiae in a normal fingerprint is approximately between 40 and 100. Figure 1.8 illustrates a fingerprint with some marked minutiae.

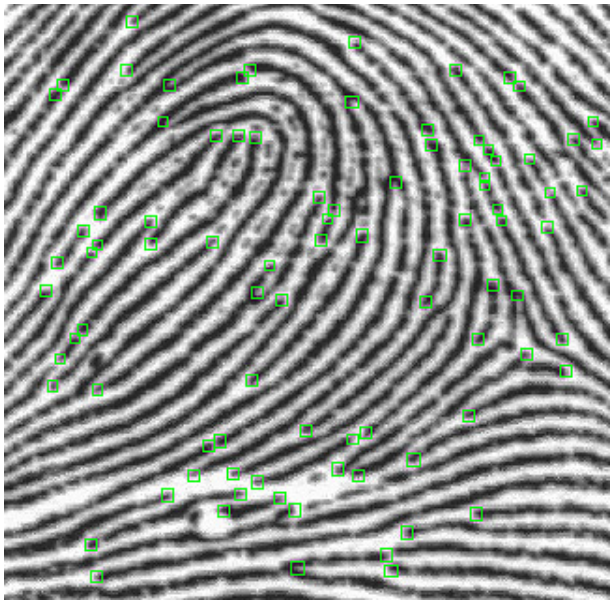


Figure 1.8: Minutiae associated to a particular fingerprint.

Once acquired the fingerprint image using some of the multiple available devices (optical sensors [33], thermal sensors [55], ultrasounds [4], etc.), the authentication task comprises the typical processing stages:

- Preprocessing of the fingerprint image
- Thinning
- Minutiae extraction
- Minutiae matching

The preprocessing of the fingerprint image is a requirement after the acquisition due to the inherent heterogeneity of this process. Illumination conditions, finger pressure on the sensor or humidity conditions are just a few of the many aspects to consider during the acquisition. Due to the high variability, the preprocessing stage is required to obtain a better and more homogeneous

representation of the fingerprint in order to enforce the repeatability on the feature extraction stage. The first step is the image enhancing. In the literature several directional filters were applied to enhance the ridges and remove spurious points in the image [37]. The direction of the filters is usually determined by the *orientation field*, which approximates the ridge line orientations based on the property that, locally, there are not significant variations in the orientation. Another additional approach in the preprocessing stage is the binarization of the image, labelling ridge pixels with 1 value and background pixels with a 0. One of the most prominent algorithms to maximize separability between two classes is presented in [69]. The boundaries of the image without ridges or with a very low density are removed, thus remaining the region of interest (ROI) of the fingerprint. Figure 1.9 illustrates the result obtained for each of the discussed preprocessing techniques.

After these preprocessing stages, the intensity values on the ridges are homogeneous. By performing a thinning, the width of the ridges will be homogeneous too (one pixel) and will make easier the feature extraction task [26]. The thinning processing is performed as a morphological operation searching matches of a structuring element (or pattern) in the image (Figure 1.10). Prior to feature extraction, it is useful to cleanse the thinned fingerprint by removing spurious branches that appeared during the thinning process and joining *broken* ridges, this is, ridges with close end points presenting similar orientations.

The minutia extraction stage avoids the need to store all the image and reduces the biometric pattern to small but significant information. Once the fingerprint is thinned, the minutiae are extracted by analyzing the neighborhood in every pixel P labelled as ridge. The number of neighbors around P, $N(P)$, also belonging to a ridge, will determine the nature of P. Therefore, if $N(P) = 1$, P is a ridge end point as it only has one neighbor. If $N(P) > 2$ then P is a bifurcation. Figure 1.11 illustrates an example of each case.

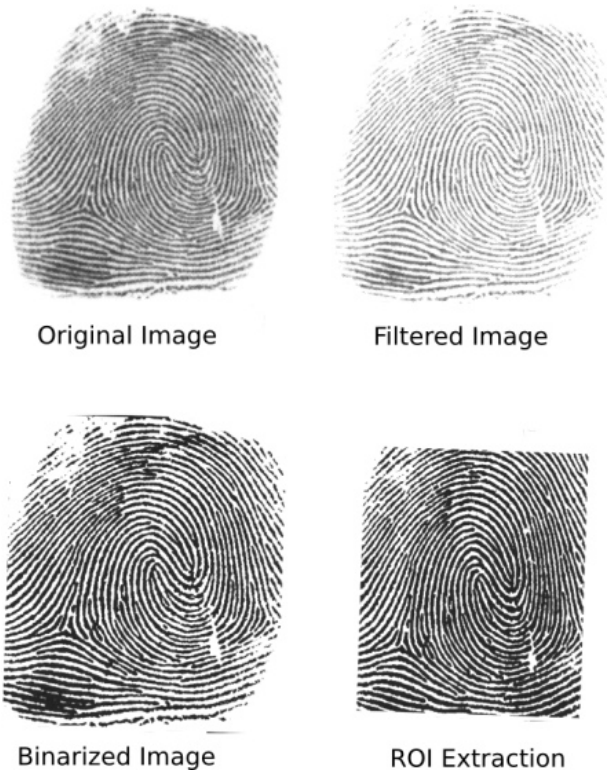


Figure 1.9: Result of applying consecutively different techniques to the fingerprint image prior to the feature extraction.

To obtain a similarity score between two sets of minutiae and due to the transformations a fingerprint image from the same individual can suffer from one acquisition session to another (finger pressure, movement, orientation to the sensor, etc.), both sets have to be aligned to match minutiae between them. The classic approach is to force two minutiae to match, calculate the transformation parameters derived and apply it to the rest of the minutiae. The more minutiae from both sets matched (being close and with similar orientation), the better the transformation is. Once the best matching parameters are selected, the final stage assigns a similarity score between fingerprints. This score will be based on the number of matched minutiae and the individual similarity for each pair [82]. There are periodically competitions held to test different fingerprint verification techniques. The best reported error rates in the last

Figure 1.10: Result of applying a thinning algorithm on the previously preprocessed image.

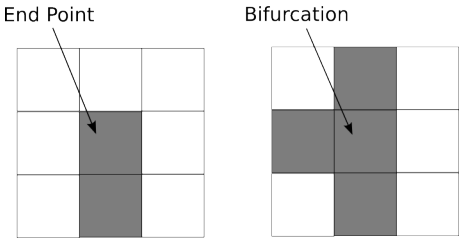


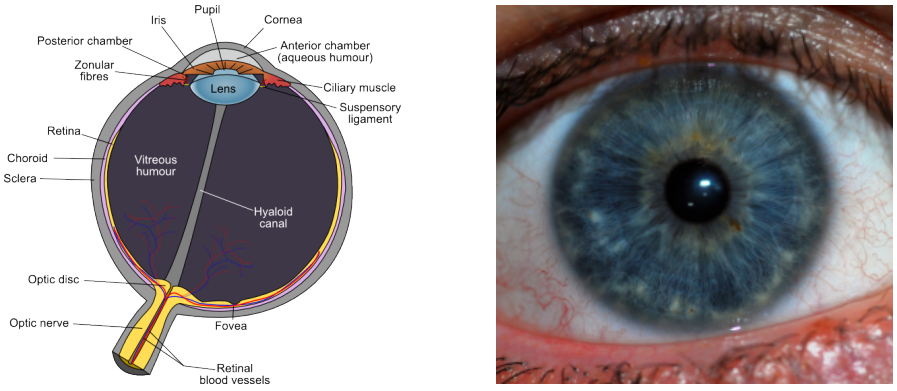
Figure 1.11: Detection and classification of minutiae. The number of neighbors of the middle point determines its class of minutia: end point or bifurcation.

fingerprint verification contest (FVC 2006) [29] are in the range $2.2\% - 2.5\%$.

1.3.2. Iris recognition

The iris is a membrane in the eye, responsible for controlling the amount of light reaching the retina. It is important to note that, although both belong to the eye, iris and retina are totally different biometric characteristics and they are even acquired in a different way. The iris consists of pigmented fiber and vessel tissue known as stroma. It is the most forward portion of the eye and the only one seen on superficial inspection. The stroma connects a sphincter muscle (sphincter pupillae), which contracts the pupil, and a set of dilator

muscles (dilator pupillae), which opens it (Figure 1.12).



(a) (b)

Figure 1.12: (a) Schema of human eye structure (b) Image of a real iris.

In the field of identity verification, uniqueness aside, the use of iris as biometric characteristic offers several advantages: it is a very *stable* characteristic in case of accident; it shows slight variations in respect to the opening range in varying illumination conditions and in fixed illumination conditions making it easier the liveness detection task; Faking an iris is very hard and normally it would take surgery; also, data can be acquired in a non-invasive way through regular cameras.

Since the 80s decade, iris recognition has been widely researched. In 1994, Professor John G. Daugman patented algorithms for iris recognition [22] that served as reference for most of the research in this field [16, 35, 45, 54, 66].

Nevertheless, just like in the case of fingerprint recognition, once acquired the iris image a preprocessing step is needed. The main goal is the localization and segmentation of the iris area. This is usually achieved in 4 stages (Figure1.13):

1. Location of the iris within the image. The image areas to look for are the brightest and darkest as they correspond to the zones surrounding

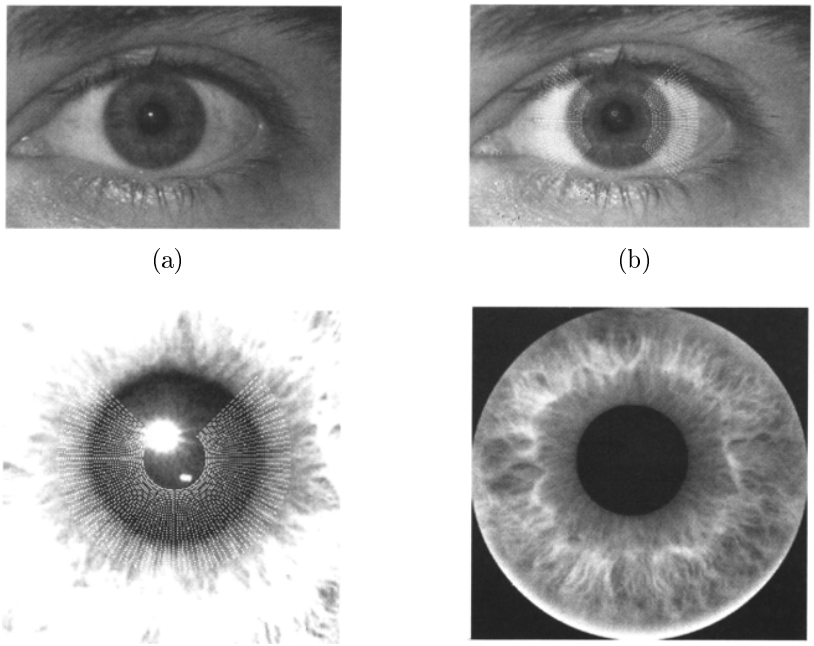


Figure 1.13: (a) Original image (b) Image with sampled points to locate external border (c) Sampled points for the location of internal border (zoom 4x) (d) Iris segmentation result

the iris. Thus, a binarization of the image is performed.

2. Definition of the external border. The border is located by selecting points in a grid that maximize the intensity variation in its neighborhood The grid is defined by a radius to the iris center. It is usually performed iteratively to refine the location.

3. Definition of the internal border. The process is similar to the location of the external border but only considering the resulting area of the previous stage.

4. Scale normalization. This is applied taking into account the distance between the subject and the camera and dilation/contraction of the pupil.

Once the iris is extracted, the most common approaches to define a biometric template are based in Gabor filters or Wavelet transforms. The first step in the Gabor filter approach is to obtain a compact representation of the iris area. Figure 1.14 illustrates the transformation. By sampling radius and angle, the iris is converted to square images where columns represent radio values and rows represent angle intervals.

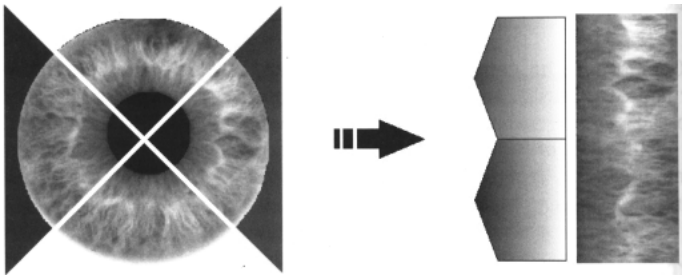


Figure 1.14: Transformation of the iris into a rectangular image.

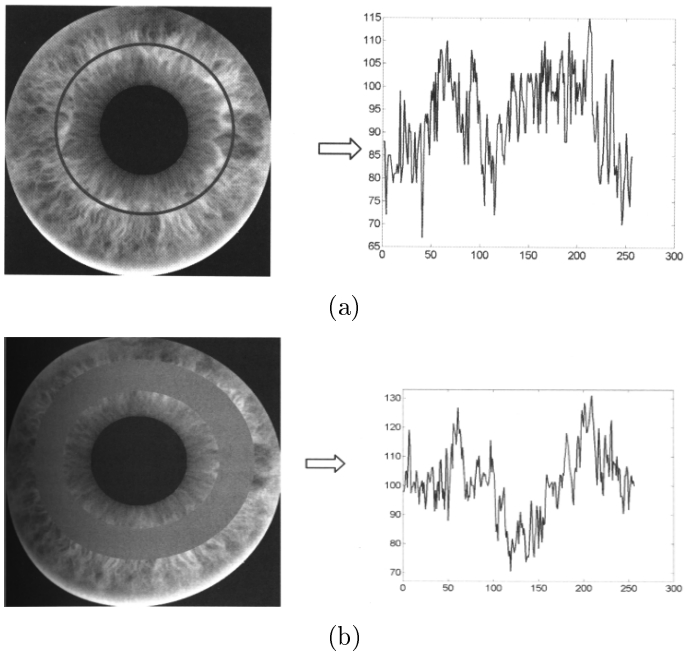


Figure 1.15: (a) Profile and associated iris signature using a single circumference (b) Profile of concentric circumferences and their average signature.

The transformed image is divided into several blocks that might be overlapped and each block is applied a bank of Gabor filters varying in scale and orientation, obtaining a coefficient value for every different block and filter. The coefficients are usually binarized (0 if negative and 1 if positive or null) to reduce the complexity of the final template.

The wavelet approach is just based in its zero-crossing representation along a profile of the iris. The profile can be obtained by a circumference (Figure 1.15(a)) or the average of several circumferences (Figure 1.15(b)).

The error rates of the iris recognition are in the range $1.5\% - 2.5\%$ [7].

1.3.3. Face recognition

Face appearance is a particularly interesting biometric because of its extensive use in the everyday life. First of all, it is the primary human recognition characteristic that humans themselves use and this facilitates its acceptance. Also, as it has been extensively used since the photography and conventional optical imaging devices easily capture faces, there are large databases to research with [96].

Face recognition systems are often required to deal with a wide variety of image acquisition devices and domains. Generally speaking, the common acquisition modes can be classified as:

- Single image. Obtained through an analog or digital camera. It is important to cope with this mode to process legacy documents.
- Video sequence. Used in surveillance domains. Regular camera footage has proved not to be very useful due to its very low resolution.
- 3D image. Least used but in growing demand, some recognition approaches are based on face geometry and require the face to be acquired

using 3D data. The 3D acquisition techniques most commonly used are stereo, structured light or phase-based ranging.

- Near infrared. Useful poor lighting conditions to enhance the face detection process.

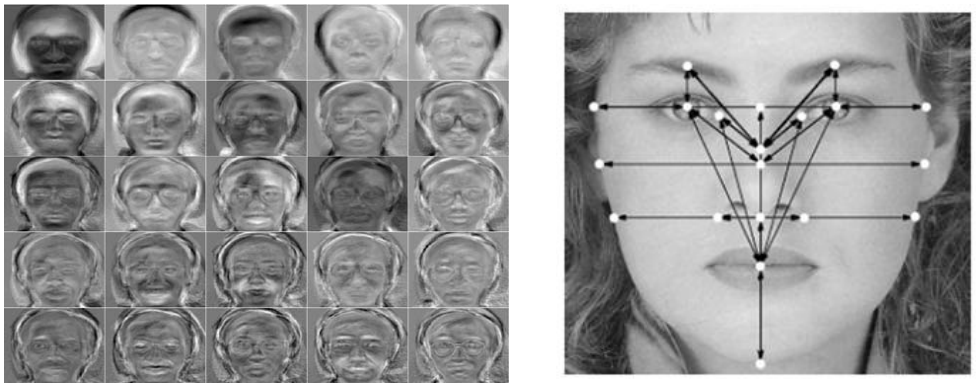
In general, face recognition systems proceed by initially detecting the face in the scene, normalizing parameters in terms of scale and rotations. Many models to find faces have been developed [88,93,95].

Once a face is located and extracted, the approaches to perform a characterization and recognition are divided in two main groups [10]:

Face appearance. The main idea with this approach is to "compress" face information into some small but characteristic representation by removing noise, lighting and other unintended external information. The face is transformed into a space spanned by basis image functions. These functions, known as *eigenfaces*, are the eigenvectors of the covariance matrix of a set of training images (Figure 1.16(a)) [44,64,94].

Face geometry. The idea here is to model the face as a compound of face features such as eyes, nose, mouth and geometric relationships between them. Face recognition is then performed in terms of general matching feature graphs (Figure 1.16(b)) [46,62,85].

Although considerable effort has been done in this field, face recognition is yet not sufficiently accurate to accomplish large-population identification. In particular, determining the identity of two photographs of the same person acquired at different times (and, possibly, different locations) is challenged by a variety of issues. The four main categories would be: *physical appearance*, including quick changes in facial expression such as blinking, emotions, speech, etc. and slower changes due to aging effects or external factors like make-up, glasses, or facial hair; *acquisition geometry* as the face location, in-



(a) (b)

Figure 1.16: (a) Example of eigenfaces. Every face image is decomposed in a linear combination of eigenfaces. (b) Example of the salient features located in a face and their geometric relations.

plane rotation, scale, etc., are totally unknown and highly variable parameters; *imaging conditions* specially the lighting conditions that can greatly vary the intensity values of several face features depending on the light focus; *compression artifacts* derived from traditional compression algorithms (like JPEG) not designed to preserve face appearance.

Several competitions are organized periodically to test the face recognition task performance of current systems. The best error rates for fully automatic systems are in the range $3-5\%$ [61].

1.3.4. Hand geometry

Hand geometry biometrics refer to geometric features or invariants of the human hand. These features typically include width and length of fingers, width of the palm, aspect ratio of the palm and so on [76,80]. Hand geometry does not usually involve the extraction of detailed features such as wrinkles on the skin, leaving them to other domains as palm print or fingerprint verification. Table 1.2 shows an example of a feature set for characterization of the

hand geometry. These features do not vary significantly across the population, but nevertheless they can be used to authenticate the identity of an individual. Each of these features is not very descriptive by itself and therefore the hand geometry recognition presents high error rates in terms of *False Acceptances* and *False Rejections*. The combination of features to obtain higher-level data in order to reduce error rates is a challenging task not very approached to. Traditional hand recognition systems rely on the hand features alone and are very widespread despite their relatively high error rates due to the user-friendliness.

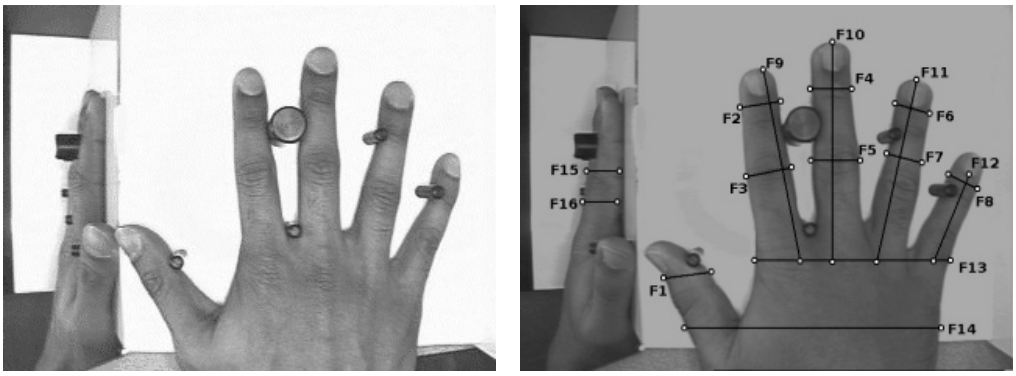
Feature	Description
F1	Width of thumb at second phalanx
F2	Width of index finger at third phalanx
F3	Width of index finger at second phalanx
F4	Width of middle finger at third phalanx
F5	Width of middle finger at second phalanx
F6	Width of ring finger at third phalanx
F7	Width of ring finger at second phalanx
F8	Width of little finger at third phalanx
F9	Length of index finger
F10	Length of middle finger
F11	Length of ring finger
F12	Length of little finger
F13	Width of palm at the base four fingers
F14	Width of palm at base of thumb
F15	Thickness of fingers at second phalanx
F16	Thickness of fingers at third phalanx

Table 1.2: Example of a set of 16 features for hand geometry [42].

Although the acquisition of the hand geometry by means of different sensors (e.g., thermal) could be more reliable, most of the systems still rely on the acquisition of a visual image of the hand [98].

Figure 1.17 shows a prototype of a hand geometry imaging system [42] capturing top and side views of the hand placed in a acquisition platform and a depiction of the features detailed in Table 1.2. The top view is necessary to

measure lengths and widths of fingers as well as palm base while the side view is necessary to measure the thickness of hand/fingers. In the figure, it can be observed as the side image is obtained via a 45-degree mirror.



(a) (b)

Figure 1.17: (a) Hand geometry acquisition devices capturing top and side of the hand. (b) Features referred to in Table 1.2.

In the enrollment stage, two acquisitions of the hand are typically obtained and their results "averaged" to store a reference template. The template consists of a series of features coded in a vector. The matching process compares the newly acquired feature vector (Q) and the stored reference one (R). In [42], a series of similarity metrics are explored (absolute, weighted absolute, Euclidean and weighted Euclidean), corresponding to the following expressions:

$$\sum_{j=1}^{d} |q_j - r_j| < \epsilon_a \tag{1.2}$$

$$\sum_{j=1}^{d} \frac{|q_j - r_j|}{\sigma_j} < \epsilon_{wa} \tag{1.3}$$

$$\sqrt{\sum_{j=1}^{d} (q_j - r_j)^2} < \epsilon_e \tag{1.4}$$

$$\sqrt{\sum_{j=1}^{d} \frac{(q_j - r_j)^2}{\sigma_j^2}} < \epsilon_{we} \tag{1.5}$$

where $Q = (q1, q2, q3, ..., q_d$ represents the query feature vector of the hand whose identity is to be verified and $R = (r1, r2, r3, ..., r_d)$ is the reference feature vector stored for the claimed identity. σ_j^2 is the variance of the jth feature component of the vectors computed over all enrolled templates. This variance weights the importance of every feature. The higher the variability of that feature among the population the more significant is its information. The verification is accepted as positive if the distance between Q and R is less than a threshold value: ϵ_a, ϵ_{wa}, ϵ_e, ϵ_{we} for each respective metric. These metrics are very commonly used in biometrics and for this particular case of hand geometry verification, the Euclidean weighted metric performed better than the rest of them.

1.3.5. Emerging biometrics

Thermograms

A thermogram measures images in the bands of the infrared spectrum, sometimes supplemented by visible spectrum imagery. Thermograms applied to biometrics are images of the parts of the body in the short (0.9-1.7 μm), mid (3-5 μm) and long (8-12 μm) infrared wavelengths. Thermograms, are highly distinctive. Even identical twins have different thermograms. Developed in the mid-1990s, thermography works much like facial recognition, except that an infrared camera is used to capture the images. The advantages of facial thermography over other biometric technologies are that it is not intrusive, no physical contact is required, every living person presents an usable image, and the image can be collected on the fly. Also, unlike visible light systems, infrared systems work accurately even in dim light or total darkness [72]. Figure1.18

illustrates samples of thermogram images.

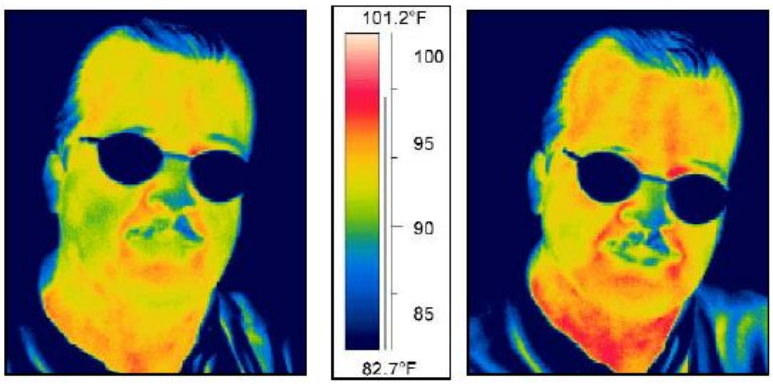


Figure 1.18: Example of facial thermograms.

Since the structures that are being imaged are beneath the skin, this biometric is practically impossible to forge or alter. It is also robust to aging but its main drawback is the high cost of the sensors.

Ear recognition

Ears were originally a part of the Bertillon system of human measurement. They were classified by trained observers into one of a few different ear types to be used as an index in a larger biometric system. In recent years, there has been an increase of the attention dedicated to the ear shape as an automated biometric. Some of the techniques include edge-finding approaches [11, 12] to extract the principal structures of the ear for sake of comparison. In these works, it has also been explored the use of thermograms (i.e. infra-red) as seen before to achieve invariance to illumination and robustness to occlusions, mainly due to hair.

This biometric is specially interesting in the Human Identification at a Distance (HID) project [21]. In [87], an analysis of principal components similar to the eigenfaces approach is performed. The recognition rates are not as good as those using face alone [7]. Figure 1.19 shows the five different ear mask images that were tried in that work.

Figure 1.19: Five different ear mask sizes considered in [87].

1.3.6. Retina biometrics

Identity verification based on retina uses the blood vessels pattern present in the retina (Figure1.20).

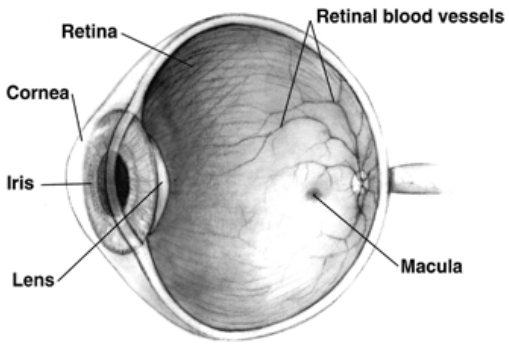


Figure 1.20: Schema of the retina in the human eye. Blood vessels are used as biometric characteristic.

Retinal blood vessel pattern is unique for each human being even in the case of identical twins. Moreover, it is a highly stable pattern over time and totally independent of genetic factors. Also, it is one of the hardest biometric to forge as the identification relies on the blood circulation along the vessels. These property make it one of the best biometric characteristic in high security environments. Its main drawback is the acquisition process which requires collaboration from the user and it is sometimes perceived as intrusive. As it will be further discussed, some advances have been done in this field but, in any case, this continues to be the weak point in retinal based authentication.

Robert Hill introduced the first identification system based on retina [36]. The general idea was that of taking advantage of the inherent properties of the retinal vessel pattern to build a secure system. The system acquired the data via a scanner that required the user to be still for a few seconds. The scanner captured a band in the blood vessels area similar to the one employed in the iris recognition as shown in Figure 1.21.

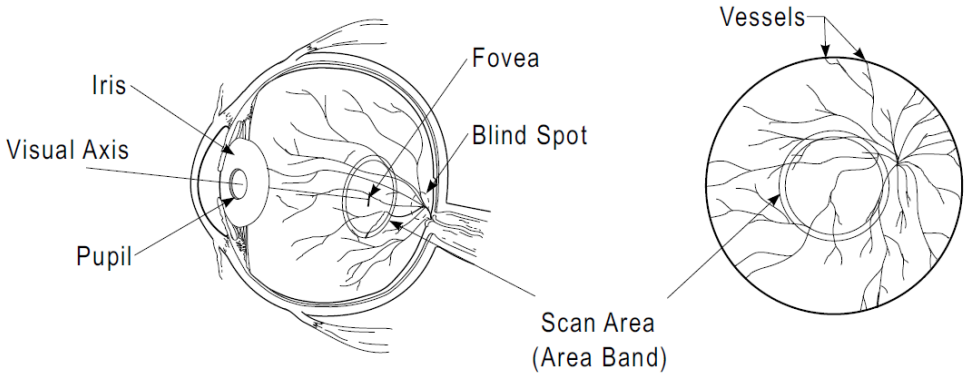


Figure 1.21: Illustration of the scan area in the retina used in the system of Robert Hill.

The scanned area is a circular band around blood vessels. This contrast information of this area is processed via fast Fourier transform. The transformed data forms the final biometric pattern considered in this system. This pattern worked good enough as the acquisition environment was very controlled. Of course, this is also the source of the major drawbacks present in the device: the data acquisition process. This process was both slow and uncomfortable for the user. Moreover, the hardware was very expensive and, therefore, it rendered the system hardly appealing. Finally, the result was that the use of retinal pattern as a biometric characteristic, despite all its convenient properties, was discontinued.

Nowadays, retinal image cameras (Figure 1.22) are capable of taking a photograph of the retina area in a human eye without any intrusive or dangerous scanning. Also, currently, the devices are cheaper and more accessible in gen-

eral. This technology reduces the perception of danger by the user during the retina acquisition process but also brings more freedom producing a more heterogeneous type of retinal images to work with. The lighting conditions and the movement of the user's eye vary between acquisitions. This produces as a result that previous systems based on contrast information of reduced areas may lack the required precision in some cases, increasing the false rejection rate.

Figure 1.22: Two retinal image cameras. The retinal image is acquired by taking an instant photograph.

In Figure 1.23 it can be observed two images from the same person acquired at different times by the same retinograph. There are some zones in the retinal vessels that can not be compared because of the lack of information in one of the images. Thus, to allow the retinal biometrics to keep and increase the acquisition comfortability, it is necessary to implement a more robust methodology that, maintaining the extremely low error rates, is capable to cope with a more heterogeneous range of retinal images.

This thesis is focused on the proposal of a novel personal authentication system based on the retinal vessel tree. This system deals with the new challenges in the retinal field where a more robust pattern has to be designed in

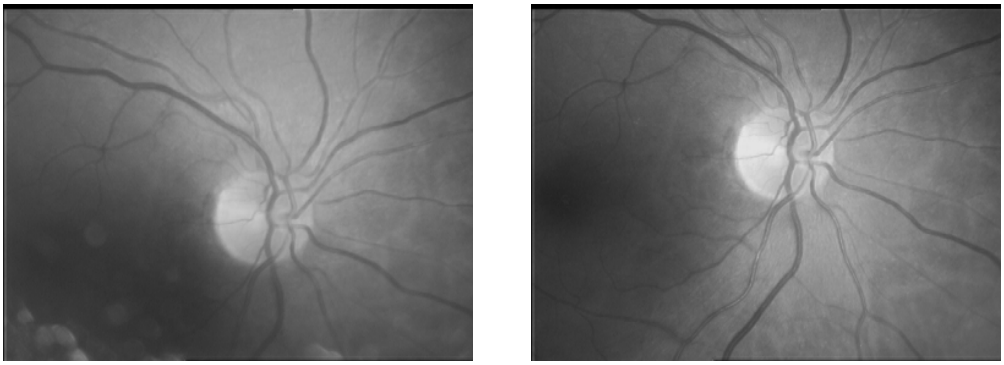


Figure 1.23: Example of two digital retina images from the same individual acquired by the same retinal camera at different times.

order to increase the usability for the acquisition stage. In this sense, the approach presented here to the retinal recognition is closer to the fingerprint developments than to the iris ones as the own structure of the retinal vessel tree suggests. Briefly, the objectives of this work are enumerated:

- Empirical evaluation of the retinal vessel tree as biometric pattern
- Design a robust, easy to store and process biometric pattern making use of the whole retinal vessel tree information
- Development of an efficient and effective methodology to compare and match such retinal patterns
- Analysis on similarity metrics performance to establish reliable thresholds in the authentication process

1.4. Outline

To deal with the suggested goals, the rest of this document is organized as follows. Second chapter introduces a study on the retinal vessel tree as biometric pattern, first from a medical point of view and also presenting the research

and results obtained in an early authentication system. *Chapter 3* presents the methodology developed to build the authentication system, including biometric template construction and template matching algorithms. *Chapter 4* discusses the experiments aimed to test the proposed methodologies, including an analysis of similarity measures. Finally, *chapter 5* offers some conclusions and final discussion.

Chapter 2

Retinal vessel tree as biometric characteristic

Awareness of the uniqueness of the retinal vascular pattern dates back to 1935 when two ophthalmologists, Drs. Carleton Simon and Isodore Goldstein, while studying eye disease, realized that every eye has its own unique pattern of blood vessels. They subsequently published a paper on the use of retinal photographs for identifying people based on their blood vessel patterns [78]. Later in the 1950s, their conclusions were supported by Dr. Paul Tower in the course of his study of identical twins. He noted that, of any two persons, identical twins would be the most likely to have similar retinal vascular patterns. However, Tower showed that, of all the factors compared between twins, retinal vascular patterns showed the least similarities [83].

Blood vessels are among the first organs to develop and are entirely derived from the mesoderm. Vascular development occurs via two processes termed vasculogenesis and angiogenesis. Vasculogenesis, this is, the blood vessel assembly during embryogenesis, begins with the clustering of primitive vascular cells or hemangioblasts into tube-like endothelial structures, which define the

pattern of the vasculature. In angiogenesis, new vessels arise by sprouting of budlike and fine endothelial extensions from preexisting vessels [68].

The application of fractals and fractal growth processes to the branching blood vessels of the normal human retinal circulation was introduced by Masters and Platt [60]. A series of papers led to an estimate of the fractal dimension for the retinal vessels of $D = 1.7$, which is in good agreement with the dimension of a diffusion-limited aggregation cluster grown in two dimensions [27]. Other studies [47, 56] have shown that the fractal nature of the retinal vasculature is independent of age and sex. Many growing, branching objects can be simulated by computer simulations in which the spatial dependence of a field satisfies the Laplace equation with moving boundary conditions. Diffusion-limited growth is a class of these models in which the concentration of diffusing particles satisfies the Laplace equation. One such model is the diffusion-limited aggregation model (DLA) formulated in 1981 by Witten and Sander [91]. The simulation of a DLA yields branching patterns similar to the branching patterns seen in the human retina. The fractal dimension of diffusion-limited aggregates is usually $D = 1.71$ for a fractal embedded in a two-dimensional space. In this way, factors controlling retinal angiogenesis would be thought to obey Laplace equation, with fluctuations in the distribution of embryonic cell-free, spaces providing the randomness needed for fractal behavior and for the uniqueness of each individual retinal vascular pattern [59]. The pattern is therefore consistent with the hypothesis that the development of retinal blood vessels involves a diffusion process. In [30, 31] DLA was developed from shear stress as a simple model of vasculogenesis. A model is proposed in which the formation of the vascular network proceeds via a progressive penetration of the vessel ramification into a capillary mesh. The driving force is of hydrodynamic origin and results in a Laplacian growth mechanism. In their model, the growth of both arteries and veins follows the directions of high shear stress caused by the blood flow on the endothelium wall of the preexisting capillary mesh. Their growth is driven by a field that

satisfies the Laplace equation.

In a more recent study [90], retinal vascular pattern images from livestock were digitally acquired in order to evaluate their pattern uniqueness. To evaluate each retinal vessel pattern, the dominate trunk vessel of bovine retinal images was positioned vertically and branches on the right and left of the trunk and other branching points were evaluated. Branches from the left (mean 6.4 and variance 2.2) and the right (mean 6.4 and variance 1.5) of the vascular trunk; total branches from the vascular trunk (mean 12.8 and variance 4.3), and total branching points (mean 20.0 and variance 13.2) showed differences across all animals (52). A paired comparison of the retinal vessel patterns from both eyes of 30 other animals confirmed that eyes from the same animal differ. Retinal images of 4 cloned sheep from the same parent line were evaluated to confirm the uniqueness of the retinal vessel patterns in genetically identical animals. This would be confirming the uniqueness of animal retinal vascular pattern suggested earlier in the 1980s also by [23].

Taking these results into account, an initial authentication system based on the retinal vessel tree structure is described in this chapter. This initial system serves two purposes: first, it will be useful to test and validate digital retina images for the authentication task and, second, it will stand as the starting point for the final authentication system introduced in this thesis.

The authentication system will extract the whole retinal vessel tree structure from the digital retina images and it will use it as biometric template. As commented in the introduction, retinal vessel trees will have to be aligned prior to the matching task in order to avoid distortions between different acquisition sessions due to the eye movement or the user position. Figure 2.1 shows a schema of the main stages of the authentication process for this system. The following sections will discuss such stages: retinal vessel tree extraction, registration and matching as well as the results of the experiments evaluating its capabilities as a biometric characteristic.

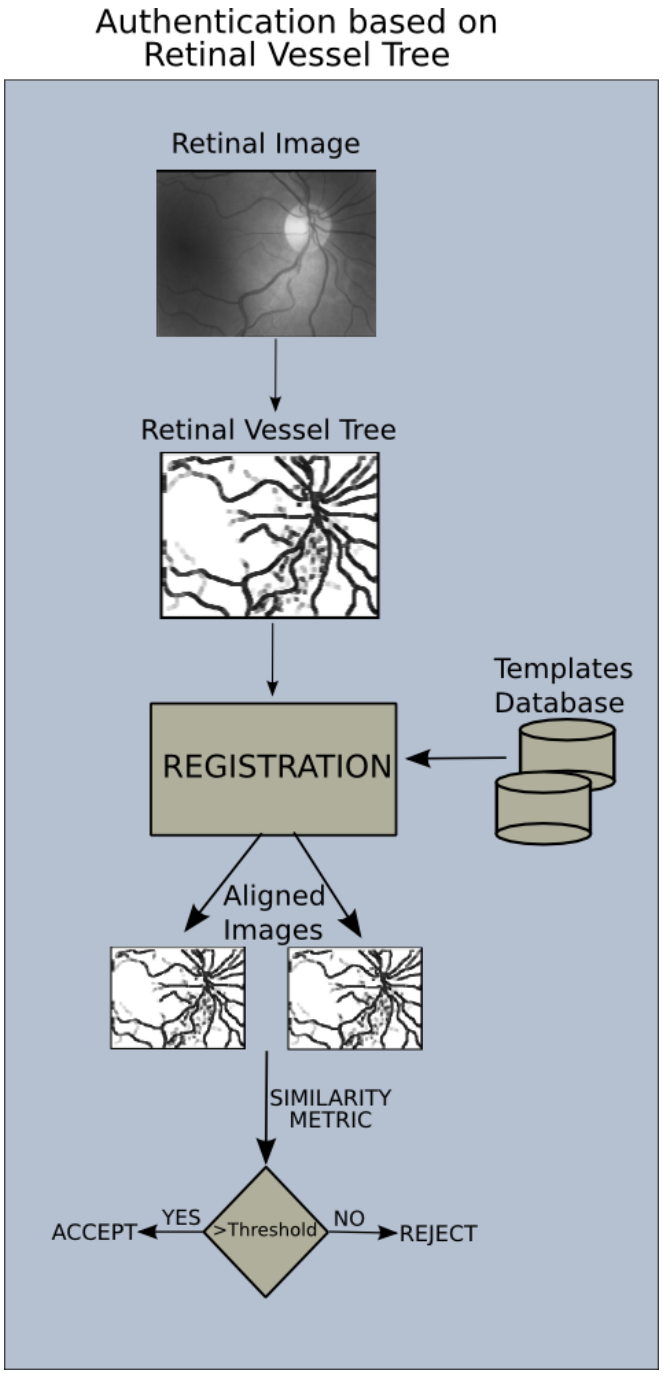


Figure 2.1: Schema of the main stages for the authentication system based in the retinal vessel tree structure.

2.1. Retinal vessel tree extraction

At first glance, the extraction of the retinal vessel tree could be performed by a vessel segmentation technique applied to the retina digital image. In this work, retinal images are optic disc centered but the methods would work in the same way for macula centered images. The segmented tree would be the biometric pattern. However, segmented retinal vessel trees from the same individual acquired from different images would lead to different width in the vessels depending on illumination and other factors, making the registration process very time consuming. Moreover, the level of detail in a full segmentation process is intuitively higher than the needed precision to characterize an individual by the vessel structure. Thus, initially the idea is to achieve a homogeneous representation of the retinal tree structure by representing the vessels as lines.

Following the idea that vessels can be thought of as creases (ridges or valleys) when images are seen as landscapes (see Figure 2.2), curvature level curves will be used to calculate the creases (ridge and valley lines). Several methods for crease detection have been proposed in the literature (see [52] for a comparison between methods), but finally a differential geometry based method [51] was selected because of its good performance in similar images [49, 50], producing very good results.

Among the many definitions of crease, the one based on Level Set Extrinsic Curvature, LSEC [53], has useful invariance properties. Given a function $L : \mathbb{R}^d \rightarrow \mathbb{R}$, the level set for a constant l consists of the set of points $\{\mathbf{x}|L(\mathbf{x}) = l\}$. For $2D$ images, L can be considered as a topographic relief or landscape and the level sets are its level curves. Negative minima of the level curve curvature κ, level by level, form valley curves, and positive maxima ridge curves. The geometry based method named LSEC gives rise to several problems, solved through the improvement of this method by a multilocal solution,

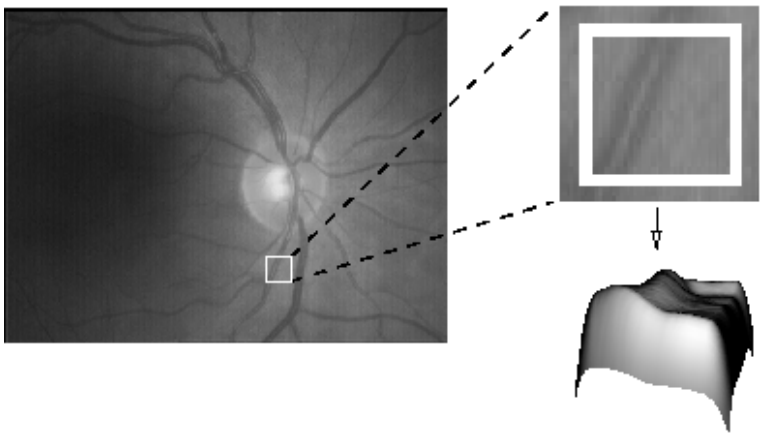


Figure 2.2: Picture of a region of the retinal image as landscape. Vessels can be represented as creases.

the MLSEC [51]. But results obtained with MLSEC can still be improved by pre-filtering the image gradient vector field using structure tensor analysis and by discarding creaseness at isotropic areas by means of the computation of a confidence measure. The methodology allows to tune several parameters to apply such filters as for creases with a concrete width range or crease length. In [13] a methodology was presented for automatic parameter tuning by analyzing contrast variance in the retinal image. For a more detailed discussion on MLSEC methodology see Appendix A.

One of the main advantages of this method is that it is invariant to changes in contrast and illumination, allowing the extraction of creases from arteries and veins independently of the characteristics of the images, avoiding a previous normalization of the input images. The final result is an image where the retinal vessel tree is represented by its crease lines. Figure 2.3 shows several examples of the creases obtained from different retinal images.

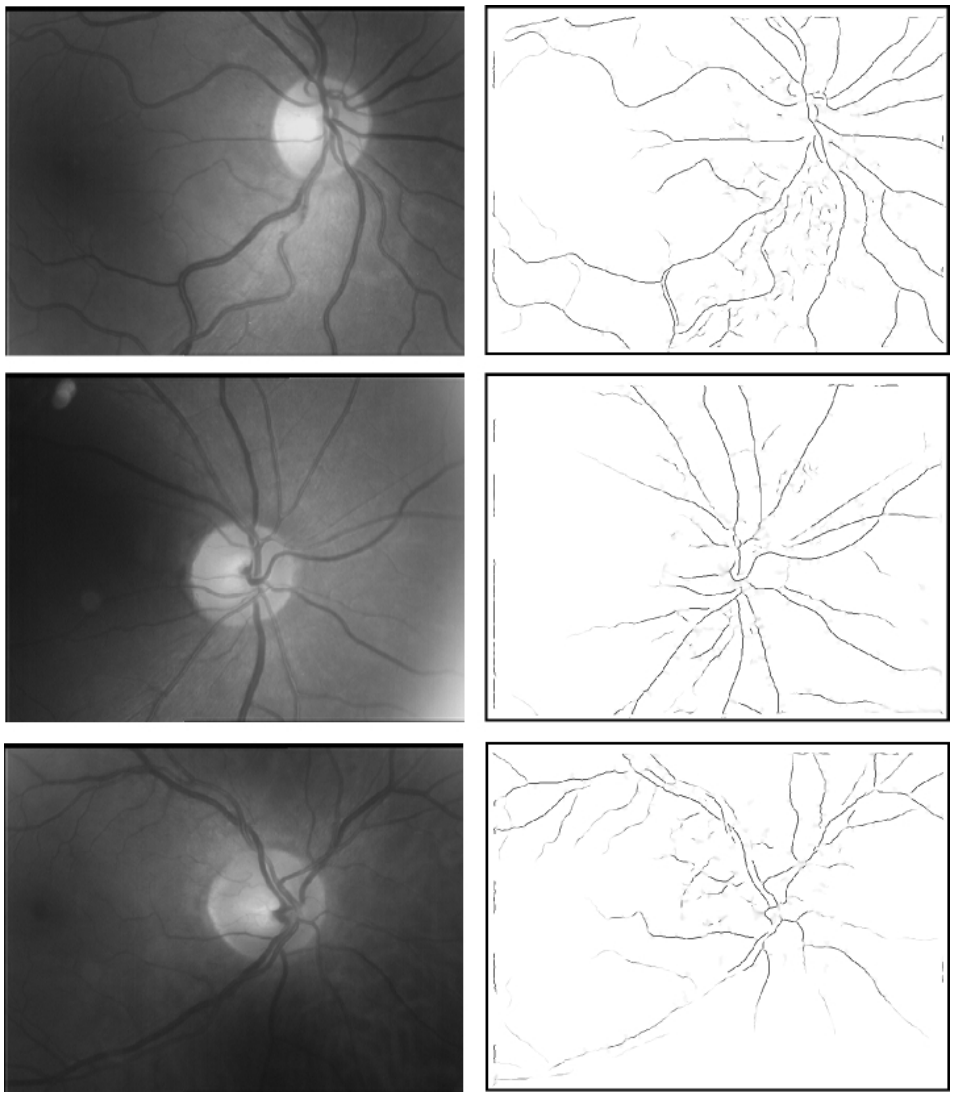


Figure 2.3: Three examples of digital retinal images, showing the variability of the vessel tree among individuals. Left column: input images. Right column: creases of images on the left column representing the main vessels.

2.2. Registration and matching

Once the creases have been extracted from the image, a registration stage is necessary to align with the previously stored reference template. This fact is depicted in Figure 2.4, with two images of the same individual acquired in different moments.

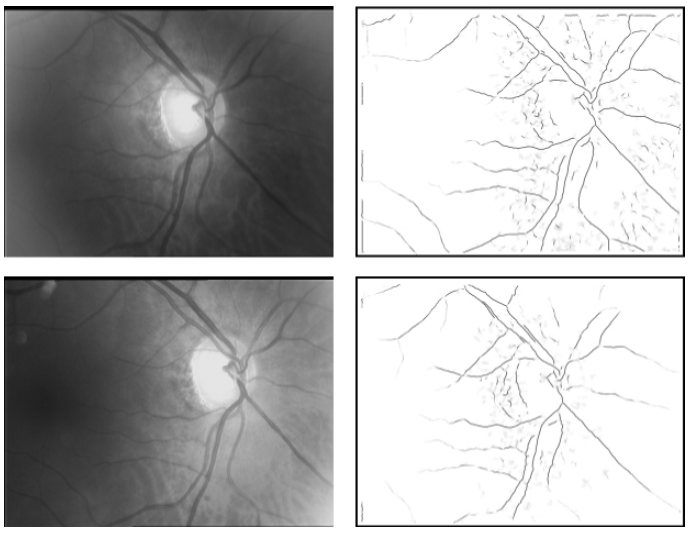


Figure 2.4: Example of the difference between acquisitions. First column: retinal images from the same individual acquired in different moments. Second column: extracted creases from images in first column.

This method resembles the human approach to image matching in the sense that the guidelines are those features common to both images. In [71], the alignment is obtained by matching sufficiently long portions of vessels, although in this case vessels are modelled by tubular structures and their extraction is carried out taking the gradient magnitude along those structures. Moreover, the whole retinal vessel tree is used here.

Since the extraction process is identical for all the images and these are unpredictably variable, the alignment must be robust to missing and non-overlapping features. The straightest approach to the alignment is to perform

an iterative optimization of some alignment function: one image is taken as reference, while the other is iteratively transformed until the function attains a hopefully global maximum. A suitable function to measure the quality of the alignment is the correlation function:

$$\mathsf{Corr}_{\mathcal{T}} = \sum_{x \in f} f(x) \cdot g(\mathcal{T}(x)) \tag{2.1}$$

where f and g are the creaseness images and $\mathcal{T}$ represents the transformation we want to test. A key step is not to transform all the pixels in the image, but only those with values higher than a small fixed threshold, usually discarding around 5% of lower intensity levels. This step saves up to 95% of the total computations and takes into account only the significant creases.

The function $\mathsf{Corr}_{\mathcal{T}}$ together with the five parameters of the transformation (x and y translation, rotation angle as well as x and y scale) define a search space which is difficult to optimize because: (a) the function is non monotonic, i.d., has many local maxima, (b) the similarity measure $\mathsf{Corr}_{\mathcal{T}}$ is expensive to compute since it involves the transformation of a large 2D image, and (c) translation and rotation parameters cannot be decoupled in order to reduce the dimensionality of the search space.

An approach to overcome the two first problems is to search within the parameter space at multiple resolutions. The multiple resolution is handled by building two pyramids, as proposed by Elsen et col. [24], where the ridgeness and valleyness images are at the bottom and each level is a sampled version of the previous level at half resolution, until images have a final size of about 64 pixels in each dimension. As depicted in Figure 2.5, the search starts at the top of the pyramid, where the small size of the images permits an exhaustive search, computing the correlation in the Fourier domain. However, the best transformation at any level often is not the one which later will lead to the final solution, because the hierarchic approach introduces false maxima. Therefore,

for the sake of robustness we might need to keep several values from each level as seeds for the next. For each level we reduce the number of seeds to its half, to end with a single value as the output of the optimization process of the highest resolution level.

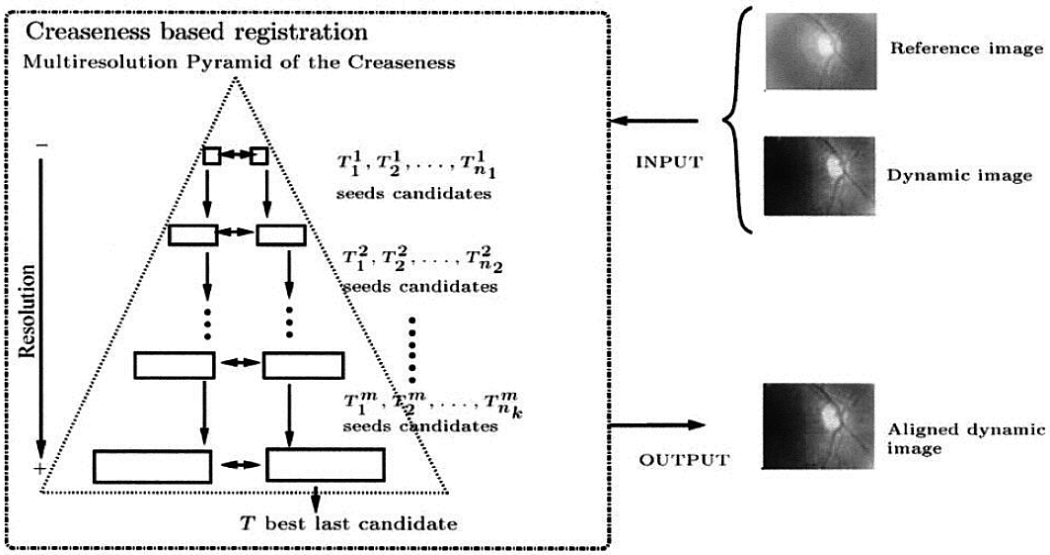


Figure 2.5: Schematic representation of the multi-resolution registration process. The input to the system is the reference image, stored in the authorized personal database, and the acquired image (dynamic image). Then creaseness based registration is used to obtain the T best last candidate transformation. The output is the aligned dynamic image.

Downhill Simplex Iterative algorithm has been selected to maximize the correlation function at all levels except the first, where a more exhaustive search in the Fourier domain is accomplished. Each search starts with the seeds from the previous level and finishes when the algorithm achieves a state where the difference between the maximum and the minimum values found in its neighborhood is lower than a threshold. The tolerance value (minimum correlation difference between iterations to consider that the solution is achieved) and the number of seeds in the highest resolution level determine to a high degree its robustness and final computation time.

A test-bench was designed and run for 20 randomly chosen retinal images for the evaluation of the crease based registration method. The set widely covered the features present at the kind of images the algorithm has to deal with. For each image, 50 random transformations were applied (with maximum ± 100 pixels of translation and $\pm 5°$ of rotation, which represent the maximum rotation and translation values present in the images analyzed). Results are depicted in Table 2.2.

	$\overline{\lvert T_a - T_r \rvert}$	**MSE**	**ST(%)**
Mean	0,033	0,097	99,100
Std.deviation	0,050	0,070	1,550

Table 2.1: Statistics of the test bench for the crease based registration. The mean and standard deviation values for twenty different retinal digital images are shown. First column represents the absolute value of the difference between the applied (T_a) and the recovered transformation (T_r), following equation 2.2. The second column represents the mean square error (MSE) of the recovered transformations, and the third column represents the percentage of successfully recovered transformations (ST), considering a successfully recovered transformation that which fulfils $-1 \leq RE \leq 1$ (equation 2.3).

The first column of Table 2.2 represents the mean and the standard deviation of the absolute value of the difference between the applied transformation and the recovered transformation, following equation 2.2.

$$\overline{\lvert T_a - T_r \rvert} = \lvert (t_x - t'_x) + (t_y - t'_y) + (\alpha - \alpha') + (\sigma - \sigma') \rvert \tag{2.2}$$

where t_x, t_y represent the applied translation, t'_x, t'_y represent the recovered translation, α and α' represent the applied and recovered rotation angles respectively and, finally, $\sigma = (\sigma_x, \sigma_y)$ stands for the scale parameters of the applied transformation and $\sigma' = (\sigma'_x, \sigma'_y)$ stands for the scale parameters of the recovered transformation.

Second column of Table 2.2 shows the mean and standard deviation of the

mean square error (MSE) of the recovered transformations. The small values in this column denote the accuracy of the registration algorithm used in the alignment process.

Finally, third column shows the percentage of recovered registrations where the recovered transformation equals the applied transformation, with a difference of ± 1 pixel, i.e., transformations where $-1 \leq RE \leq 1$ (Equation 2.3). To measure the misalignment of the images after registration, four points were manually selected in the dynamic image, so that they could be used as reference points. The registration error for each frame is defined as:

$$RE = \frac{1}{4} \sum_{j=1}^{4} D(\mathcal{T}(x_j), x_j) \tag{2.3}$$

where D is the Euclidean distance and $\mathcal{T}(x_j)$ is the point x_j after the transformation $\mathcal{T}$ has been applied. As Table 2.2 shows, almost all the set of applied transformations were successfully recovered.

Once the registration process has been performed and the images are aligned, the creases of the registered images are used to obtain a similarity measure between them. So, if two images belong to the same person, the aligned crease images will be more similar than images from different persons where the registration process is successfully performed. The similarity measure must be robust against changes in image amplitude such as those caused by variable lighting conditions, and also against the number of points obtained in the creases extraction process. One possibility could be to use the correlation computed in the registration process as a similarity measure, but it does not fulfil those two conditions, since it is affected by the changes in the signal's intensity and a increasing number of points in the extracted creases would produce a higher level in the correlation value. Such conditions are fulfilled by the Normalized Cross-Correlation coefficient γ, defined as [43]:

$$\gamma = \frac{\sum_{x,y}[f(x,y) - \overline{f}][g(x,y) - \overline{g}]}{\left\{\sum_{x,y}[f(x,y) - \overline{f}]^2 \sum_{x,y}[g(x,y) - \overline{g}]^2\right\}^{0.5}} \tag{2.4}$$

where $\overline{g}$ is the mean of the registered image, and $\overline{f}$ is the mean of the image. It must be noted that although the sums are over all of the image, only the overlapping areas of them are not null. The pixels outside the overlapping area are set to zero, avoiding its influence in the computation of γ.

Once the Normalized Cross-Correlation Coefficient γ is computed, a confidence value will determine if two images belong to the same person. Two images will be considered from the same person if the γ value of the aligned crease images is higher than this confidence value.

2.3. Experimental validation

The images used in our experiments were acquired during a period of 15 months and in different centers of the Complejo Hospitalario Universitario de Santiago de Compostela (CHUS), all of them with the same camera, a Canon CR6-45NM Non-Mydriatic Retinal Camera, with a 768×584 pixel resolution and centered in optic disc. Although the camera originally capture color images, a conversion to gray-level images (with 256 gray levels) was performed prior to the storage in the database, since color does not provide any useful information.

Validation of the method has been performed by means of four experiments. The goal of the first experiment was to determine the confidence level which sets the rejection or acceptance of an individual. To compute this level and test the parameters of the registration technique, a set of images, some of them from the same persons, and some of them from different persons, was input to

the system. In the second experiment, the reliability of the system was tested through a bigger set of images and, by analyzing the results, the accuracy of the tuned parameters has also been established. In the third experiment, new images obtained from two centers of the CHUS, located in different places, were input to the system, which served to guarantee that results from previous experiments were reliable and accurate. Finally, a set of images formed by the union of the second and third sets served as input for the fourth experiment.

To avoid false acceptance cases caused by errors in the acquisition, where only small creases could be extracted, an acquired image is considered valid for the authentication algorithm if the number of points in the creases is above a minimum. This threshold is obtained from the application of the Tchebycheff theorem [25], and the acceptance of the image is defined by the next equation:

$$Accept(N_c) = \begin{cases} true & if\ N_c > 3\mu \\ false & otherwise \end{cases} \tag{2.5}$$

where N_c is the number of points in the creases, and μ is the average number of points in the creases of a set of well acquired images.

To set and tune the parameters and the confidence value of the system, a set of 20 images (5 different persons, 4 images per person) were evaluated. Images belonged to a wide spectrum of people (males and females, with ages ranging from 15 to 45). These results showed that the value of γ, normalized to the interval $[0, 1]$, of the images belonging to the same individual, although acquired in different times, is always above the value 0.6, and the γ value of images acquired from different persons is always below value 0.36.

From these results, a bigger experiment was designed to test the reliability of our system. A set of 119 retinal images was introduced to the system, with a priori knowledge of the results that should be found by the system. In the benchmark, two kind of images were analyzed: the most of the images (113)

belonged to different individuals, and 6 were images from the same persons taken at different times (3 individuals, 2 images each of them). From the set of 113 images, three were removed by the system using equation 2.5, because they presented very poor contrast, and creases were too small.

From the remaining 116 images, the Normalized Cross-Correlation γ of the Cartesian product was calculated, and the result is depicted in Figure2.6. It is clear that the values of the diagonal are all 1, since it corresponds to the correlation of the images with themselves. The other values belong to the other two categories: values bigger than 0.60 are obtained correlating images pertaining to the same person but acquired in different moments, and the rest of the values, which are all under the peak value 0.36, correspond to the value γ for images of different individuals.

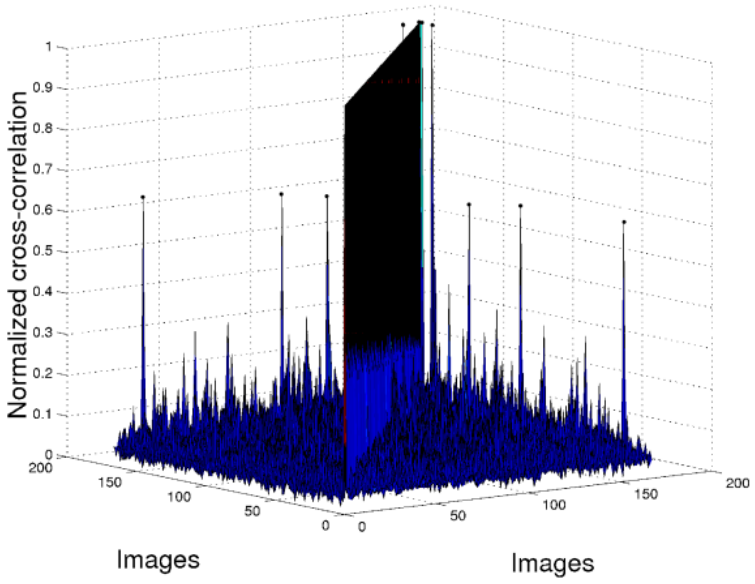


Figure 2.6: Graph representing the values of the correlation obtained in the experiment with 116 images. Main diagonal is always 1, since it corresponds to the γ value of each image with itself, and the other 6 peaks with value 0.6 correspond to the correlation of images from the same person taken in different moments (marked with a dot).

From these results, a third experiment was designed. 168 new images were input to the system, to test if it was able to detect images belonging to the same person. One point must be remarked: in this experiment there was no a priori knowledge of the number of images from the same individuals in the test set. Results are depicted in Figure 2.7.

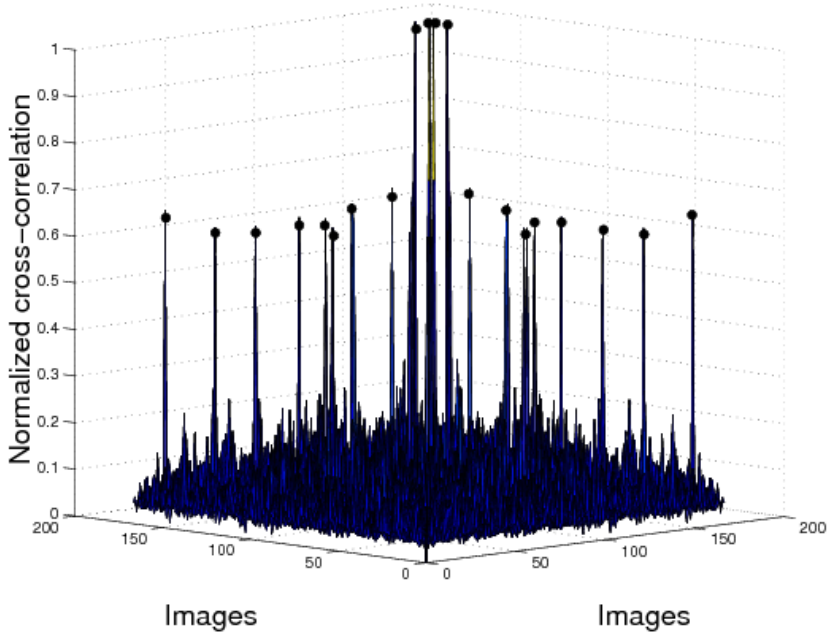


Figure 2.7: Graph representing the values of the correlation obtained in the third experiment with 168 images. Values corresponding to the correlation of images from the same person are marked with a dot. Since the main diagonal corresponds to the correlation of each image with itself, their correlation will always be 1, so it has been removed for a better visualization.

The values of the diagonal were again 1, since they correspond to the correlation of the images with themselves, so they have been removed from the graph in Figure 2.7 for a better visualization. The other values correspond to the other two categories: values bigger than 0.60 are images pertaining to the same person but acquired in different moments, and the rest of the values, which are all of them below 0.41, correspond to the γ value for images of different individuals.

In a new experiment, images from the second and third test sets were joined to test the system with a bigger number of cases, so now the test set is composed of 284 images (116 from the 2nd and 168 from the 3rd experiment). With a higher number of cases, variability among images grows and a more realistic environment is obtained. As expected, the results obtained confirmed the accuracy of the method, and the extreme values of γ were the same: values bigger than 0.61 were obtained by correlating images pertaining to the same individual but acquired in different moments, and the rest of the values, which are all of them below the peak value 0.41, correspond to the value of γ for images from different individuals. With these thresholds, no new false positive cases were found although the variability of the images from the test set was bigger than in the previous experiments.

The confidence level represents a very important parameter in the system, since a too low level could lead to the acceptance of false individuals, but a too high level would reject legitimate individuals. Figure 2.8 shows the percentages of false positive and true positive cases from the combination of the second and third experiments, without counting the correlation values of the main diagonals, corresponding to the correlation of each image with itself. It can be clearly seen that till the threshold value 0.60, the rate of true positive cases is 1, meaning that no true positive is rejected. From value 0.608 the curve decay as it appears some false negative. From threshold 0.685 until 1 the true positive cases are just the values of γ of each image with itself, which is always 1.0. As we have eliminated those cases, the graph shows a null true positive percentage. On the contrary, when the threshold goes down, false negative cases do not appear until it reaches 0.414, growing exponentially from that value. From this analysis, we get to the conclusion that when the threshold is in the interval $(0.415, 0.686)$ the successful percentage of the system is 100%.

All the conclusions exposed in this approach were tested by expert clinicians of the CHUS since they knew, before the experiments were performed, which

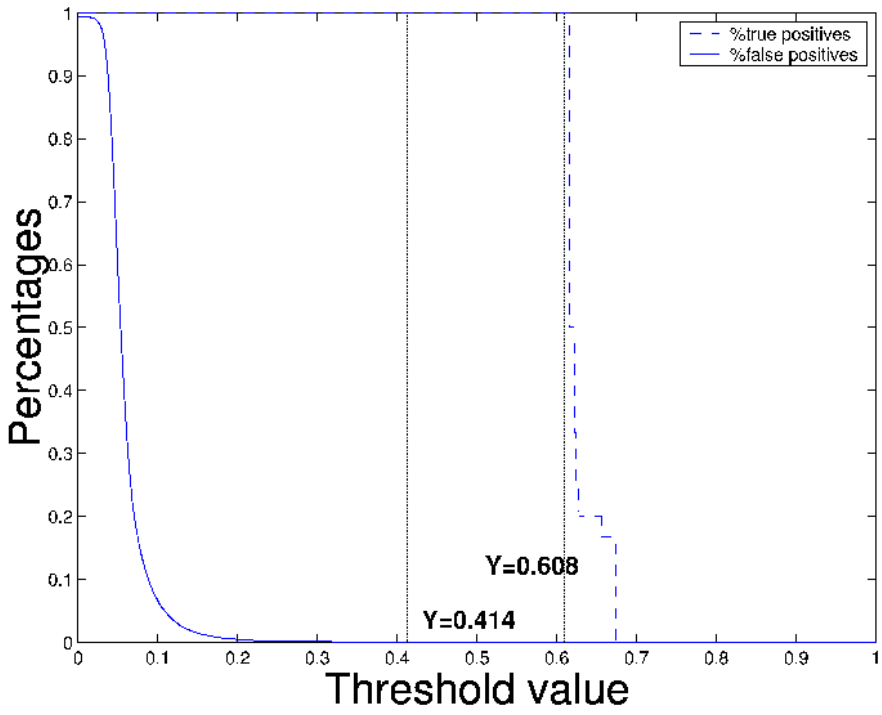


Figure 2.8: False positive and true positive percentages with respect to the threshold value.

images belonged to the same individuals and which not, concluding that our results were correct, and that matching images were effectively taken from the same patients and that false rejections did not exist. These experiments validate the use of the retinal vessel tree as biometric characteristic. However, the template in this system is an image which leads to high demand in terms of storage and memory usage for the authentication process. Currently, it is not realistic to define a whole image as a biometric template because of some of the environments where the patterns are expected to be stored, for instance, in small portable devices or card chips. In the next chapter a new authentication system is discussed based on a more compact template from the vessel tree while keeping the previously discussed advantages of this physical characteristic. The main idea of the system will be to obtain feature landmarks from the vessel tree significant enough to be able to uniquely characterize any individual. The landmarks considered will be the nodes in the graph depicted by the vessel tree, in particular, bifurcations and crossovers between vessels.

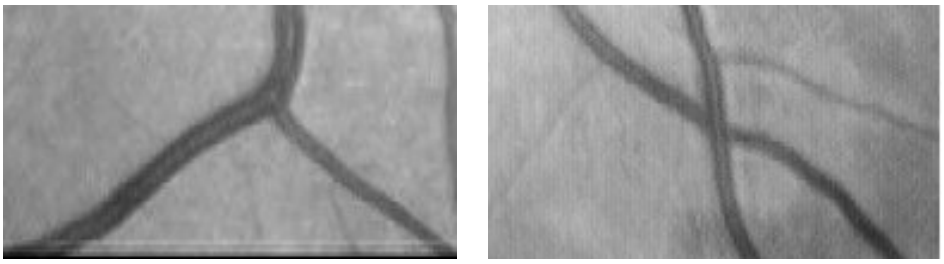
Chapter 3

Retinal verification based on feature points

In this thesis it is proposed a biometric system for authentication that uses the retinal blood vessel pattern as biometric characteristic. As discussed previously, this is a unique pattern in every individual and it is almost impossible to forge it in a false individual. Of course, the pattern does not change through the individual's life, unless a serious pathology appears in the eye. Most common diseases like diabetes do not change the pattern in a way that its topology is affected. Some lesions (points or small regions) can appear but they are easily avoided in the vessel extraction methods discussed later. Thus, the retinal vessel tree pattern has been proved a valid biometric trait for personal authentication as discussed in previous chapter, which methodology was published in [18, 58]. The results showed a high confidence band in the authentication process but only 110 out of 116 images were from individuals with only that sample in the dataset. One of the weak points of the proposed system was the necessity of storing and handling a whole image as the biometric pattern. This greatly increases the difficulty in the storing of the pattern in databases and even in different devices with memory restrictions like cards or mobile

devices. In [28] a pattern is defined using the optic disc as reference structure and using multi scale analysis to compute a feature vector around it. Good results were obtained using an artificial scenario created by randomly rotating one image per user for different users. The dataset size was 60 images, each of them rotated 5 times. The performance of the system is about a 99% accuracy. However, the experimental results do not offer error measures in a real case scenario where different images from the same individual are compared.

Based on the idea of fingerprint minutiae discussed in the Introduction chapter, a robust pattern is introduced here where a set of landmarks are extracted and used as feature points. In this scenario, the pattern matching problem is reduced to a point pattern matching problem and the similarity metric has to be defined in terms of matched points. The feature points considered in this work are the most typical characteristic points in the retinal vessel tree, bifurcations and crossovers (Figure 3.1). More specifically, a bifurcation is a point where a blood vessel splits into two smaller vessels. A crossover can be defined as a point where, with different depth level, two blood vessel coincide.



(a) (b)

Figure 3.1: Types of feature points where (a) shows a bifurcation and (b) shows a crossover.

Figure 3.2 illustrates the general schema for the new feature point based authentication approach. The newly introduced stages are the feature point extraction and the feature point matching. The following chapter sections will discuss the methodology on these new stages of the system.

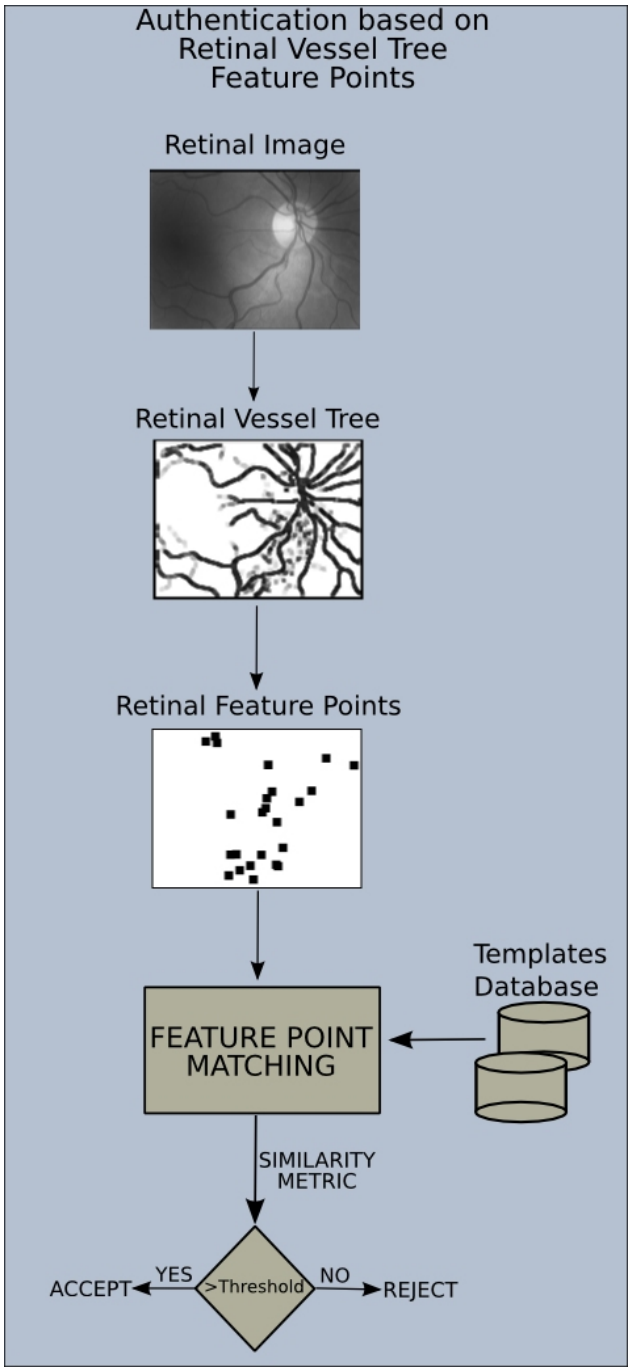


Figure 3.2: Schema of the main stages for the authentication system based in the retinal vessel tree structure.

3.1. Feature Points Extraction

In this section, the methodologies for feature point detection and extraction will be introduced. Detecting vascular tree feature points is a complex task particularly due to the complexity of the vessel structure where illumination and size is highly heterogeneous between images and also between regions from the same image. A common problem in previous approaches is that the optic disc is needed as a reference structure in the image. The detection of the optic disc is a complex problem and it cannot be achieved correctly in some individuals with eye diseases. In this work, the use of reference structures to obtain the biometric pattern is avoided to allow the system to cope with a wider range of images and users.

We have used two different approaches for the feature point extraction task. The first methodology discussed will start from the vessel creases obtained in the previous chapter. Starting from them, crease relationships will be analyzed to detect feature points. The other methodology will be based in a vessel tree segmentation (instead of the valley/ridge computation) and post processing to get to the feature points. In the results chapter, a comparison between the performance of both methods is presented.

3.1.1. Point extraction based on crease computation

The idea of utilizing the crease computation lies in the good representation of the vessel tree that the crease lines already offer.

The landmarks of interest are points where two different vessels are connected. Therefore, it is necessary to study the existing relationships between vessels in the image. The first step is to track and label the vessels to be able to establish their relationships between them.

In Figure 3.3, it can be observed that the crease images show discontinuities in the crossovers and bifurcations points. This occurs because of the two different vessels (valleys or ridges) coming together into a region where the crease direction can not be set. Moreover, due to some illumination or intensity loss issues, the crease images can also show some discontinuities along a vessel (Figure 3.3). This issue requires a process of joining segments to build the whole vessels prior to the bifurcation/crossover analysis.

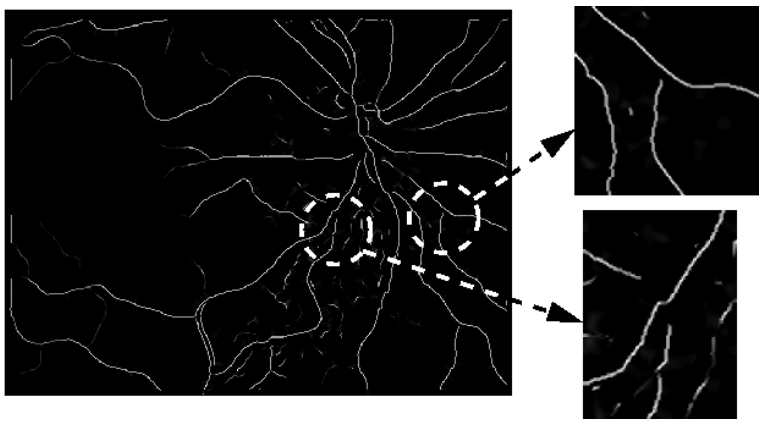


Figure 3.3: Example of discontinuities in the creases of the retinal vessels. Discontinuities in bifurcations and crossovers are due to two creases with different directions joining in the same region. But, also, some other discontinuities along a vessel can happen due to illumination and contrast variations in the image.

Once the relationships between segments are established, a final stage will take place to remove some possible spurious feature points. Thus, the four main stages in the feature point extraction process are:

1. Labelling of the vessels segments
2. Establishing the joint or union relationships between vessels
3. Establishing crossover and bifurcation relationships between vessels
4. Filtering of the crossovers and bifurcations

- Tracking and Labelling of Vessel Segments

To detect and label the vessel segments, an image tracking process is performed. As the crease images eliminate background information, any non-null pixel (intensity greater than zero) belongs to a vessel segment. Taking this into account, each row in the image is tracked (from top to bottom) and when a non-null pixel is found, the segment tracking process takes place. The aim is to label the vessel segment found, as a line of 1 pixel width. This is, every pixel will have only two neighbors (previous and next) avoiding ambiguity to track the resulting segment in further processes.

To start the tracking process, the configuration of the 4 pixels which have not been analyzed by the initially detected pixel is calculated. This leads to 16 possible configurations depending on whether there is a segment pixel or not in each one of the 4 positions. If the initial pixel has no neighbors, it is discarded and the image tracking continues. In the other cases there are two main possibilities: either the initial pixel is an endpoint for the segment, so this is tracked in one way only or the initial pixel is a middle point and the segment is tracked in two ways from it. Figure 3.4 shows the 16 possible neighborhood configurations and how the tracking directions are established in any case.

Once the segment tracking process has started, in every step a neighbor of the last pixel flagged as segment is selected to be the next. This choice is made using the following criterion: the best neighbor is the one with the most non-flagged neighbors corresponding to segment pixels. This heuristic contains the idea of keeping the 1-pixel width segment to track along the middle of the crease, where pixels have more segment pixel neighbors. In case of a tie, the heuristic tries to preserve the most repeated orientation in the last steps. When the whole image tracking process finishes, every segment is a 1 pixel width line with its endpoints defined. The endpoints are very useful

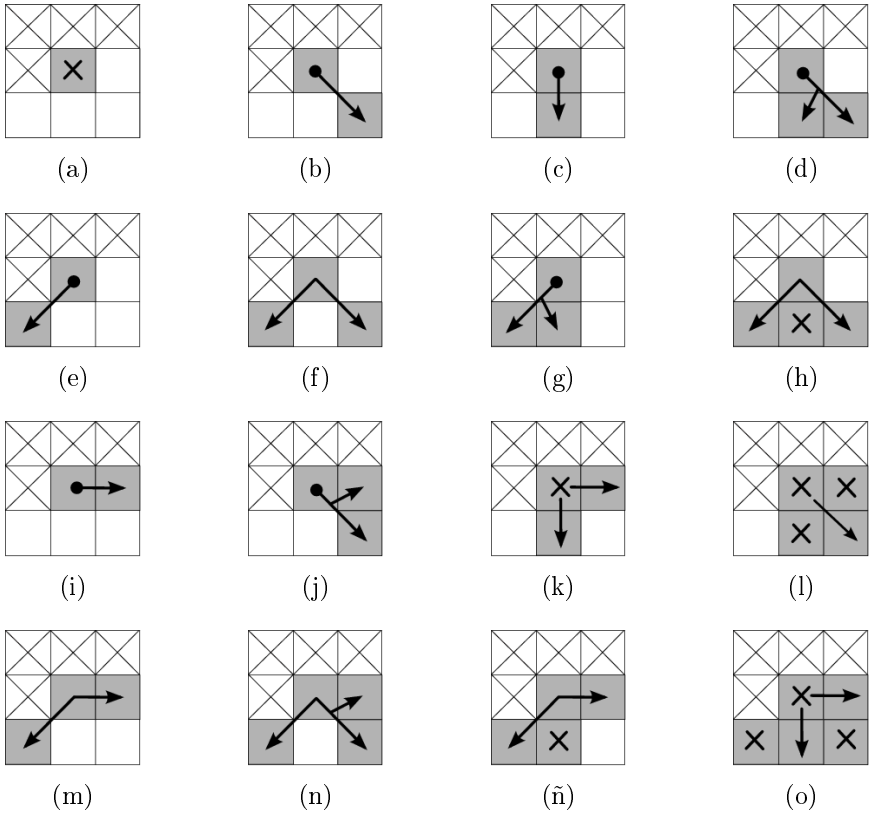


Figure 3.4: Initial tracking process for a segment depending on the neighbor pixels surrounding the first pixel found for the new segment in a 8-neighborhood. As there are 4 neighbors not tracked yet (the bottom row and the one to the right), there are a total of 16 possible configurations. Gray squares represent crease (vessel) pixels and the white ones, background pixels. The upper row neighbors and the left one are ignored as they have already been tracked due to the image tracking direction. Arrows point to the next pixels to track while crosses flag pixels to be ignored. In 3.4(d), 3.4(g), 3.4(j) and 3.4(n) the forked arrows mean that only the best of the pointed pixels (i.e. the one with more new vessel pixel neighbors) is selected for continuing the tracking. Arrows starting with a black circle flag the central pixel as an endpoint for the segment (3.4(b), 3.4(c), 3.4(d), 3.4(e), 3.4(g), 3.4(i), 3.4(j)).

to establish relationships between segments because these relationships can always be detected in the surroundings of a segment endpoint. This avoids the analysis of every pixel belonging to a vessel, considerably reducing the complexity of the algorithm and therefore the running time.

- Union Relationships

As stated before, the union detection is needed to build the vessels out of their segments. Aside the segments from the crease image, no additional information is required and therefore is the first kind of relationship to be detected in the image. An union or joint between two segments exists when one of the segments is the continuation of the other in the same retinal vessel. Figure 3.5 shows some examples of union relationships between segments.

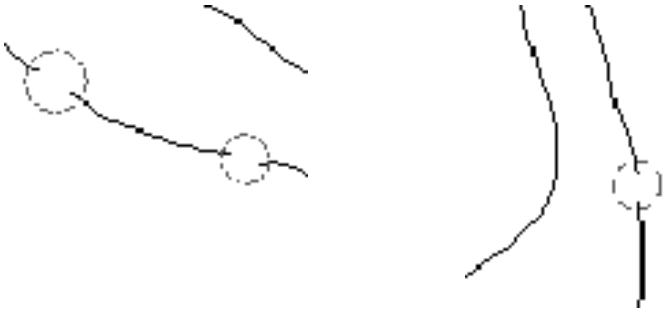


Figure 3.5: Examples of union relationships. Some of the vessels present discontinuities leading to different segments. These discontinuities are detected in the union relationships detection process.

To find these relationships, the developed algorithm uses the segment endpoints calculated and labelled in the previous subsection. The main idea is to analyze pairs of close endpoints from different segments and quantify the likelihood of one being the prolongation of the other. The proposed algorithm connects both endpoints and measures the smoothness of the connection.

An efficient approach to connect the segments is using an straight line

between both endpoints. In Figure 3.6, a graphical description of the detection process for an union is shown. The smoothness measurement is obtained from the angles between the straight line and the segment direction. The segment direction is calculated by the endpoint direction. The maximum smoothness occurs when both angles are π rad., i.e. both segments are parallel and belong to the straight line connecting it. The smoothness decreases as both angles decrease. A criterion to accept the candidate relationship must be established. A minimum angle θ_{min} is set as the threshold for both angles. This way, the criterion to accept an union relationship is defined as

$$Union(r, s) = (\alpha > \theta_{min}) \wedge (\beta > \theta_{min}) \tag{3.1}$$

where r, s are the segments involved in the union and α, β their respective endpoint directions. It has been observed that for values of θ_{min} close to $\frac{3}{4}\pi$ rad. the algorithm delivers good results in all cases.

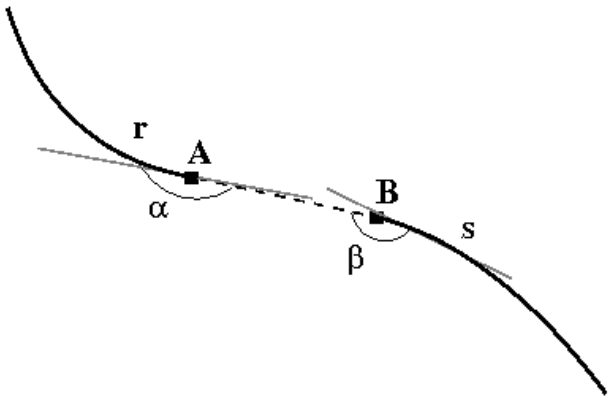


Figure 3.6: Union of the crease segments r and s. The angles between the new segment $\overline{AB}$ and the crease segments r (α) and s (β) are near π rad, so they are above the required threshold ($\frac{3}{4}\pi$) and the union is finally accepted.

- Bifurcation/Crossover Relationships

Bifurcations and crossovers are the feature interest points in this work for characterizing individuals by a biometric pattern. A crossover is an intersection between two segments. A bifurcation is a point in a segment where another one starts from. While unions allow to build the vessels, bifurcations allow to build the vessel tree by establishing relationships between them. Using both types, the retinal vessel tree can be reconstructed by joining all segments. An example of this is shown in Figure 3.7.

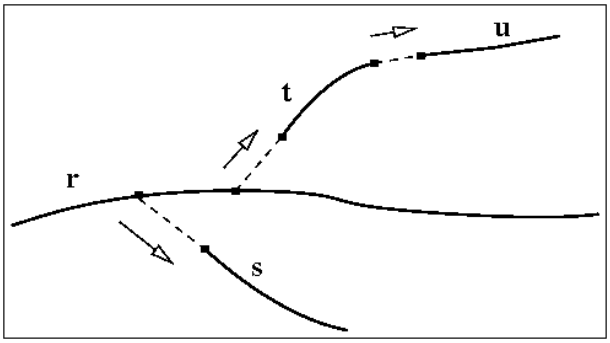


Figure 3.7: Retinal Vessel Tree reconstruction by unions (t, u) and bifurcations (r, s) and (r, t).

A crossover can be seen in the segment image, as two close bifurcations forking from the same segment. Therefore, finding bifurcation and crossover relationships between segments can be initially reduced to find only bifurcations. Crossovers can then be detected analyzing close bifurcations.

In order to find bifurcations in the image, an idea similar to the union algorithm is followed based on the search of the bifurcations from the segments endpoints. The criterion in this case is finding a segment close to an endpoint whose segment can be assumed to start in the found one. This way, the algorithm does not require to track the whole segments, bounding complexity to the number of segments and not to their length.

For every endpoint in the image, the process is as follows (Figure 3.8):

1. Compute the endpoint direction.
2. Extend the segment in that direction a fixed length l_{max}.
3. Analyze the points in and nearby the prolongation segment to find candidate segments.
4. If a point of a different segment is found, compute the angle (α) associated to that bifurcation, defined by the direction of this point and the extreme direction from step 1.

The parameter l_{max} is inserted in the model to avoid indefinite prolongation of the segments. If it follows that $l <= l_{max}$, the segments will be joined and a bifurcation will be detected, being l the distance from the endpoint of the segment to the other segment.

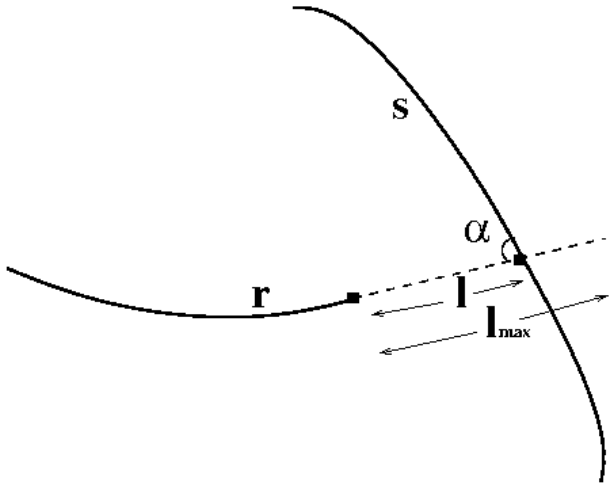


Figure 3.8: Bifurcation between segment r and s. The endpoint of r is prolonged a maximum distance l_{max} and eventually a point of segment s is found.

Figure 3.9 shows an example of results after this stage where feature points are marked. Also, spurious detected points are identified in the image. These spurious points may occur for different reasons such as wrongly detected segments. In the image test set used (over 100 images) the approximate mean

number of feature points detected per image was 28. The mean of spurious points corresponded to 5 points per image. To improve the performance of the matching process is convenient to eliminate as spurious points as possible. Thus, the last stage in the biometric pattern extraction process will be the filtering of spurious points in order to obtain an accurate biometric pattern for an individual.

- Filtering of Feature Points

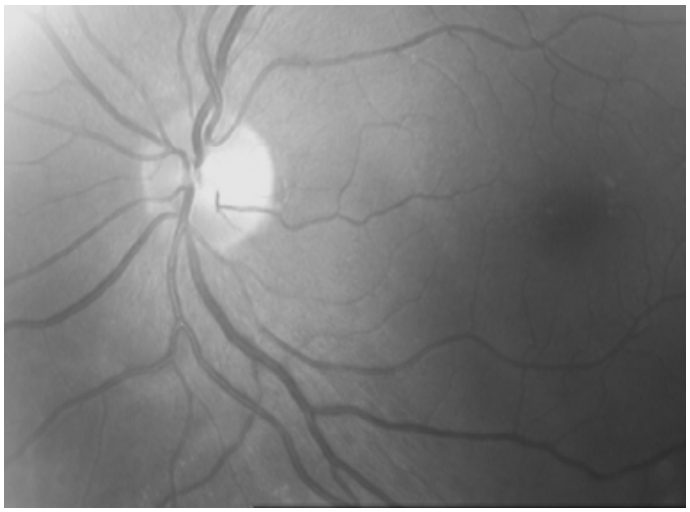
A segment filtering process takes place in the tracking stage, filtering detected segments by their length using a threshold, T_{min}. This leads to images with minimum false segments and with only important segments in the vessel tree.

Finally, since crossover points are detected as two bifurcation points, as Figure 3.9(b) shows, these bifurcation points are merged into an unique feature point by calculating the midpoint between them.

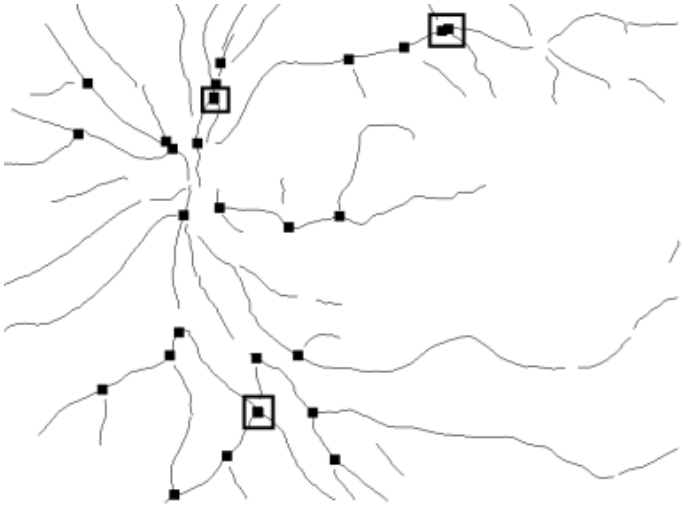
Figure 3.10 shows an example of the filtering process result, i.e. the biometric pattern obtained from an individual. Briefly, in the initial test set of images used to tune the parameters, the reduction of false detected points was about from 5 to 2 in the average. In the Chapter 4, a deeper analysis will be discussed to test the impact of the filtering also in the removal of true feature points.

3.1.2. Point extraction based on vessel segmentation

The goal in this first stage is to detect the feature points of the retinal vessel tree. This detection implies an analysis of the vascular structure. The first step is to perform a segmentation of the vascular tree. In this approach it has been used a technique with a particularly high sensitivity and specificity

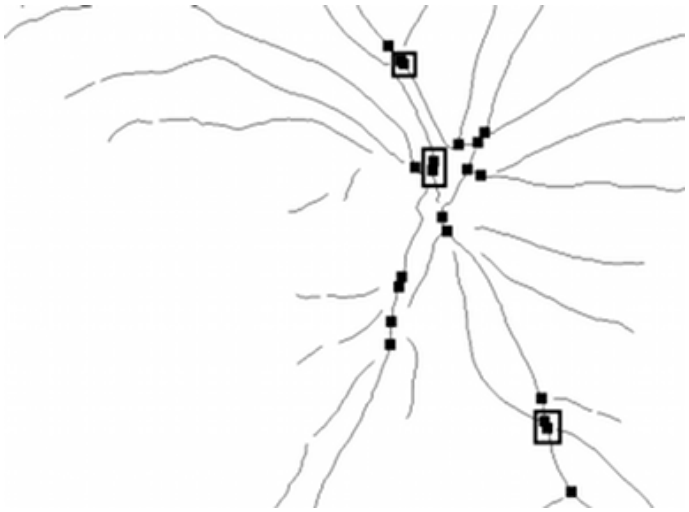


(a)

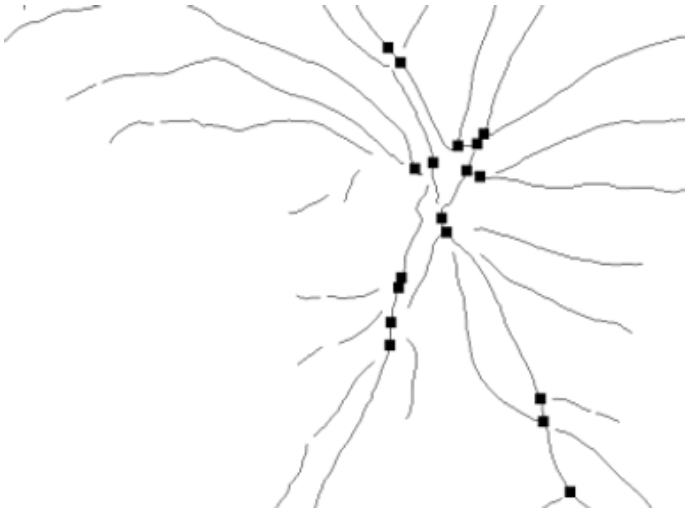


(b)

Figure 3.9: Example of feature points extracted from original image after the bifurcation/crossover stage. (a) Original Image. (b) Feature points marked over the segment image. Spurious points corresponding to the same crossover (detected as two bifurcations) are signalled in squares.



(a)



(b)

Figure 3.10: Example of the result after the feature point filtering. (a) Image containing feature points before filtering. (b) Image containing feature points after filtering. Spurious points from duplicate crossover points have been eliminated.

at classifying points as vessel or non vessel points, described in [19]. As discussed before, properties are not constant along all the structure, like the vessel width, that decreases as the branch level of the structure becomes deeper. To unify this property, a method able to reduce vessel width to one pixel without changing either vessel direction or connectivity is needed. The skeleton is the structure that meets all these properties.

However, the results of the segmentation process force a previous preprocessing step before the skeletonization. Figure 3.11 (a) shows gaps inside the vessels in the segmented image that would give a wrong skeleton structure if the next step is applied to images with this problem. A vessel with gaps in the segmented image would produce two parallel vessels in the skeletonized image (one for each border of the gap) creating false feature points, as shown in Figure 3.11 (b).

(a) (b)

Figure 3.11: Segmentation problems, creating gaps inside the vessels. Subfigure (a) shows the segmentation problem with inside vessel gaps colored in red. Subfigure (b) shows the skeleton of a vessel with gaps, false feature points are marked in red.

To avoid these false positive feature points it is necessary to "fill" the gaps inside the vessels. To perform this task, a dilation process is applied making

the lateral vessel borders grow towards the center filling the mentioned gaps. The dilation process is done using a modified median filter. As in this case the filter is applied to a binary image the result central pixel value will be the most repeated value in the original window. In order to avoid an erosion when the filter is applied to the external border of vessels, the result value will only be set if it is a vessel pixel. To "fill" as much white gaps as possible, the dilation process is applied in an iterative way, this is, dilation is applied to the previous dilation result N times. The value of N must be big enough to fill as much gaps as possible and, at the same time, small enough to avoid merging not connected vessels. The value of N depends on the spatial resolution of the images used. For example, with the images used in this work (768x584) it was determined empirically that optimal values for N were around 4. The iterative process is shown in Figure 3.12.

Usually, the skeletonization goal is to represent global object properties reducing the original image as much as possible. The skeleton, as stated before, expresses the structural connectivity of the objects with a width of one pixel. The basic method to obtain the skeleton is thinning, an iterative technique that erases pixels of the borders with, at least, one background neighbor if this erasing does not change the connectivity. The skeleton is defined by the medial axis function (MAF) [6], defined as the set of points which are the center of the maximum radius circles that fit inside the object. Calculating directly the MAF is a very expensive task and thus template based methods are used due to its versatility and effectiveness. In this work, the Stentiford thinning method [79] is applied. This method uses four templates (one for each of the four different borders of the objects) to erase the pixels only when the template matches and the connectivity is not affected. Figure 3.13 shows the results obtained with this approach. A deeper analysis of the Stentiford approach can be found in Appendix B.

As defined previously, feature points are landmarks in the vessel tree where

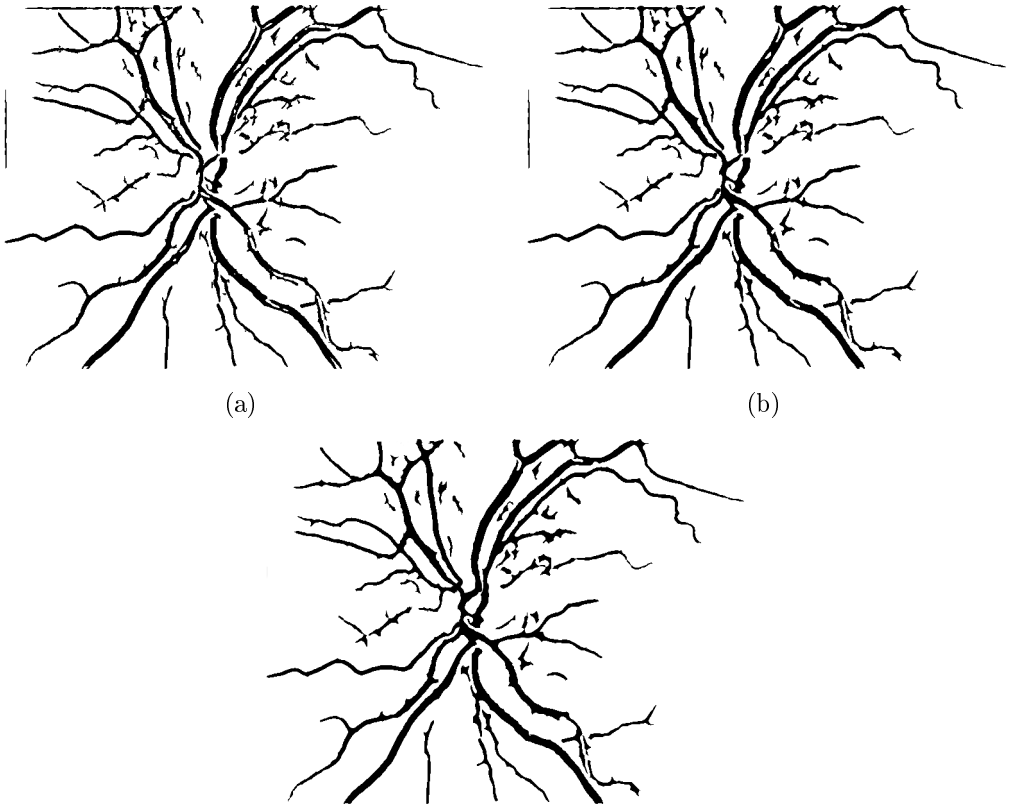
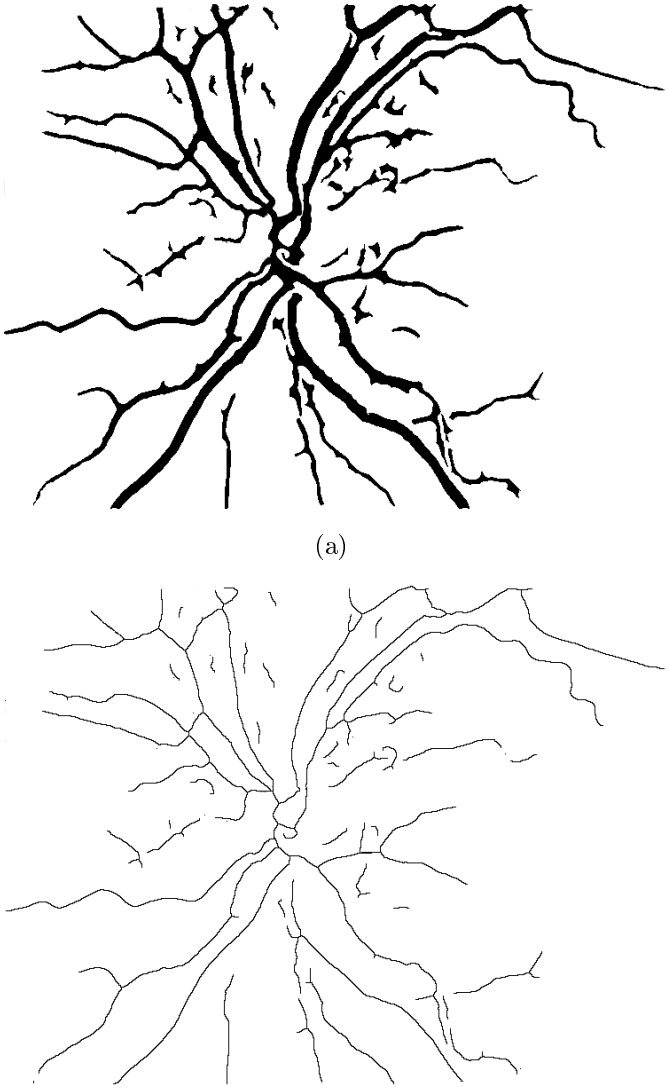


Figure 3.12: Original segmented image (a) and result of the dilation process with $N = 2$ (b) and $N = 4$ (c).

several vessels appear together in the 2D representation. This allows to locate the feature points in the vessel tree using local information along it. This information is obtained with the analysis of the neighbors of each point. This way, the intersection number, $I(v)$, is calculated for each point, v, of the structure as showed in Equation 3.2, where the $N_i(v)$ are the neighbors of the analyzed point, v, named clockwise consecutively.

$$I(v) = \frac{1}{2}\left(\sum_{i=1}^{8} |N_i(v) - N_{i+1}(v)|\right) \tag{3.2}$$



(a)

(b)

Figure 3.13: Thinning process example. (a) Dilated image. (b) Result of the thinning process obtained from the dilated image.

According to its intersection number each point will be marked as,

- Vessel end point if $I(v) = 1$
- Vessel internal point if $I(v) = 2$
- Vessel bifurcation or crossover if $I(v) > 2$

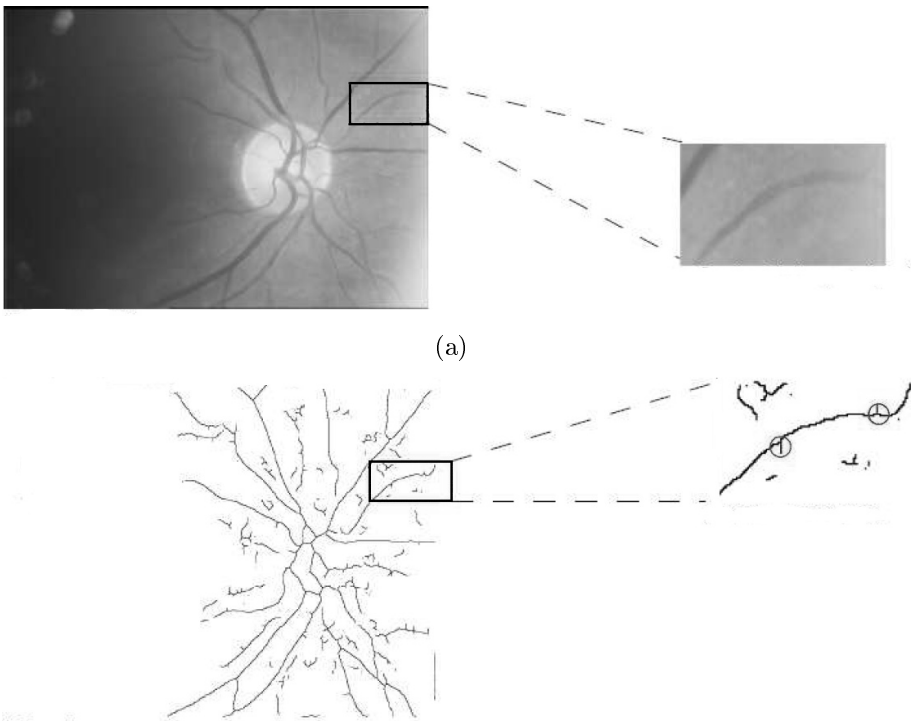
In this approach, points are labelled as feature points when their intersection number $I(v)$ is greater than two, therefore corresponding to bifurcations or crossovers.

The problem in this detection is that not all the points are real points, this is, not every point detected exists in the real image due to the small branches that the skeletonization process creates in the border of the vessels as Figure 3.14 shows.

The skeleton of the retinal vascular tree, as shown before, is obtained from a segmented image through a thinning process that erases the pixels from the borders towards the vessel center without affecting the connectivity. To adapt this structure to a correct point detection, it is necessary to erase the branches that do not actually belong to the retinal tree but its appearance is due to small waves in the borders of the vessels. The process to remove the spurious branches is performed following the next methodology:

1. The points previously detected are divided into two sets,
   - C_1: Set of points labelled as vessel end points. $(I(v) = 1)$
   - C_2: Set of points labelled as bifurcation or crossover. $(I(v) > 2)$
2. With these two sets, the extraction algorithm is as follows,

   a) A point, $c \in C_1$ is taken as initial point (seed).



(a)

(b)

Figure 3.14: Example of branches appearing after the skeletonization process. (a) A region is zoomed in the original image and (b) associated skeleton where circles surround branches not corresponding to any real vessel.

b) Starting in c, the vessel is tracked following the direction of neighbor pixels. Note that every pixel has only one predecessor and one successor.

c) When a previously labelled point, v, is found,

1) If $v \in C_1$, the segment is labelled as an independent segment and the process ends.

2) If $v \in C_2$, the segment is labelled as a branch and the process ends.

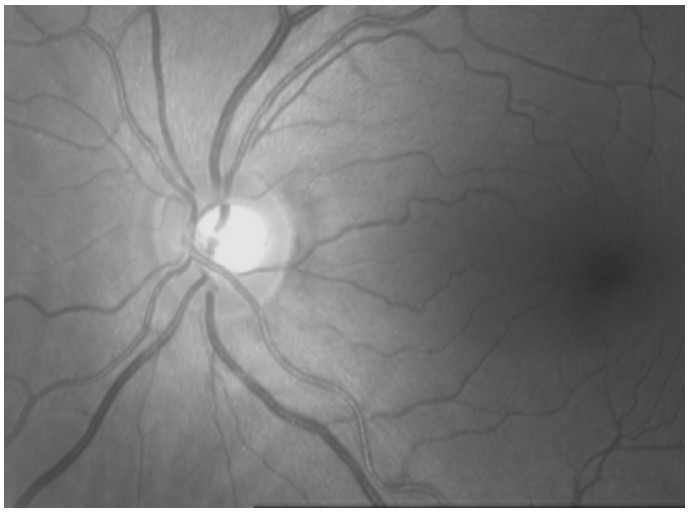
Once obtained all the segments labelled as branches, and defined each of them by its final points (initial and end point), its internal points and its length, the pruning task consists of an analysis of all the branches, deleting the ones shorter than the established threshold (ζ). Erasing a branch implies erasing the intersections associated to it, removing that particular intersection point from the list of feature points.

The chosen value for ζ is given by the origin of the false branches, the ones due to small undulations in vessel borders. So, ζ is the minimum vessel width to be considered in the image. Figure 3.15 shows an example of final feature points extracted with this approach.

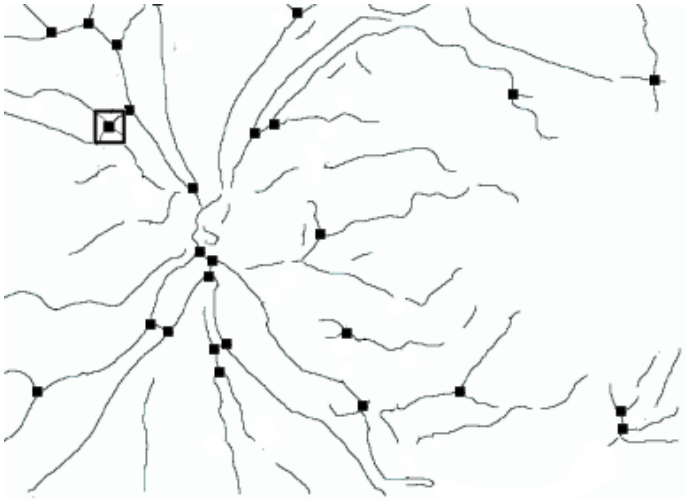
In the results chapter, experiments to tune the parameters are performed and both feature point extraction methodologies are compared.

3.2. Feature Point Matching

In the matching stage, the stored reference pattern, ν, for the claimed identity is compared to the pattern extracted, ν', during the previous stage. Due to the eye movement during the image acquisition stage, it is necessary to align ν' with ν in order to be matched [48,65,97]. This fact is illustrated in



(a)



(b)

Figure 3.15: Example of feature points extracted from original image with the vessel segmentation approach. *(a)* Original Image. *(b)* Feature points marked over the image after the pruning of branches. Again, spurious points are signalled. Squares surround pairs of points corresponding to the same crossover (detected as two bifurcations). The same heuristics than in the crease approach may be followed to avoid those problems.

Figure 3.16 where two images from the same individual, 3.16(a) and 3.16(c), and the obtained results in each case, 3.16(b) and 3.16(d), are shown using the crease approach.

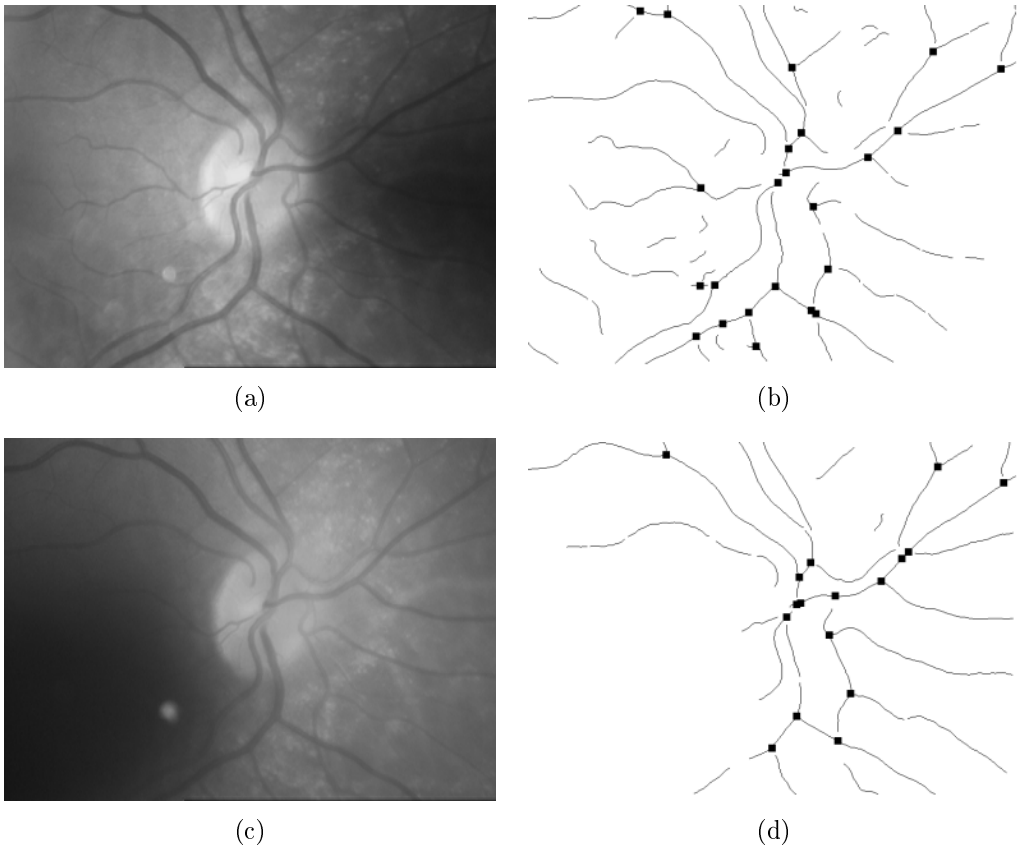


Figure 3.16: Examples of feature points obtained from images of the same individual acquired in different times. (a) and (c) original images. (b) Feature point image from (a). A set of 23 points is obtained. (d) Feature point image from (c). A set of 17 points are obtained.

Depending on several factors, such as the eye location in the objective, patterns may suffer some deformations. A reliable and efficient model is necessary to deal with these deformations allowing to transform the candidate pattern in order to get another similar to the reference one. The movement of the eye in the image acquisition process basically consists of translations in both axis,

rotations and sometimes a very small change in scale. It is also important to note that both patterns, ν and ν', could have a different number of points, as seen in Figure3.16 where, from the same individual, two patterns are extracted with 23 and 17 points. Note that this issue is independent of the point extraction algorithm taken since the transform parameters come determined by different conditions of illumination and orientation in the image acquisition stage.

The transformation considered in this work is the Similarity Transformation (ST), which is a special case of the Global Affine Transformation (GAT) [32]. ST can model translation, rotation and isotropic scaling using 4 parameters. The ST works fine with this kind of images where the rotation angle is moderate. It has also been observed that the scaling, due to eye proximity to the camera, is nearly constant for all the images. Also, the rotations are very slight as the eye orientation when facing the camera is very similar. Under these circumstances, the ST model appears to be very suitable, with the additional benefit that it is a very efficient model compared to the higher-level ones.

The ST uses four parameters (Δx,Δy,S,θ) to model the translation in x-axis, the translation in y-axis, the scaling and the rotation respectively. The transformation is defined as follows:

$$\begin{pmatrix} x_r \\ y_r \end{pmatrix} = \begin{pmatrix} \Delta x & S\cos\theta & -S\sin\theta \\ \Delta y & S\sin\theta & S\cos\theta \end{pmatrix} \begin{pmatrix} 1 \\ x_c \\ y_c \end{pmatrix} \tag{3.3}$$

where $(x_r, y_r)^T$ and $(x_c, y_c)^T$ are the coordinates of a point in the reference and candidate image, respectively.

The ultimate goal is to achieve a final value indicating the similarity between the two feature point sets, in order to decide about the acceptance or the rejection of the hypothesis that both images correspond to the same indi-

vidual. To develop this task, the matching pairings between both images must be determined. A transformation has to be applied to the candidate image in order to register its feature points with respect to the corresponding points in the reference image. The set of possible transformations is built based on some restrictions and a matching process is performed for each one of them. The transformation with the highest matching score will be accepted as the best transformation.

To obtain the four parameters of a concrete ST, three pairs of feature points between the reference and candidate images are considered. The set of possible transformations is obtained by building an initial set similar to that described in [73]. Using Equation 3.3, the parameters of the transformation are computed. Defining both patterns as a set of feature points of different sizes: $\nu = \{v_1, v_2, ..., v_M\}$, $\nu' = \{v'_1, v'_2, ..., v'_N\}$, being M the total number of feature points in the reference image and N the total number of points in the candidate one, the size of the set T of possible transformations is computed using the following equation:

$$T = \frac{(M^2 - M)(N^2 - N)}{2} \tag{3.4}$$

Since T represents a high number of transformations, some restrictions must be applied in order to reduce it. As the scale factor between patterns is always very small in this acquisition process, a constraint can be set to the pairs of points to be associated. In this scenario, the distance between both points in each pattern has to be very similar. As it cannot be assumed that it will be the same, two thresholds are defined, S_{min} and S_{max}, to bound the scale factor. This way, elements from T are removed where the scale factor is greater or lower than the respective thresholds S_{min} and S_{max}. Equation 3.5 formalizes this restriction:

$$S_{min} < \frac{distance(v_1, v_2)}{distance(v'_1, v'_2)} < S_{max} \tag{3.5}$$

where v_1, v_2 are points from ν pattern, and v'_1, v'_2 are the matched points from the ν' pattern. Using this technique, the number of possible matches greatly decreases and, in consequence, the set of possible transformations decreases accordingly. The mean percentage of not considered transformations by these restrictions is around 70%.

In order to check feature points, a similarity measure between points (SIM) is defined. The distance between these two points will be used to compute that value. For two points v and v', their similarity value is defined by:

$$SIM(v, v') = 1 - \frac{distance(v, v')}{D_{max}} \tag{3.6}$$

where D_{max} is a threshold that stands for the maximum distance allowed for those points to be considered a possible match. This threshold controls the quality loss and discontinuities during the crease extraction process leading to mislocation of feature points by some pixels. Therefore, if $distance(v, v') > D_{max}$ then $SIM(v, v') = 0$.

In some cases, two points $v'_1, v'_2 \in \nu'$ could have both a good value of similarity with one point $v \in \nu$ in the reference pattern. This happens because v'_1 and v'_2 are close to each other in the candidate pattern. To identify the most suitable matching pair, the possibility of correspondence is defined comparing the similarity value between those points to the rest of similarity values of each one of them:

$$P(v, v') = \frac{SIM(v, v')^2}{\left(\sum_{i=1}^{M} SIM(v_i, v') + \sum_{j=1}^{N} SIM(v, v'_j) - SIM(v, v')\right)} \tag{3.7}$$

A matrix Q of size $M \times N$, is constructed such that position (i, j) holds $P(v_i, v'_j)$. Note that if the similarity value is 0, the possibility value is also 0. This means that only valid matchings will have a non-zero value in Q. The desired set C of matching feature points is obtained from P using a greedy algorithm. The element (i, j) inserted in C is the position in Q where the maximum value is stored. Then, to prevent the selection of the same point in one of the images again, the row (i) and the column (j) associated to that pair are set to 0. The algorithm finishes when no more non-zero elements can be selected from Q. The final set of matched points between patterns is C.

Consider the next example to illustrate the process: Q_0 (4×5) is the initial similarity matrix for two patterns of sizes 4 and 5:

$$Q_0 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0.1 & 0 & 0.2 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.4 & 0.7 \end{pmatrix} \tag{3.8}$$

The highest similarity is 0.7 pertaining to row 4 and column 5, thus the set of matched points, C, is initialized and the similarity matrix in the next iteration, Q_1 , is updated with zeros in row 4 and column 5:

$$Q_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0.1 & 0 & 0.2 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, C = [(4, 5)] \tag{3.9}$$

The now highest value (0.5) is detected and its row and column are updated along with C:

$$Q_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, C = [(4,5),(3,1)] \tag{3.10}$$

Finally, the last available value is picked (0.2 in the position $(2,3)$), Q_3 only contains zeros so the process stops with the final matched points set being:

$$C = [(4,5),(3,1),(2,3)] \tag{3.11}$$

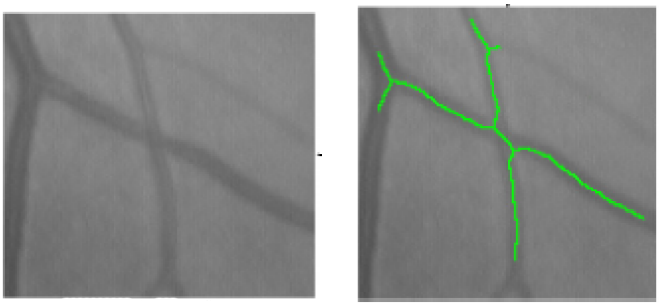
Using the information of matched points, a similarity metric must be established to obtain a final criterion of comparison between patterns. Performance of several metrics using these data is analyzed in the next chapter.

3.2.1. Matching optimization: Feature point classification

The matching process weakness is the high amount of transformations to be computed in order to find the optimal one. To reduce the amount of computation time, more information about the domain is introduced. By characterizing the feature points into crossovers and bifurcations, an efficient constraint can be added only allowing points of the same class (or unclassified) to match.

A similar approach to the one introduced in this work for the detection based on the Intersection Number (Equation 3.2) is taken in [3] for the classification between bifurcations and crossovers in such a way that every point with $I(v) = 3$ is classified as a bifurcation and points with $I(v) = 4$ are classified as crossovers. However, this criterion used to classify the points makes unlikely to classify a point as crossover. In crossovers, on the contrary to bifurcations, the vessels that create the feature point do not coincide in the same pixel in

the skeleton. This is due to the angle and width of the vessels involved in the feature point. They cause, in most cases, that the central axis of the vessels do not intersect in the same pixel. This problem, shown in Figure 3.17, produces a misclassification of crossovers as two close bifurcations.



(a) (b)

Figure 3.17: Problem of the representation, in the skeleton, of a crossover as two bifurcations. (a) Crossover in the original image. (b) Skeleton over the original image.

To solve this problem and produce a more robust and valid classification it is needed to perform a further analysis on the feature points. This classification is done according to local features of the points and a topological analysis necessary to spot the cases of the close bifurcations. Feature point classification is done in two parts. In a first step the points are labelled according only to environment features, this is, using only local information. In a second step, this classification is refined using information of the relationship between points, understood this information as the distance between points.

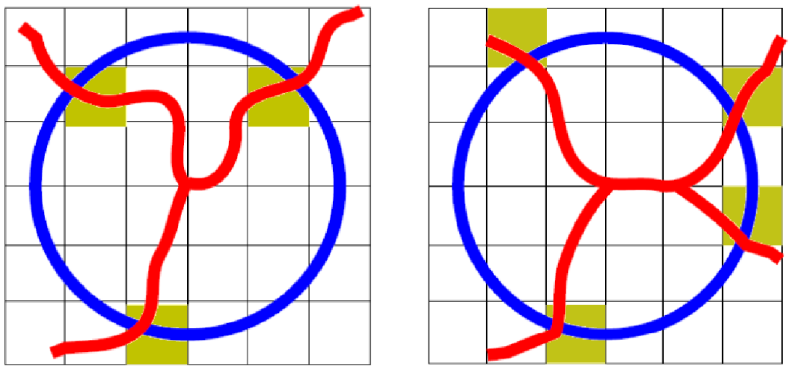
- Local analysis for classification

The first classification step is done according to local features of the points without considering the effect of the other points. So, to define a classification

for a point, the number of vessel segments that create the intersection is studied. Each detected feature point, v, is used as center of a circumference with radius R_c used for the analysis. $n(v)$ gives the number of vessel segments that intersect the circumference being the point, v, classified as follows,

- v classified as bifurcation candidate $\Leftrightarrow n(v) = 3$
- v classified as crossover candidate $\Leftrightarrow n(v) = 4$

Figure 3.18 shows these two possible classifications. The images shows the blood vessels, the circumference used to do the analysis, and, colored darker, the pixels where the vessels intersect the circumference.



(a) (b)

Figure 3.18: Preliminary feature point classification according to the number of vessel intersections where (a) represents a bifurcation and (b), a crossover.

This classification has some problems when classifying crossovers due to its representation as two close bifurcations, previously explained and shown in Figure 3.17. Due to this, the radius R_c of the circumference used to the presented classification has to be big enough to be intersected by the two bifurcations of each crossover and, in this way, be intersected by the four vessel segments to be classified as a crossover. The problem of increasing the

size of the radius is that, due to the complexity of the vascular structure and the classification method used, the circumference can be intersected by vessels that do not belong to the feature point analyzed and, because of this, the point is wrongly classified. To avoid this problem, a vote system with three radius sizes is used (Figure 3.19).

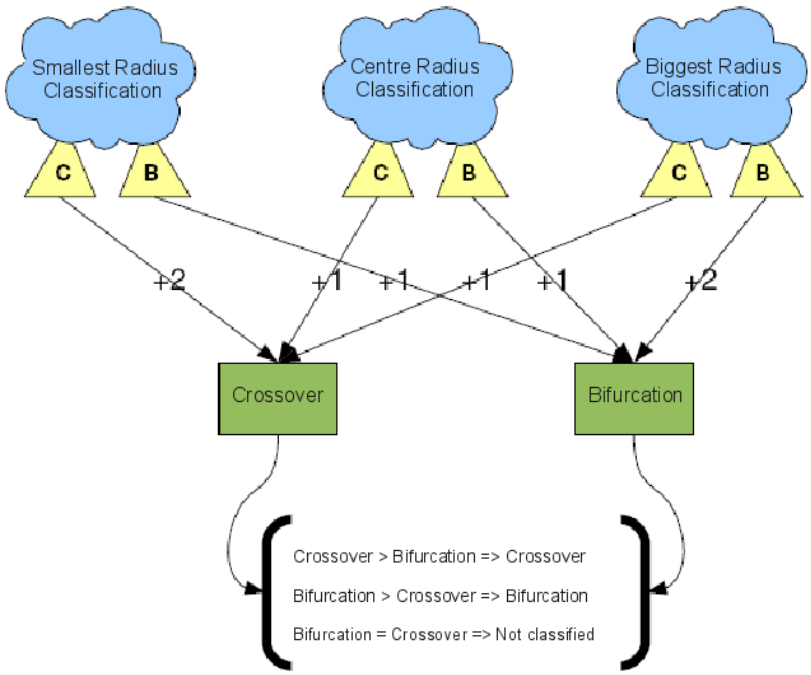


Figure 3.19: Voting system for classification. Small radius information is more significant for crossovers while big radius is for bifurcations.

In the vote system, three different classifications according to three different radius sizes (R_1, R_2, R_3) for the analysis of the point are set. The selected radius are defined as: $R_1 = R_c - \rho$, $R_2 = R_c$ and $R_3 = R_c + \rho$ where ρ is a fixed amount. With these definitions, two values are calculated, $C(v)$ and $B(v)$, meaning the number of votes for a point v to be classified, respectively, as a crossover and a bifurcation:

$$C(v) = 2 * C(v, R_1) + C(v, R_2) + C(v, R_3) \tag{3.12}$$

$$B(v) = B(v, R_1) + B(v, R_2) + 2 * B(v, R_3) \tag{3.13}$$

where $C(v, R_i)$ and $B(v, R_i)$ are binary values indicating if v is classified, respectively, as a crossover or a bifurcation using a radius R_i. Note that the contribution of the smallest radius is more valuable, and therefore weighted, in the crossover classification while for bifurcations the biggest radius adds more information. Feature point v will be classified as a crossover when $C(v) > B(v)$ and as a bifurcation otherwise.

As shown before, the contribution is not the same for all classifications of all radius because, the smaller the radius is that classifies the feature point as a crossover, the more probability to be properly classified. This is due to the fact that, the bigger the radius of the circumference is, the more likely is to be intersected by a vessel segment that does not belong to the feature point. If a point is classified as a crossover using the small radius (R_1), as this is less probable, it adds more information to the final result. Analogously, and because all the crossovers are represented as two bifurcations, the classification of a feature point as a bifurcation is more significant in the final result.

As a result of this step, a preliminary classification of the feature point detected in the vascular structure is given. This classification will guide the last analysis in the next step.

- Topological analysis for classification

The previously presented classification method, as stated before, is only based on the local features of the point to classify it. However, due to the rep-

resentation of crossovers in the skeleton (see Figure 3.17), this information is not enough to properly classify a crossover while being a necessary condition. According to this, a topological classification is needed to analyze the feature points in pairs.

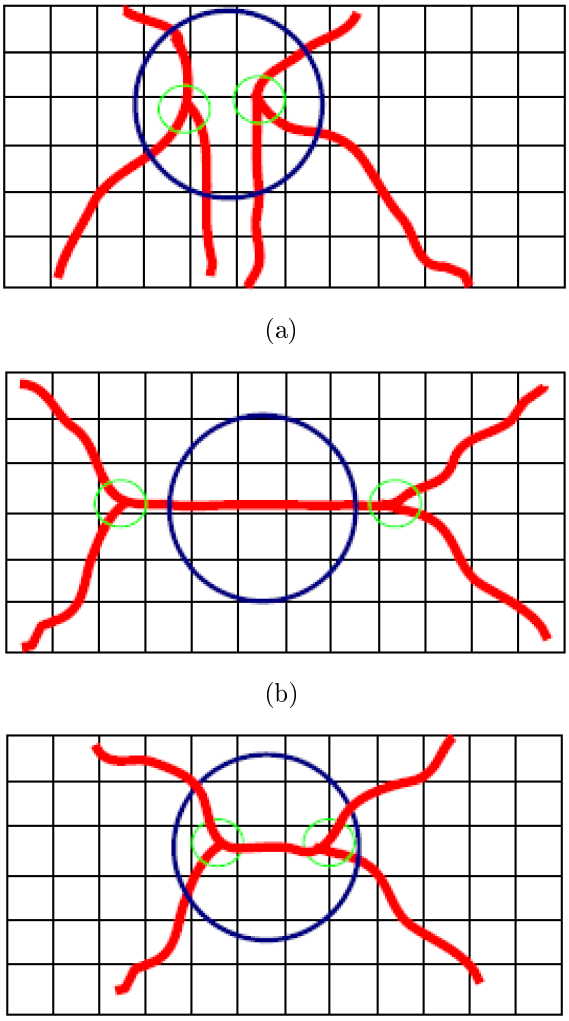
In order to group the feature points in pairs, a criterium based on Euclidean distance, d, is set. In particular, two feature points (v_i, v_j) are paired when they minimize $d(v_i, v_j)$ among the available feature points, this is, the feature points not paired in a former iteration.

Each of these pairs of points must satisfy two conditions to be classified as a crossover:

- Both points must be connected. This is, it must exist a vessel segment between the two points.
- $d(v_i, v_j) \leqslant 2 * R_c$, this is, they must be close enough to fall inside a circle of radius R_c

Figure 3.20 shows different cases of the process representing the blood vessels. The position of the feature points are shown in light grey and the circumference with the maximum radius considered for crossover classification is shown in darker grey.

Up to this stage, the feature point classification step was centered in whether or not a pair of points must be classified as a crossover in the real image. The goal of this step is to find the real position of these crossovers that, obviously, does not match any of the two bifurcations of the pair. Empirically, it was determined that a good approach was to set the real point for the crossover in the middle point of the segment that joins the two bifurcations. Figure 3.21



(a)

(b)

(c)

Figure 3.20: Schema of the different cases in the crossover classification. (a) shows a feature point that fulfils the condition of distance but not connectivity, so it is not classified as a crossover. (b) shows a feature point that fulfils the condition of connectivity but not distance, so it is not classified as a crossover. Finally, (c) shows a feature point classified as a crossover since it fulfills the conditions of distance and connectivity.

shows the final position of the crossover, marked over the real retina image.

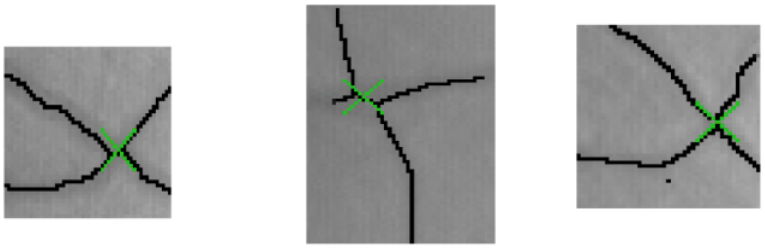


Figure 3.21: Position of the real crossover point computed as the middle point of the segment between the bifurcations. The real point, lighter, is shown over the real image and the skeleton, in black.

At this point of the process, it could be assured that every point not classified as a crossover is a bifurcation. Having classified the crossovers in the previous step, the classification of every point could be considered complete. However, assuming that every point not marked as crossover is a bifurcation makes the classification of the last ones too dependant of the success of the crossover classification. So, if this idea is accepted, for each misclassified crossover in the previous step two bifurcations would be misclassified. For this reason, it is necessary to use another threshold (R_b) to take a decision of which points are accepted as bifurcations.

This process is analogous to the previously presented for crossover classification. In this way, for each pair of feature points, a circumference with radius R_b centered in the middle point of the segment between the points is used. This circumference must not contain both points in order to be considered as bifurcations. Every pair of points not fulfilling the conditions is marked as not classified in the final result. This is due to the fact that points are not close enough to be considered as one crossover but not far enough to be considered

as two independent bifurcations.

Once finished the feature point classification, three categories can be distinguished (shown in Figure 3.22),

- The feature points classified as crossovers, those fulfilling the conditions of morphology and proximity (Figure 3.22(a)).
- The feature points classified as bifurcations, those that fulfilling the conditions of morphology are further than the established threshold (Figure 3.22(b)).
- The feature points not classified, those that not being close enough to be classified as crossovers and not either far enough to be classified as bifurcations (Figure 3.22(c)).

Figure 3.23 shows the result of the detection and classification of the feature points of the vascular structure in an eye fundus image. The points classified as crossovers are marked with an asterisk. The points classified as bifurcations are marked with a circle. Finally, the points not classified are marked with a square. Note that the R_c and R_b parameters allow to tune the system in terms of specificity and sensitivity. In the authentication domain a high specificity is preferred over sensitivity to avoid a worse decision performance in the authentication task. In the next chapter, performance of the classifier is studied as a function of both radius to determine the optimal setup.

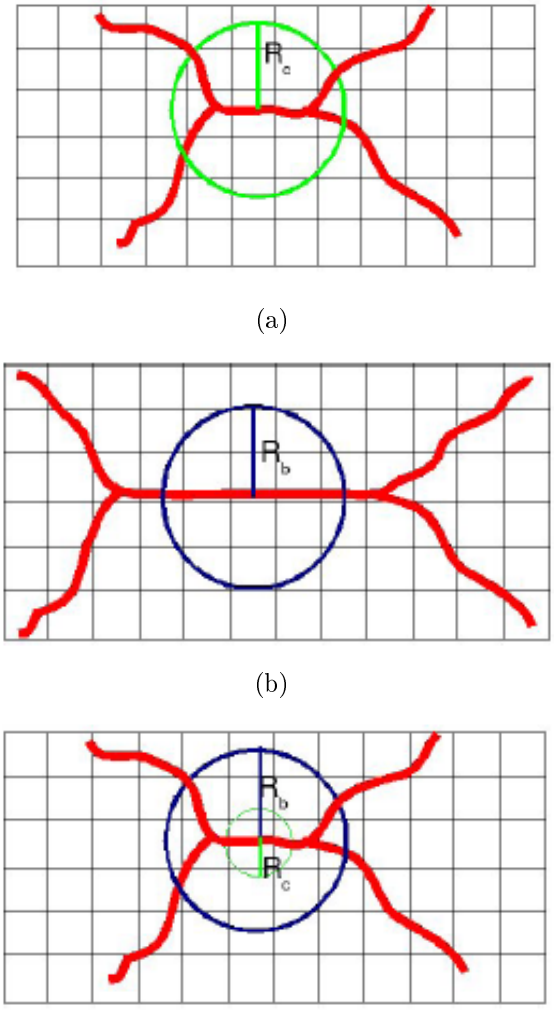


Figure 3.22: Final categories of the feature points where (a) shows a point classified as crossover, (b) two independent bifurcations and (c) two feature points not classified

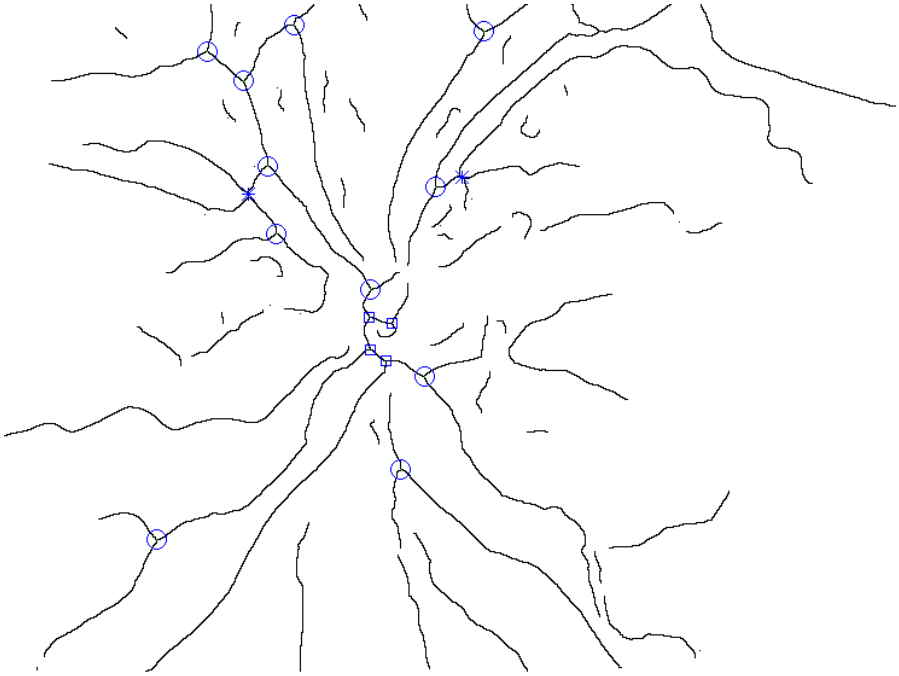


Figure 3.23: Classification of the feature points. Circles mark bifurcations, crosses mark crossovers and squares mark unclassified points.

Chapter 4

Experimental Results

In this chapter, a series of experiments are presented in order to test the different techniques discussed in the work. The final objective is to evaluate the global performance of the methodology for personal authentication based on retinal tree feature points. With that goal in mind, each approach introduced in the different stages of the system will be evaluated to verify the suitability of such proposals.

The first stage of the evaluation is the biometric pattern processing, in our case the feature point extraction and characterization. For the extraction part, two different techniques were introduced, the first one based on crease extraction and the second one based on vessel tree segmentation. Both approaches will be compared in terms of effectiveness and efficiency. In the characterization part, the goal will be to evaluate the performance of the feature point classification. This performance analysis covers two different issues: first, the classification error in terms of sensitivity and specificity depending on the control parameters discussed in section 3.2.1. Second, the analysis of the computational improvement during the registration stage associated to the constraint of the feature point type. The possible error in the authentication task introduced due to the characterization stage is analyzed at the end of the

chapter.

Once the feature points is addressed, the second part of the experiments will be focused on testing the matching capabilities of the authentication system. The similarity metrics determine the system ability to successfully classify authentications as authorized or unauthorized. Thus, an analysis of the performance of similarity metrics is done in order to maximize the confidence band between match (authorized) and non-match (unauthorized) distributions.

In the final experiment, a global test will be performed with the final parameters on the whole test set in order to analyze the capabilities of the system. In this part, we include the analysis associated to the characterization of feature points to assess the impact on the authentication error variation of the biometric system related to the error in the characterization stage.

The images used for the experiments were extracted from the VARIA database [84]. This database was created during the development of this thesis and consists of 233 retinal images from 139 different individuals, 59 of which possess at least two samples. The images have been acquired over a span of several years with a TopCon NW-100 model non-mydriatic retinal camera. They are optic disc centered with a resolution of 768x584. These images have a high variability in contrast and illumination, allowing the system to be tested in quite hard conditions, and simulating a more realistic environment. The different conditions are also due to the fact that different experts with different illumination configurations on the camera have acquired the images. Figure 4.1 illustrates several examples of retinal images from the database where the heterogeneity between samples can be observed.

All the tests were run using a Pentium IV 2.4Ghz Desktop PC with 2Gb of RAM memory. The following sections of this chapter will further discuss the experiments and results obtained for the different approaches introduced in this work.

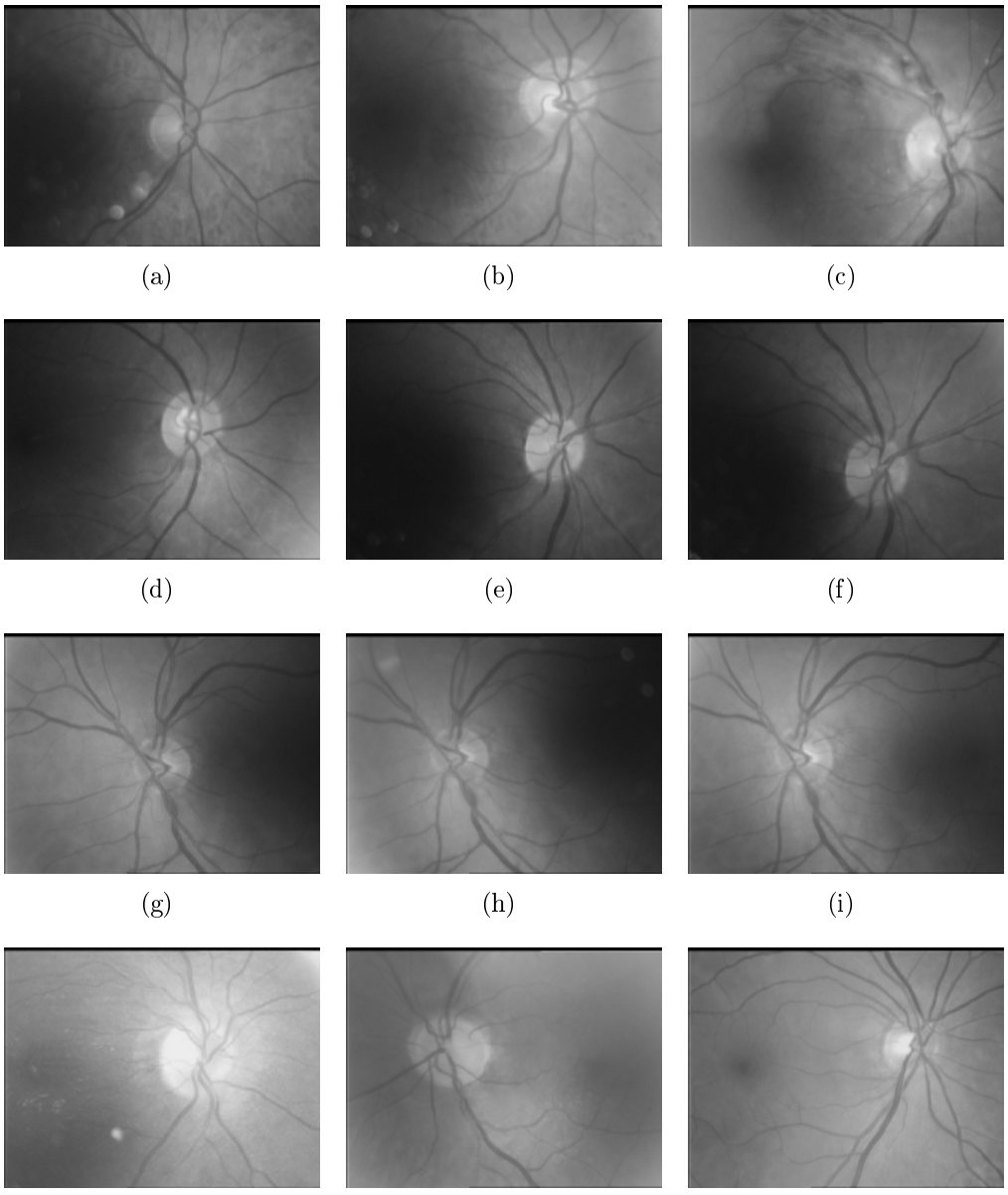


Figure 4.1: Retinal images extracted from VARIA database [84], created and used in this work for the validation experiments. The database contains 233 images from several users and the heterogeneity between samples is very high, simulating a real-case scenario where there is not an exhaustive control of the acquisition environment. In the examples, this heterogeneity can be observed even in cases of images from the same individual as 4.1(e),4.1(f) or 4.1(g),4.1(h),4.1(i).

4.1. Feature point validation

4.1.1. Feature point extraction

To evaluate both proposals for the feature point extraction (based in creases and based in vessel tree segmentation), we define some quality metrics the methods must maximize. The metrics attend to three main quality criteria in the point detection process:

1. *Efficacy.* The goal is to detect as many feature points contained within the image as possible while avoiding to detect false points.

   The two proposed metrics to control the efficacy are *precision* and *recall*. Precision measures the rate of correctly detected points versus all detected points. The higher precision the lower the percentage of spurious points. Recall, in the other hand, measures the rate of existing feature points that are correctly detected. In probabilistic terms, precision measures the probability that a detected point is a valid one while recall indicates probability that a valid point is detected:

   $$Precision = \frac{TP}{TP + FP} \tag{4.1}$$

   $$Recall = \frac{TP}{TP + FN} \tag{4.2}$$

   where TP (True Positives) refers to correctly detected points, FP (False Positives) refers to spurious points and FN (False Negatives) refers to undetected existing points.

2. *Efficiency.* The computation should be light and, therefore, the running time low.

The most interesting metric is the averaged running time of the process, taking also into account the minimum and maximum times along with the variance. The tests must be run several times to reduce the effect of external factors in the total time.

3. *Accuracy.* The points should be detected in their correct position on the image. This criterium is measured in terms of the deviation of the detected point from the original existing point location. This deviation is expressed as a Euclidean distance (in pixels):

$$Deviation = \sqrt{(D_x - O_x)^2 + (D_y - O_y)^2} \tag{4.3}$$

where (D_x, D_y) are the coordinates where the point was detected and (O_x, O_y) is the original position of the feature point.

Initially, 50 images were randomly selected from VARIA database to tune the parameters in both methods. Feature points were labelled by medical experts to allow to compute the precision and recall rates. In order to make execution times more reliable, the tests were repeated 100 times for each image and method in a random order.

Table 4.1 shows the best results obtained with and without running the spurious point filtering algorithm using the crease extraction approach. The filtering of points improves the precision significantly enough without affecting the recall, i.e. practically no real points are lost during the filtering. Table 4.2 shows the best values for the parameters using the 50 images.

Analogously, Table 4.3 shows the best results obtained with and without running the spurious point filtering algorithm (branch pruning) for the segmentation approach. The filtering of points improves the specificity even more significantly than the crease case due to the nature itself of the skeletonization. In fact, the filtering stage with this approach is unavoidable because the

	Efficacy		Efficiency			Accuracy
	Recall	**Precision**	T_{avg}	T_{min}	T_{max}	**Deviation**
No filtering	84.8%	84.7%	0.512s	0.348s	0.652s	6.44px
Filtering	82.2%	93.9%	0.564s	0.411s	0.712s	5.97px

Table 4.1: Performance of crease-based approach for feature point extraction using metrics of efficacy, efficiency and accuracy. First row represents the results obtained without applying filtering of points and the second row applying the filtering.

Parameter	Description	Value
θ_{min}	Minimum angle to accept an union	$\frac{3}{4}\pi rad$
l_{max}	Maximum extension to search for bifurcations	25 pixels
T_{min}	Minimum segment length be kept	18 pixels

Table 4.2: Parameter configuration for the feature point extraction using the crease extraction approach.

precision would be too low otherwise. Table 4.4 shows the best values for the parameters using the same set of images.

Once the parameters have been tuned, both methods were tested using the whole VARIA database. Table 4.5 shows how both methods of feature point extraction performed using the previously exposed quality metrics for the database.

The extraction method based in segmentation offers slightly better results in terms of detected points and accuracy. However, the crease based method is very efficient because the analysis of the vessel lines is made only in terms

	Efficacy		Efficiency			Accuracy
	Recall	**Precision**	T_{avg}	T_{min}	T_{max}	**Deviation**
No filtering	93.2%	71.7%	4.14s	1.84s	5.93s	4.77px
Filtering	91.5%	99.2%	4.56s	2.41s	6.11s	3.69px

Table 4.3: Performance of segmentation-based approach for feature points extraction using metrics of efficacy, efficiency and accuracy. First row represents the results obtained without applying filtering of branches and the second row applying the filtering.

Parameter	Description	Value
N	# of dilations to fill holes in the segmented tree	4
ζ	Threshold to prune tree branches	15 pixels

Table 4.4: Parameter configuration for the feature point extraction using the vessel segmentation approach.

	Efficacy		Efficiency			Accuracy
	Recall	**Precision**	T_{avg}	T_{min}	T_{max}	**Deviation**
Creases	83.5%	93.7%	0.542s	0.386s	0.725s	6.23px
Segment.	89.7%	98.9%	4.68s	2.21s	6.36s	3.72px

Table 4.5: Performance of creases- and segmentation-based approaches for feature point extraction using metrics of efficacy, efficiency and accuracy.

of segment endpoints. This greatly reduces the processing as features are only searched for every endpoint instead of tracking the whole vessels. In conclusion, the crease method is more appropriate due to its applicability in a wider range of devices and needs. The segmentation based method could be useful in situations where the pattern extracted from the creases is incomplete due to quality image issues.

4.1.2. Feature point characterization

In this section the classification of feature points into bifurcations and crossovers is analyzed, in terms of classification performance and impact on the authentication system. For the analysis of the methodology, the same set of 50 images from experiment 4.1.1 was used.

For the quantification of the results obtained in these steps, we use the sensitivity and specificity measures. Both concepts are used in terms of precision related to total detected feature points, excluding end points. Sensitivity measures the capacity to detect the feature when it exists while specificity measures the capacity not to detect the feature when it does not exists. Im-

age preprocessing parameters are necessary to the correct performance of later steps.

Radius parameters, R_c and R_b, tune the permissiveness in the classification task and, thus, a deeper analysis of the influence of both parameters is presented. The previously used images in the detection process are used now for obtaining these results. The parameter R_c is the radius of the circumference centered in the feature point analyzed. This value affects directly to the point classification increasing the crossover detection probability proportional to the radius size. For the bifurcation, the parameter to consider is R_b, that represents the minimum distance two bifurcation must be separated to be considered as two independent bifurcations. This is, the smaller the R_b is, the more bifurcations are classified.

For a right choice of the mentioned parameters, R_c and R_b, a quantitative study is presented. The results allow to select the adequate parameters for a specific domain where the desired sensibility or specificity levels can change. Table 4.6 shows the results for crossover classification for 1116 feature points (846 bifurcations and 270 crossovers), according to the chosen R_c radius and Figure 4.2(a) shows how can the parameter be adjusted to fit different domains.

Radius(R_c)	**T.P.**	**F.N.**	**F.P.**	**Sensitivity**	**Specificity**
5	54	216	0	20,00%	100%
10	108	162	0	40,00%	100%
15	162	108	0	60,00%	100%
20	204	66	18	75,56%	96,20%
25	204	66	54	75,56%	94,94%

Table 4.6: Results for crossover classification according to R_c showing True Positives (T.P.), False Negatives (F.N.) and False Positives (F.P.)

The table shows how the number of correct classified crossovers increases with the radius size. This tendency could throw the idea of increasing the

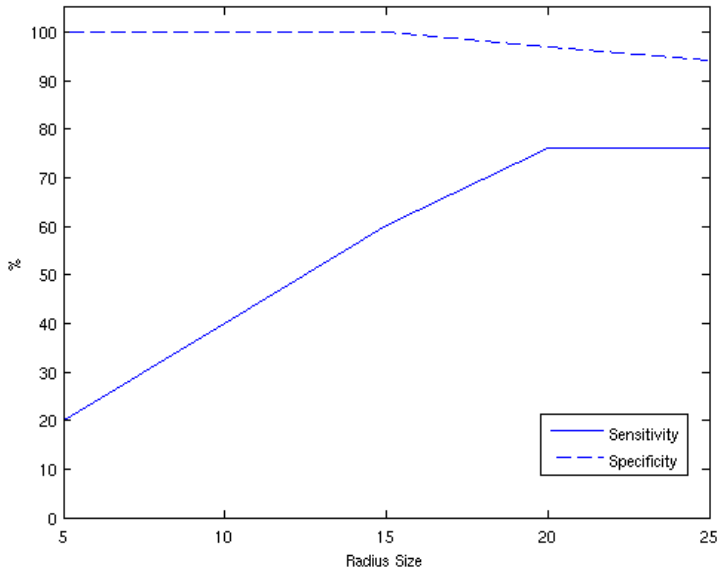
radius size until obtaining a big number of classified crossovers, however, increasing the radius also increases the number of misclassified crossovers. For the remaining 978 feature points with 846 bifurcations and 66 crossovers (that means 132 points due to crossover representation) and with $R_c = 20$ as the selected value, table 4.7 shows the results for bifurcation classification according to the chosen R_b radius. This table shows a new category, the non classified points, that includes the points that fulfilling the morphology conditions are not close enough to be classified as crossover and not far enough not to be classified as independent bifurcations.

Radius(R_b)	**T.P.**	**F.N.**	**F.P.**	**N.C.**	**Sensitivity**	**Specificity**
25	726	120	132	84	85,82%	0,00%
30	690	156	48	216	80,85%	63,63%
35	630	216	12	324	74,47%	90,90%
40	534	312	12	396	63,12%	90,90%

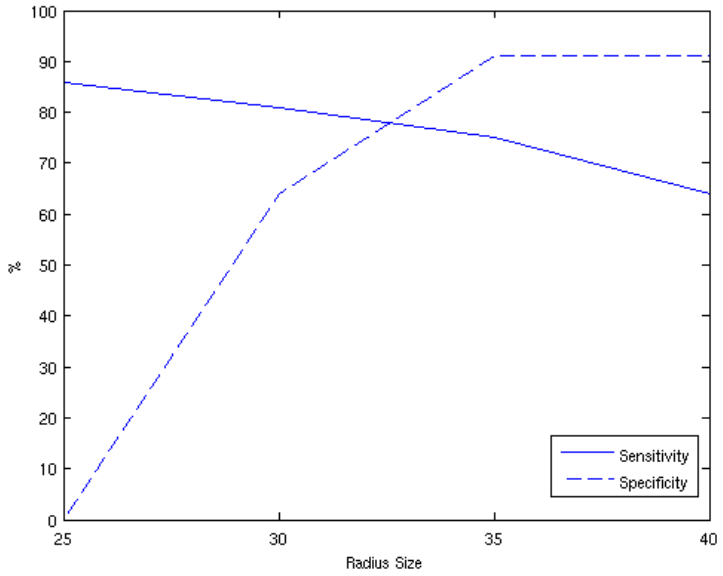
Table 4.7: Results for bifurcation classification according to R_b showing True Positives (T.P.), False Negatives (F.N.), False Positives (F.P.) and Not Classified (N.C.)

The results, represented in Figure 4.2(b), show how the parameter R_b allows to adjust the results according to the domain. The bigger R_b is, the more number of points not classified although the number of false positives will be below 1%. Opposite to this, if a big level of true positives is needed with a small radius, the sensitivity is over 80%.

The technique proposed in [3] showed a problem with the crossover misclassification due to the skeleton representation, where a crossover turns into two close bifurcations. This work does not offer quantitative results. However, our implementation of this technique shows that nearly every point is classified as a bifurcation, being capable to classify correctly only 3% of the crossovers. The work proposed in [34] extracts the structure using a vessel tracking based with the results shown in Table 4.8. Other previous techniques do not offer



(a)



(b)

Figure 4.2: (a) Influence of the parameter R_c in crossover classification. (b) Influence of the parameter R_b in bifurcation classification after the crossover stage.

quantitative results in the characterization task to compare with so Table 4.8 compares the obtained results for the VARIA database to the results given in [34]. The main improvement comes in the crossover rate, due to the radius proposed. In general, the system exhibits a very high specificity rate for both classes making it specially suitable for the registration task as, otherwise, the classification would introduce errors in this stage leading occasionally to an inaccurate registration.

	Bifurcations		**Crossovers**	
	Sensitivity	**Specificity**	**Sensitivity**	**Specificity**
E. Grisan *et al.*	76%	87%	62%	74%
Our work	74%	92%	77%	97%

Table 4.8: Obtained results for characterization of feature points compared to the results presented in [34]

Nevertheless, the impact of the characterization in the computation load of the matching process is the main objective to evaluate. By avoiding matchings between differently classified points, a lot of irrelevant transformations are automatically removed from the computation thus reducing the computation process. Table 4.9 shows the impact in the VARIA database. The average reduction of computed transformations was 21.36%.

Total	Removed	Mean	Std
240593	51389	21.36%	7.58%

Table 4.9: Statistics on the transformations removed in the matching process. The columns refer to total possible transformations without considering the point classification restriction, the number of transformations avoided by including the restriction and the mean and standard deviation (std) percentages of transformation removed per image.

4.2. Similarity metric analysis

In this section, the goal is to analyze the identity verification capabilities of the system. For that purpose, we define a series of similarity metrics and evaluate them using retinal images from the VARIA database. These similarity metrics are defined incrementally as one is usually the result of refining the previous one in order to extract a better performance out of the available information from the matching process. In our case, the data obtained via the matching process are the matched points (C) and sizes of both patterns to compare (M and N).

For the metric analysis, a set of 150 images (100 images, 2 images per individual and 50 different images more) from VARIA database were used. The remaining images will be used for testing global performance at the end of this chapter. In order to build the training set of matchings, all images are matched versus all the images (a total of 150x150 matchings) for each metric. The matchings are classified into attacks or clients accesses depending if the images belong to the same individual or not. Distributions of similarity values for both classes are compared in order to analyze the classification capabilities of the metrics.

Based on the matched points (C) and pattern sizes (M,N) the following similarity metrics are presented to test the system:

- C, number of matched points
- $S = \frac{C}{f(M,N)}$, where C is normalized using some function depending on pattern sizes. Three functions $f(M,N)$ are defined:
  - $f(M,N) = min(M,N)$
  - $f(M,N) = \frac{M+N}{2}$
  - $f(M,N) = \sqrt{MN}$

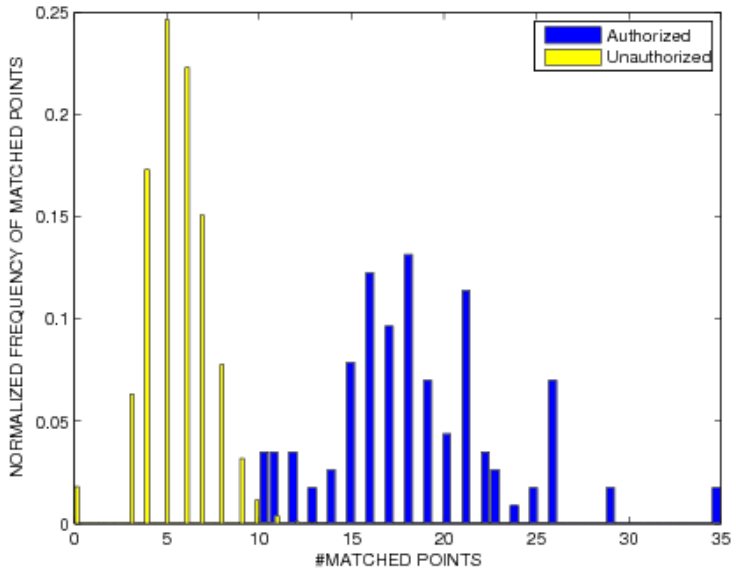
- $S_\gamma = S \cdot C^{\gamma-1} = \frac{C^\gamma}{f(M,N)}$: a parameter γ is introduced to tune the absolute influence of number of matched points independently of the pattern size. This metric is normalized in two ways:
  - $S_{\gamma R} = min\left\{\frac{S_\gamma}{R}, 1\right\}$, dividing by a *safe* reference value and setting to 1 all metrics with values above it.
  - $S_{\gamma T} = T(S_\gamma)$ where $T(x) = \frac{1}{1+e^{s\cdot(x-0.5)}}$ is a sigmoidal transference function.

4.2.1. Matched points

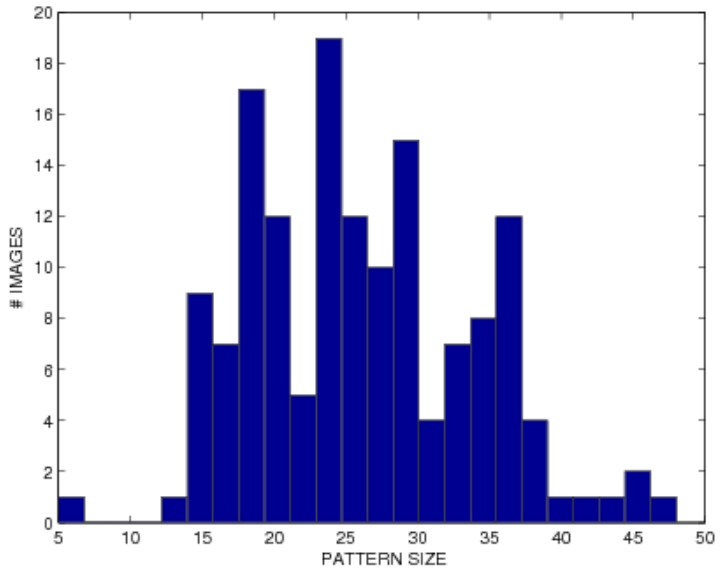
The main information to measure similarity between two patterns is the number of feature points successfully matched between them. Figure 4.3(a) shows the histogram of matched points for both classes of authentications in the training set. As it can be observed, matched point information is by itself quite significant but insufficient to completely separate both populations since there is an overlapping between them in the interval $[10, 13]$.

This overlapping is caused by the variability of the patterns size in the training set due to the different illumination and contrast conditions in the acquisition stage. Figure 4.3(b) shows the histogram for the biometric pattern size, i.e. the number of feature points detected. A high variability can be observed, as some patterns have more than twice the number of feature points of other patterns. As a result, some patterns have a small size, capping the possible number of matched points (Figure 4.4). Also, using the matched points information alone lacks a well bounded and normalized metric space.

To combine information of patterns size and normalize the metric, a function f will be used. Normalized metrics [15] are very common as they make easier to compare class separability or establishing valid thresholds as the following section shows.

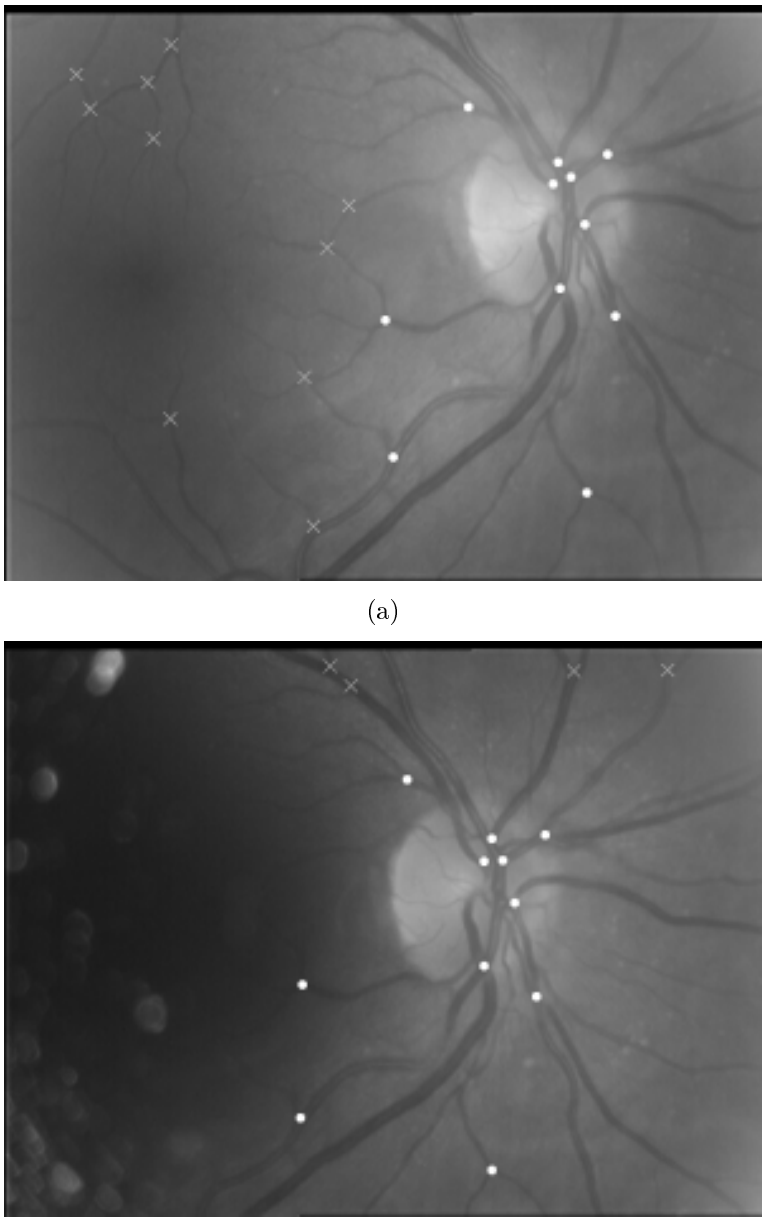


(a)



(b)

Figure 4.3: (a) Matched points histogram in the attack (unauthorized) and client (authorized) authentication cases. Both distributions overlap in the interval $[10, 13]$. (b) histogram of detected points for the patterns extracted from the training set.



(a)

(b)

Figure 4.4: Example of matching between two samples from the same individual in the VARIA database. White circles mark the matched points between both images while crosses mark the unmatched points. In (b) the illumination conditions of the image lead to miss some features from the left region of the image. Therefore, a small amount of detected feature points is obtained capping the total amount of matched points.

4.2.2. Normalized metrics

The normalized similarity measure (S) between two patterns is defined by:

$$S = \frac{C}{f(M,N)} \tag{4.4}$$

where C is the number of matched points between patterns, and M and N are the matching patterns sizes. The first f function defined and tested is:

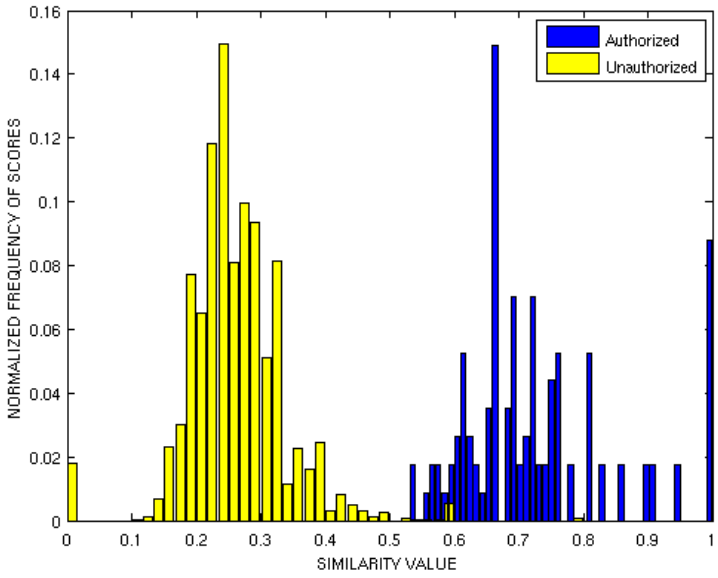
$$f(M,N) = min(M,N) \tag{4.5}$$

The *min* function is the less conservative as it allows to obtain a maximum similarity even in cases of different sized patterns. Figure 4.5(a) shows the distributions of similarity scores for clients and attacks classes in the training set using the normalization function defined in Equation 4.5, and Figure 4.5(b) shows the FAR and FRR curves versus the decision threshold.

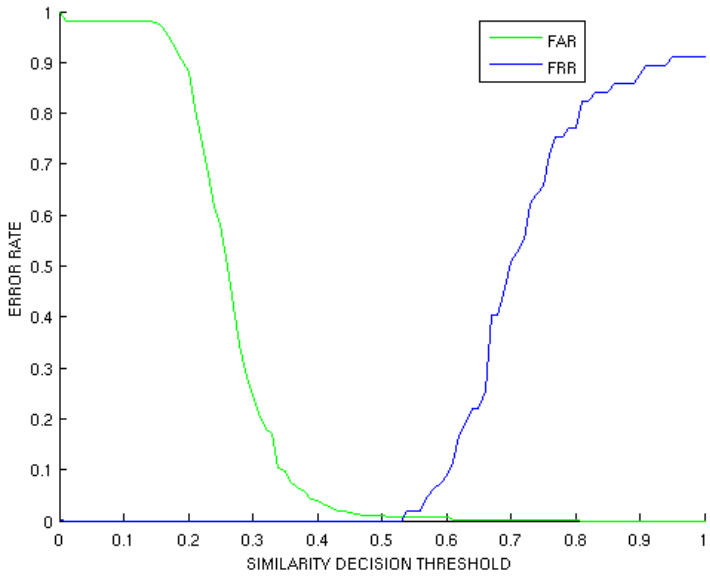
Although the results are good when using the normalization function defined in Equation 4.5, a few cases of attacks show high similarity values, overlapping with the client class. This is caused by matchings involving patterns with a low number of feature points. As $min(M,N)$ is usally very small, only a few points are needed to match in order to get a high similarity value. This suggests, as it will be later revised in this chapter, that some minimum quality constraint in terms of detected points would improve performance for this metric.

A new normalization function is defined to cope with the issue of low sized patterns:

$$f(M,N) = \frac{M+N}{2} \tag{4.6}$$



(a)



(b)

Figure 4.5: (a) Similarity value distribution for authorized and unauthorized accesses using $f = min(M, N)$ as the normalization function for the metric. (b) False Accept Rate (FAR) and False Rejection Rate (FRR) for the same metric.

Using the arithmetic mean, pattern sizes are combined to reduce the impact of patterns with a small number of points. Figure 4.6(a) shows the distributions of similarity scores for client and attack classes in the training set using the normalization function defined in Equation 4.6 and Figure 4.6(b) shows the FAR and FRR curves versus the decision threshold.

To further improve the class separability, a new normalization function f is defined:

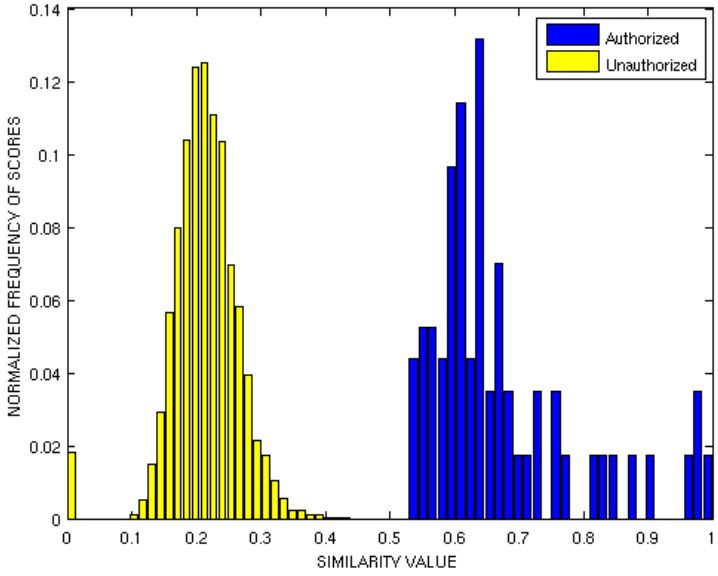
$$f(M, N) = \sqrt{MN} \tag{4.7}$$

Figure4.7(a) shows the distributions of similarity scores for clients and attacks classes in the training set using the normalization function defined in Equation 4.7 and Figure 4.7(b) shows the FAR and FRR curves versus the decision threshold.

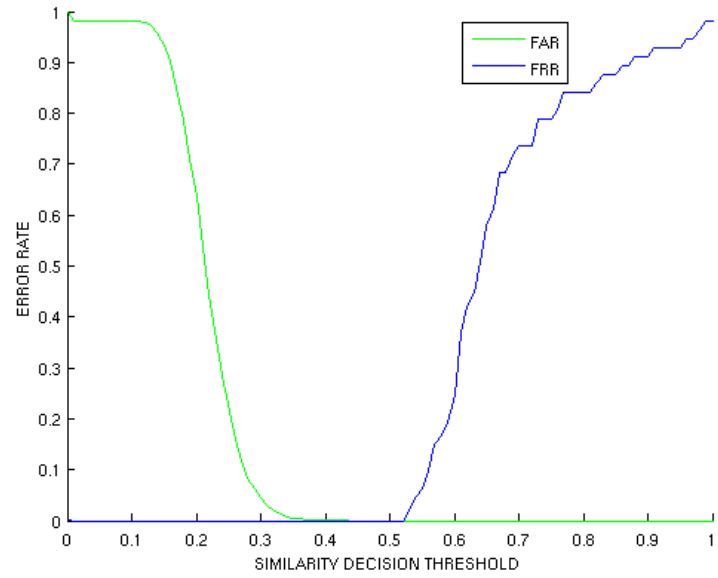
Function defined in Equation(4.7) combines both patterns size in a more conservative way, preventing the system to obtain a high similarity value if one pattern in the matching process contains a low number of points. This reduces the attack class variability and, moreover, separates its values away from the client class as this class remains in a similar value range. As a result of the new attack class boundaries, a decision threshold can be safely established where $FAR = FRR = 0$ in the interval $[0.38, 0.5]$, as Figure 4.7(b) clearly exposes. Although this metric shows good results, it also has some issues due to the normalization process, which can be corrected by tuning the impact of the normalization.

4.2.3. Gamma weighted metrics

Normalizing the metric has the side effect of reducing the similarity between patterns of the same individual where one of them had a much greater number

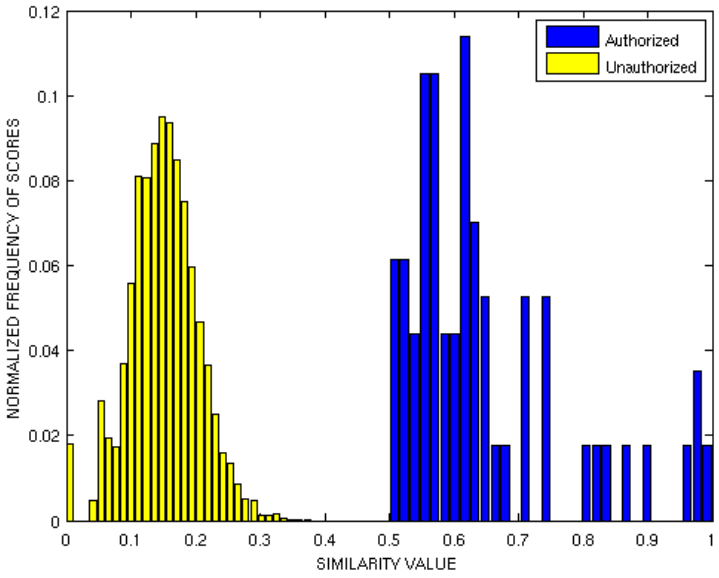


(a)

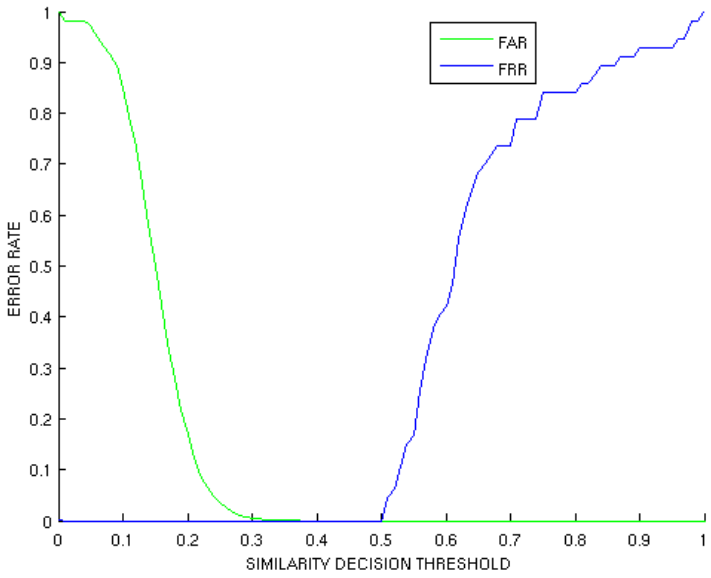


(b)

Figure 4.6: (a) Distribution of similarity values for authorized and unauthorized accesses using $f = \frac{M+N}{2}$ as normalization function for the metric. (b) False Accept Rate (FAR) and False Rejection Rate (FRR) for the same metric.



(a)



(b)

Figure 4.7: (a) Distribution of similarity values for authorized and unauthorized accesses using $f = \sqrt{MN}$ as normalization function for the metric. (b) False Accept Rate (FAR) and False Rejection Rate (FRR) for the same metric.

of points than the other, even in cases with a high number of matched points. This means that some cases easily distinguishable based on the number of matched points are now near the confidence band borders. To take a closer look at this region surrounding the confidence band, the cases of unauthorized accesses with the highest similarity values (S) and authorized accesses with the lowest values are evaluated. Figure 4.8 shows the histogram of matched points for cases in the marked region of Figure 4.7(b). It can be observed that there is an overlapping but both histograms are highly distinguishable.

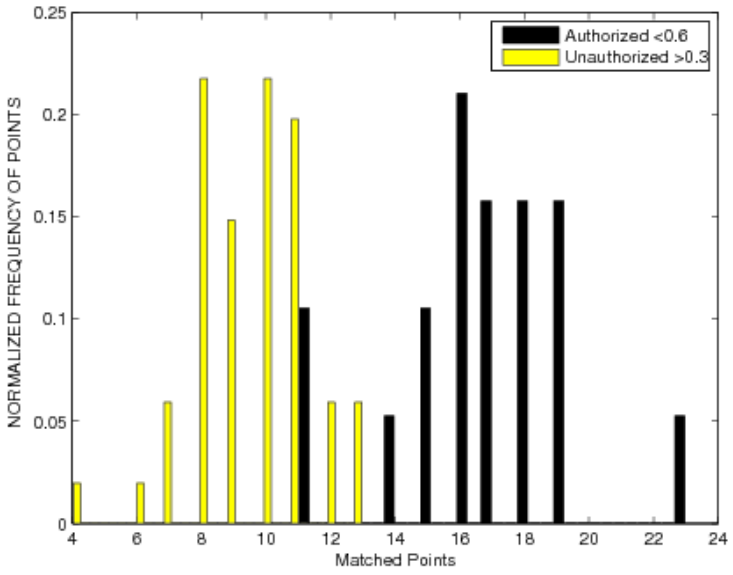


Figure 4.8: Histogram of matched points in the populations of attacks where the similarity value is higher than 0.3 and client accesses where the similarity value is lower than 0.6.

To correct this situation, the influence of the number of matched points and the pattern size have to be balanced. A correction parameter (γ) is introduced in the similarity measure to control this influence. The new metric is defined as:

$$S_\gamma = S \cdot C^{\gamma-1} = \frac{C^\gamma}{\sqrt{MN}} \tag{4.8}$$

where S, C, M and N are the same parameters from Equation 4.7. The γ correction parameter allows to improve the similarity values when a high number of matched points is obtained, specially in cases of patterns with a high number of points.

Using the gamma parameter, values can be higher than 1. In order to normalize the metric back into a $[0, 1]$ range, two proposals are presented.

First, in order to normalize the metric to the $[0, 1]$ space again, S_γ is divided by a reference value, R, representing a similarity value in the S_γ space which is certain to be an authorized access case. The new normalized metric will be defined as:

$$S_{\gamma R} = min\left\{\frac{S_\gamma}{R}, 1\right\} \tag{4.9}$$

R can be defined in the same space as S_γ in Equation 4.8 as $R = S_R C_R^{\gamma-1}$, where S_R and C_R are values in the similarity and matched point space, respectively. These values must have a very high probability to belong to a match between patterns from the same individual. Moreover, these parameters should not be very high in order to allow a good number of positive cases to get closer to a similarity value of 1. Ideally, mean values for the similarity and matched points distributions should be used.

In Figure 4.3(a) and 4.7(a), the distribution of the unauthorized and authorized cases was shown for the matched points and normalized metric, respectively. Mean values for the client accesses are, respectively, 18 points and 0.65. Distributions for the unauthorized accesses have a mean and standard deviation values of $\mu_m = 5.58, \sigma_m = 1.74$ for matched points and $\mu_s = 0.1508, \sigma_s = 0.0537$ for similarity values.

Given that $18 > \mu_m + 7\sigma_m$ and $0.65 > \mu_s + 9\sigma_s$, $S_R = 0.65$ and $C_R = 18$ (i.e. the mean values of matched points and normalized metric distributions for clients accesses) are safe values to be used as they are far enough from their respective attack distribution means.

To select a good γ parameter, the confidence band improvement has been evaluated for different values of γ (Fig 4.9). The maximum improvement is achieved at $\gamma = 1.12$ with a confidence band of 0.2304, twice the original from previous section. The distribution of the whole training set (with $\gamma = 1.12$, $S_R = 0.65$ and $C_R = 18$) is shown in Figure4.10(a) where the wide separation between classes can be observed. Figure 4.10(b) shows the associated FAR vs FRR curves.

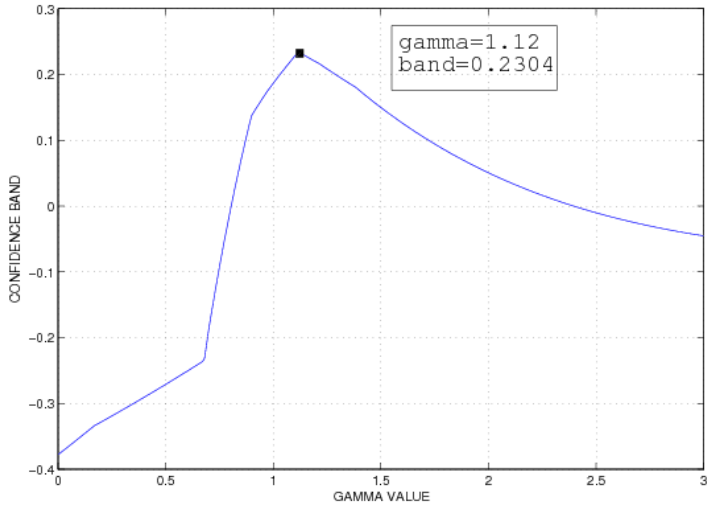
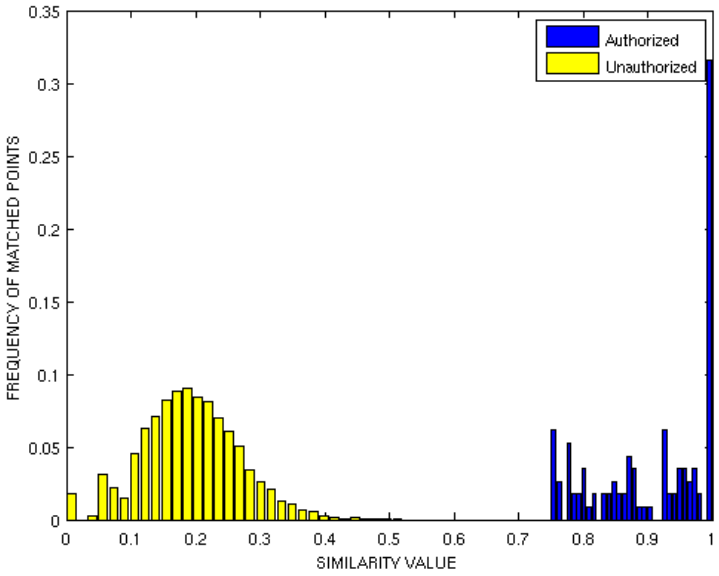


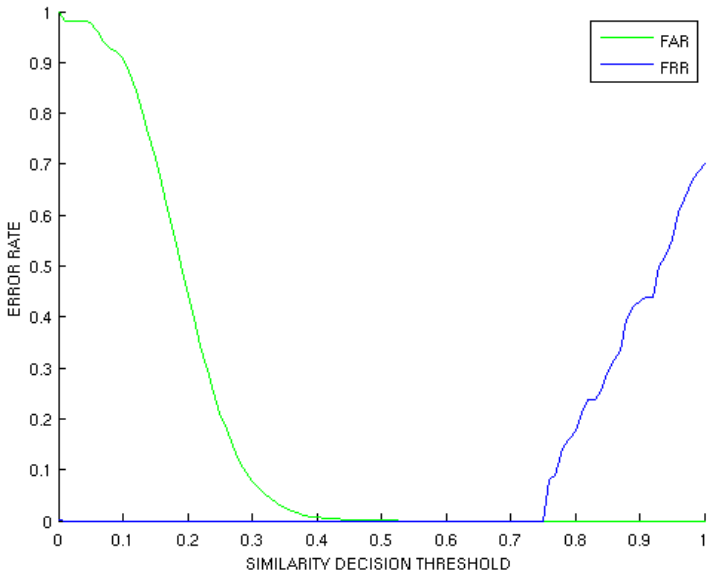
Figure 4.9: Confidence band size vs gamma (γ) parameter value using reference value to normalize. Maximum band is obtained at $\gamma = 1.12$.

The second normalization proposal is the use of a sigmoid transference function, $T(x)$:

$$T(x) = \frac{1}{1 + e^{s \cdot (x - 0.5)}} \tag{4.10}$$



(a)



(b)

Figure 4.10: (a) Distributions of similarity values using $S_{\gamma R}$ with γ=1.12. (b) FAR and FRR curves obtained with the same metric.

where s is a scale factor to adjust the function to the correct domain as S_γ does not return negatives or much higher than 1 values when a typical $\gamma \in [1, 2]$ is used. In this work, $s = 6$ was chosen empirically. The normalized gamma-corrected metric, $S_{\gamma T}(x)$, is defined by:

$$S_{\gamma T} = T(S_\gamma) \tag{4.11}$$

The maximum confidence band is obtained at $\gamma = 1.12$ as it was the case for the previous metric (Figure 4.11). However, the improvement is higher in this case as the confidence band reaches 0.3288. Figure 4.12(a) shows the wide separation between classes can be observed. Figure 4.12(b) shows the associated FAR vs FRR curves.

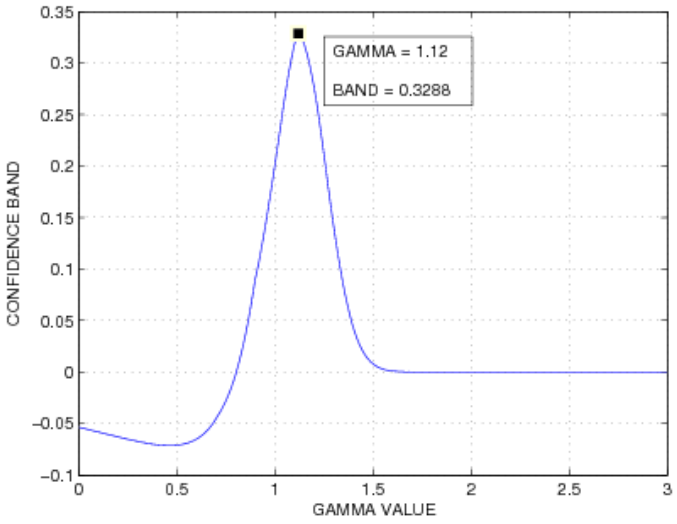
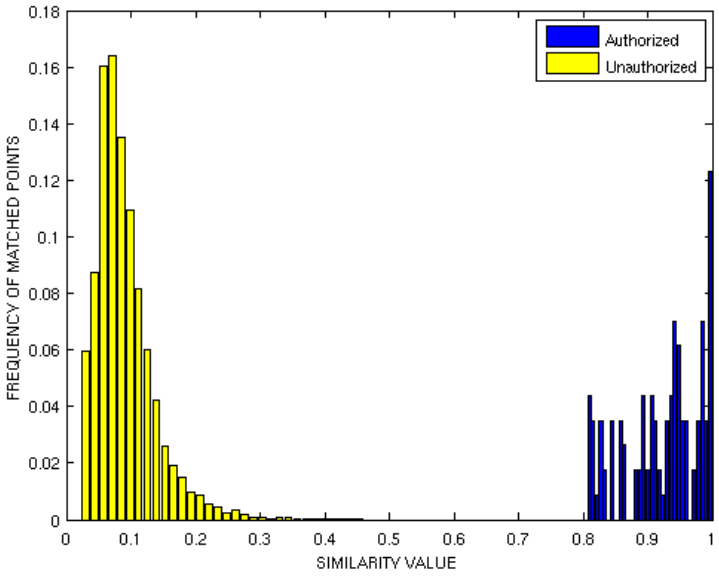
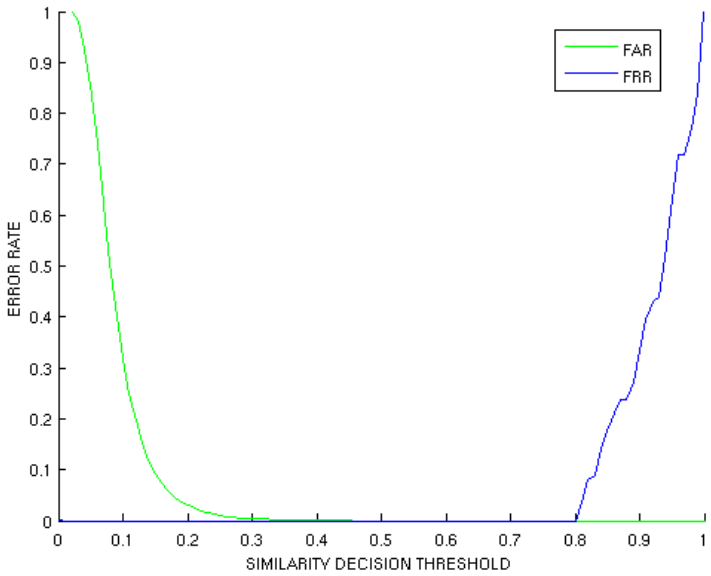


Figure 4.11: Confidence band size vs gamma (γ) parameter value using a normalization transfer function. Maximum band is obtained at $\gamma = 1.12$.

In conclusion, the normalized metric using a sigmoidal transference function with a matched points weighting parameter (γ) offers the best classification performance for this system.



(a)



(b)

Figure 4.12: (a) Distributions of similarity values using $S_{\gamma T}$ with γ=1.12. (b) FAR and FRR curves obtained with the same metric.

4.3. Retinal verification test

In this section, the metrics were tested using all the VARIA images to evaluate the global performance of the system once the parameters have been fixed. The test was performed without using the matching optimization derived from the characterization of points as this process is later evaluated by itself to check that it does not affect to the authentication capabilities. The False Acceptance Rate and False Rejection Rate were calculated for several metrics to test the metrics performance (the metrics normalized by Equation 4.5, Equation 4.7 and the gamma-corrected normalized metric defined in Equation 4.9. The results are shown in Fig 4.13.

The establishment of a wide confidence band is specially important in this scenario of different images from users acquired on different times and with different configurations of the capture hardware.

The characterization of feature points is included in the system to evaluate its impact and the new FAR and FRR curves are calculated (Figure 4.14). The impact is almost null due to the fact that this confidence band is limited to real life similarities between different retinas. The negative impact is avoided as the specificity is over 96%, minimizing the misclassification.

Finally, a test is run to evaluate the influence of the image quality. The images with a pattern size below a threshold are removed from the set and the confidence band is computed. Figure 4.15 shows the confidence band versus the minimum detected points constraint. The confidence band does not grow significantly until a fairly high threshold is set. Considering as threshold the mean value of detected points for all the test set, i.e. 25.2, the confidence band begins to grow by a margin greater than 0.1. So, it is necessary to remove almost half of the images to increase the band by more than 0.1, suggesting that the gamma-corrected metric is very robust to low quality images.

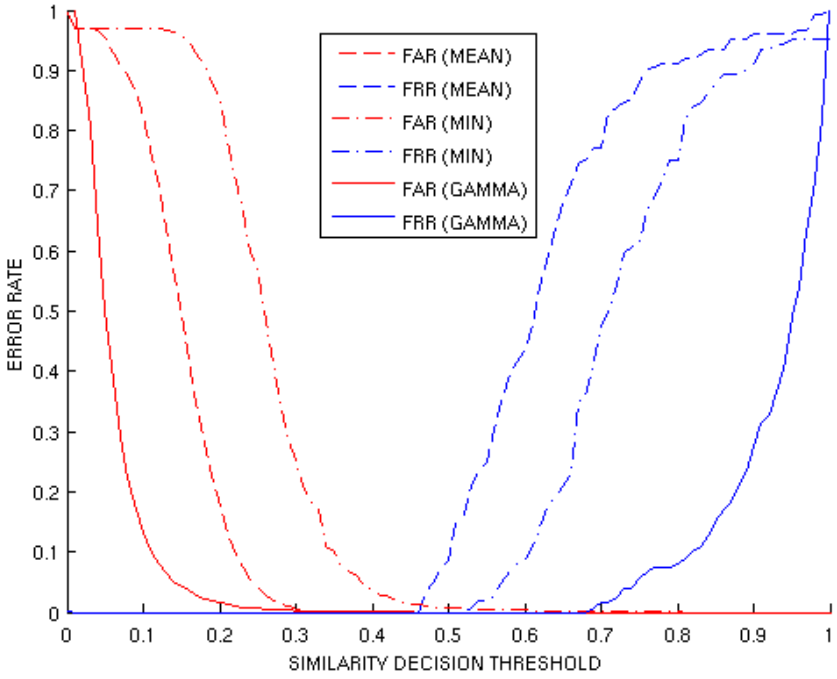


Figure 4.13: FAR and FRR curves for the normalized similarity metrics (min: normalized by minimum points, mean: normalized by geometrical mean and gamma: gamma corrected metric using a sigmoid function). The best confidence band belongs to the gamma corrected metric, i.e. a band of 0.2317.

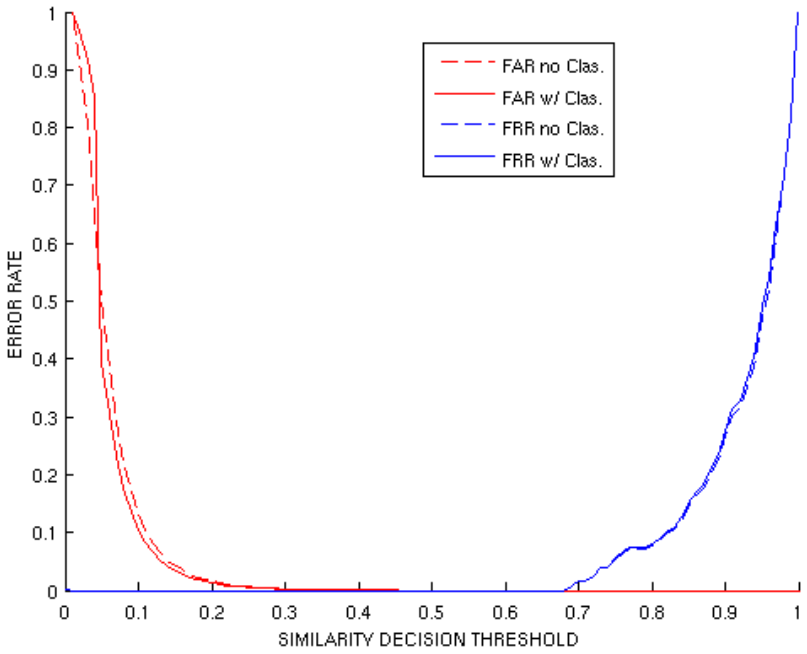


Figure 4.14: FAR and FRR curves for the authentication system using $S_{\gamma T}$ as decision metric, with and without classification of feature points.

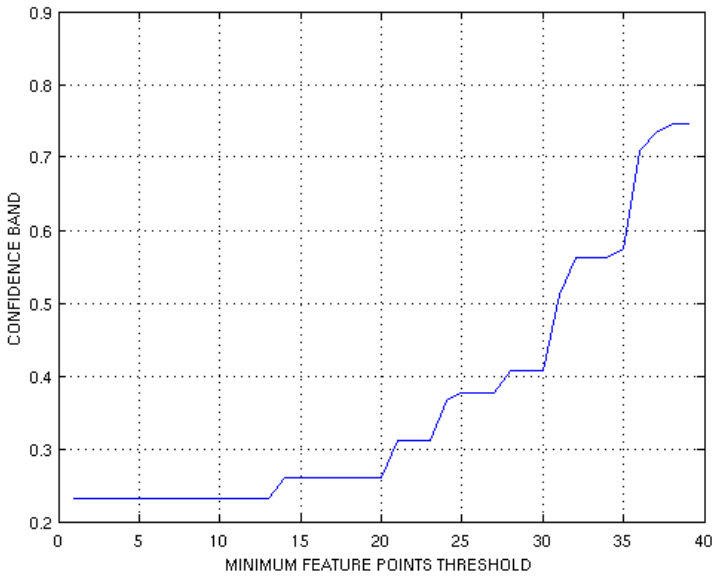


Figure 4.15: Evolution of the confidence band using a threshold of minimum detected points per pattern.

Chapter 5

Conclusions and future research

In this thesis an automatic authentication system based on retinal vessel tree is introduced. A biometric template is designed and extracted and a matching algorithm is developed to compare templates. Experimental results performed in this work show that the system is valid for authentication use even in high security environments, as it was the initial intention given the nature of human retina.

The suitability of the retinal vessel tree as biometric characteristic has been evaluated and tested based on earlier medical studies and making use of an initial authentication methodology that uses the whole vessel tree structure as biometric template. Once the suitability was proved, a more compact template was designed which reduces storage requirements and, at the same time, reduces the computation to perform a matching. Instead of using an image of the vessel tree structure (creases), feature points from it were extracted, i.e. bifurcations and crossovers, and used as new template. Thus, individuals can be reliably characterized by a reduced set of feature points from their retinas.

Two methodologies were introduced in order to extract feature points from the retinal vessel tree, one based on the previous crease extraction method-

ology employed for the initial authentication system and another based on a vessel segmentation approach. Segmentation approach proved to be a bit more precise in terms of point extraction (approximately. 5% more of precision and recall) but required far more computation time. Thus, the crease-based point extraction method is the preferred as it showed enough detection capabilities to offer a reliable set of points to characterize individuals in a very limited time.

A feature point matching algorithm was introduced taking into account the external alterations that acquired retinal images might suffer (geometric transformations, illumination conditions changes, etc.). To reduce the computation of the optimal transformation parameters, a methodology of characterization of feature points was introduced. This methodology used local (neighbors) and topological information (close feature points) of the feature points to determine their nature. The method got better performance results than previous approaches, and it can be tuned in terms of two radius of control. This is specially important in the biometrics domain, as a highest as possible specificity is desired to avoid provoking misclassification of points, leading to miss matchings between legitimate points or to match incorrect pairs of points. Overall, the improvement of this approach was nearly a 20% computation reduction for the matching stage while keeping the same confidence band range for authentication. This improvement is specially important in domains such as identification where templates are matched through several stored templates and, thus, the matching stage requires lowest possible computation.

Also, a similarity metric study was carried out in order to test the authentication capabilities of the system in a retinal image database. Several ideas over traditional metrics were proposed, specially the introduction of a parameter to tune the relevance of absolute number of matched points versus the normalization function.

Overall, the authentication system offered excellent results with a mean

time of 0.542s for the whole biometric template extraction stage since acquisition and 0.112s for pattern matching. These times were taken on a Pentium IV 2.4Ghz Desktop PC.

As future research goes, some improvements can be designed and tested for the system. One of the most interesting points, would be to advance in the extraction of higher-level data from the images. Information obtained via relevant structures such as optical disk or vessel higher-level data such as bifurcations/crossovers accurate angles, would ideally help to improve the performance of the system specially in terms of computation. At the time of this thesis, methods for such tasks were not reliable enough to be included on it as the errors in this stage cause matchings to be less accurate and, subsequently, similarity metrics less precise in terms of class separability.

Another interesting future line of work is the extension of the system proposed here to cope with identification tasks, defining an optimal indexing algorithm for templates in order to enhance the search through database of the most similar template.

Appendix A

Multilocal Level-set Extrinsic Curvature

Vessels are reliable landmarks in retinal images as they are almost rigid structures visible in every modality of acquisition. Moreover, they can be seen as ridges or valleys taking images as topographic surfaces as shown in Figure A.1.

Among the many definitions of *crease*, that based on level set extrinsic curvature ($LSEC$) has very useful invariance properties. Given a function $L : \mathbb{R}^d \rightarrow \mathbb{R}$, the set of levels for a constant l consists of the set of points $\{\mathbf{x}|L(\mathbf{x}) = l\}$ (Figure A.2). For $2D$ images L can be seen as a topographic surface, and the level sets as its level curvatures. The negative minimum of the curvature κ, level by level, conforms the so called *valley curves*, while positive maxima conform *ridge curves*. LSEC in $2D$ can be expressed in terms of L derivatives following methods of tensor calculus [81] as indicated in Equation A.1:

$$\kappa = (2L_xL_yL_{xy} - L_y^2L_{xx} - L_x^2L_{yy})(L_x^2 + L_y^2)^{-\frac{3}{2}} \tag{A.1}$$

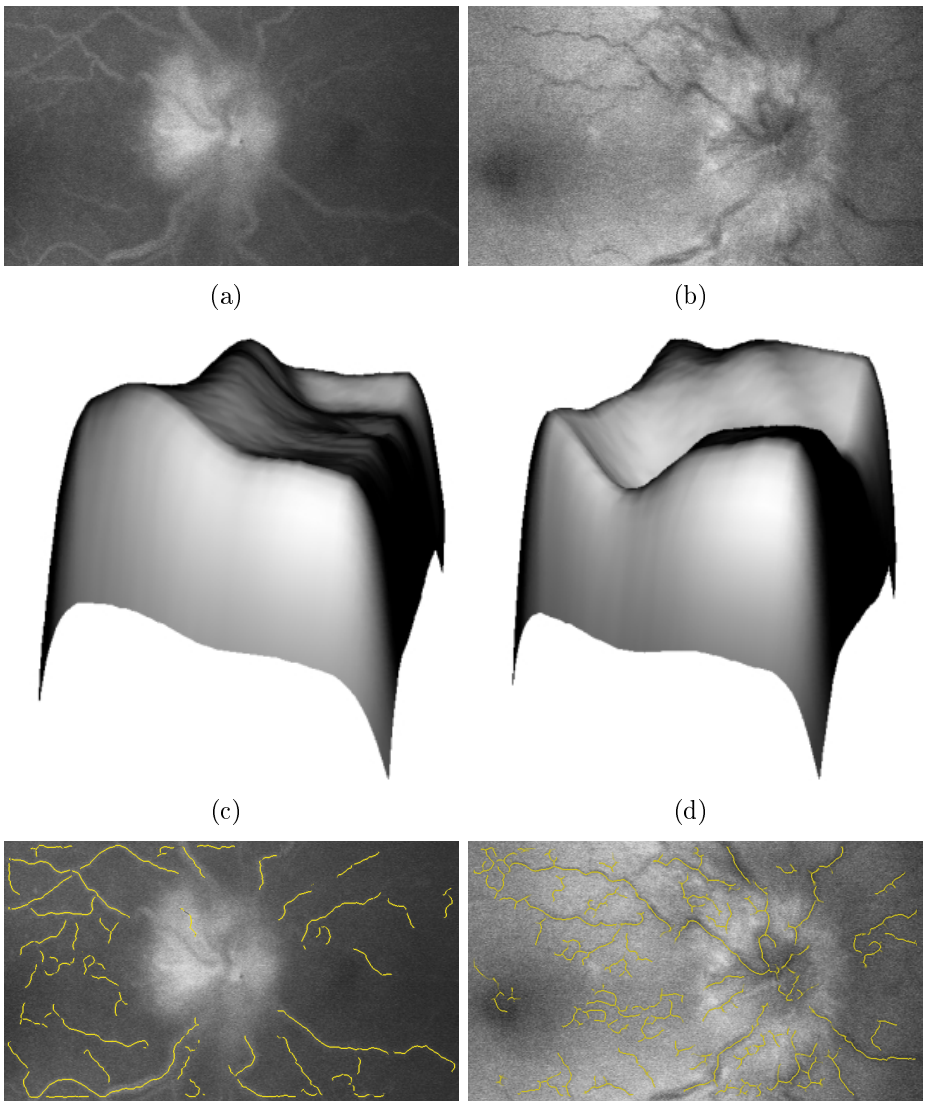


Figure A.1: (a) and (b) show different images. (c) and (d) show the interest region of images (a) and (b) represented as a topographic surface. It can be observed that, regardless of the intensity and contrast variation, ridges and valleys are clearly visible. Finally, (e) and (f) show the original images with the extracted creases (in yellow) over them.

where

$$L_\alpha = \frac{\partial L}{\partial \alpha}, L_{\alpha\beta} = \frac{\partial^2 L}{\partial \alpha \partial \beta}, \alpha, \beta \in \{x, y\}$$

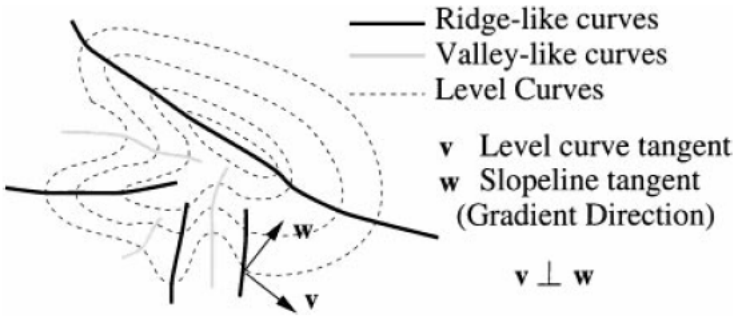


Figure A.2: Graphical interpretation of terms related to level curves: negative minima of curvature κ form *valley curves* while positive maxima form *ridge curves*. Vectors v and w represent tangent and gradient direction, respectively, in each point of level curves, being additionally orthogonal one respect to the other.

However, the usual discretization of the LSEC is ill-defined in some cases, introducing some unexpected discontinuities in the center of long objects. These problems are due to the too local definition of the LSEC, which is not very appropriate in the discrete domain.

To avoid discontinuity problems around critical points, a new operator is used; the *Multilocal Level Set Extrinsic Curvature - Structure Tensor*, $\tilde{\kappa}_{\mathrm{d}}$, based on the LSEC.

In $2D$ κ can be defined by means of its relation to slope lines, this is, the lines that form the vector field of the gradient $\mathbf{w}$ and, therefore, orthogonal to level curves. Due to this property, when level curves are parallel, slope lines are straight and parallel too, and when level curves bend, slope lines converge/diverge. In vector calculus there is a *divergence* operator which measures this parallelism degree. The divergence of a $d-$dimensional vector field $\mathbf{u} : \mathbb{R}^d \rightarrow \mathbb{R}^d, \mathbf{u}(\mathbf{x}) = (\mathbf{u}^1(\mathbf{x}), \ldots, \mathbf{u}^d(\mathbf{x}))^t$ is defined as [74]:

$$\mathrm{div}(\mathbf{u}) = \sum_{i=1}^{d} \frac{\partial u^i}{\partial x^i} \tag{A.2}$$

If we name $\mathbf{0}_d$ as the $d-$dimensional zero vector, the normalized vector field of the gradient of $L : \mathbb{R}^d \rightarrow \mathbb{R}$, $\bar{\mathbf{w}}$ is defined as:

$$\bar{\mathbf{w}} = \begin{cases} \mathbf{w}/||\mathbf{w}|| & if||\mathbf{w}|| > 0 \\ \mathbf{0}_d & if||\mathbf{w}|| = 0 \end{cases} \tag{A.3}$$

and then it can be proved:

$$\kappa_d = -\mathrm{div}(\bar{\mathbf{w}}) \tag{A.4}$$

In $2D$, let $\mathbf{x}$ be the point when we want to compute divergence of a $2D$ vector field. Let $\mathcal{C}$ be a closed curve simple in $\mathbb{R}^2$ with parameter l, surrounding point $\mathbf{x}$. Let $\mathbf{n}$ be its unitary normal vector and ω the region inside $\mathcal{C}$ (FigureA.3).

Then, the divergence of $\mathbf{u}$ in $\mathbf{x}$ can also be expressed as [74]:

$$\mathrm{div}(\mathbf{u}) = \lim_{\omega \rightarrow 0} \frac{1}{\omega} \int_{\mathcal{C}} \mathbf{u}^t \cdot \mathbf{n} \partial l \tag{A.5}$$

The local definition of κ_d is replaced by a multilocal definition based on the discretized version of EquationA.5, where multilocal property is obtained assuming that the neighborhood $\mathcal{W}$ around a point $\mathbf{x}$, or analogously, its closed limit $\mathcal{C}$, is a selectable parameter. This means that gradient vectors are taken into account along the path $\mathcal{C}$ in order to compute $\mathrm{div}(\bar{\mathbf{w}})$ in $\mathbf{x}$.

Following this reasoning, for a given dimension d, we denote by $\overline{\kappa}_d$ the MLSEC operator (*Multilocal Level-Set Extrinsic Curvature*) based in equations A.3, A.4 y A.5, given a particular $\mathcal{C}$. Operator MLSEC for a discrete domain

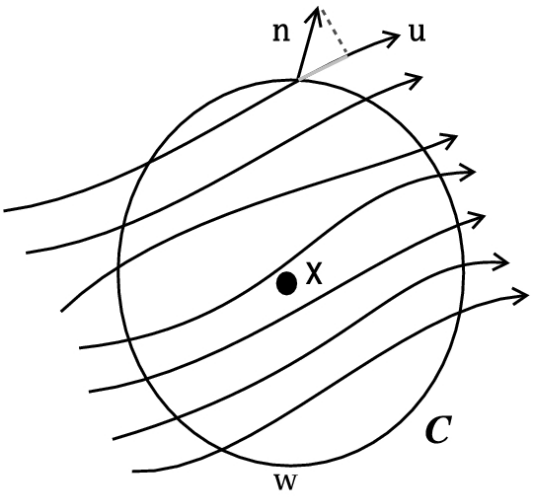


Figure A.3: Implicated geometry in the definition of the divergence for a vector field $\mathbf{u}$ in a point $\mathbf{x}$. $\mathcal{C}$ is a simple closed curve $\mathbb{R}^2$ parameterized by l surrounding point $\mathbf{x}$, with unitary normal vector $\mathbf{n}$. ω represents the area inside $\mathcal{C}$.

is defined as:

$$\overline{\kappa}_d = -\text{div}(\bar{\mathbf{w}}) = -\frac{d}{r}\sum_{k=1}^{r} \bar{\mathbf{w}}_k^t \cdot \mathbf{n}_k \tag{A.6}$$

where r is the selected adjacency given by the specific $\mathcal{C}$ (for instance, in $2D$ we can choose between 4 adjacency or 8 adjacency), and d is the space dimensionality.

Once $\overline{\kappa}_d$ has been established as a good measure for ridges and valleys, it can be still improved by prefiltering the gradient vector field of the image to increase the attraction/repulsion in the ridge curves and valley curves, measured by κ. This can be performed using an structural tensor, which is a good known technique for analysis of oriented textures [5, 38, 39].

In the $d-$dimensional space, given a symmetric neighborhood of size σ_{I} centered in a given point $\mathbf{x}$, denoted by $\mathcal{N}(\mathbf{x}; \sigma_{\text{I}})$, the structural tensor is defined as a semi defined positive symmetric matrix of size $d \times d$:

$$\mathbf{M}(\mathbf{x};\sigma_I) = \mathcal{N}(\mathbf{x};\sigma_{\mathrm{I}}) * (\mathbf{w}(\mathbf{x}) \cdot \mathbf{w}^t(\mathbf{x})) \tag{A.7}$$

where "$*$" denotes convolution operator.

The eigenvector corresponding to the minimum eigenvalue of $\mathbf{M}(\mathbf{x};\sigma_I)$, $\mathbf{w}'(\mathbf{x};\sigma_{\mathrm{I}})$, expresses the *predominant orientation* in $\mathbf{x}$, and is perpendicular to the *gradient predominant orientation*. Note that the gradient of a function points to the maximum variation direction, and the predominant orientation is perpendicular to this direction because the anisotropy appears as similar grey values along an orientation and great variations perpendicular to it.

This analysis assumes that in every neighborhood there is only one orientation. To verify this supposition, a normalized *confidence measure* is introduced. Each orientation is associated to a real value $\mathsf{C} \in [0,1]$ that can be obtained from the eigenvalues of the structural tensor. The eigenvalue similarity of the structural tensor implies isotropy and, as a result, C should be close to 0. Therefore, denoting $\lambda_1, \ldots, \lambda_d$ the eigenvalues of $\mathbf{M}$, a logical option is to check if the sum of their quadratic differences (Equation A.8) is greater than a established characteristic threshold c for $\lambda_\triangle$ in the structure we intend to enhance. An adequate function is [67]:

$$\lambda_\triangle(\mathbf{x};\sigma_{\mathrm{I}}) = \sum_{i=1}^{d} \sum_{j=i+1}^{d} (\lambda_i(\mathbf{x};\sigma_{\mathrm{I}}) - \lambda_j(\mathbf{x};\sigma_{\mathrm{I}}))^2 \tag{A.8}$$

$$\mathsf{C}(\mathbf{x};\sigma_{\mathrm{I}};c) = 1 - e^{-\lambda_\triangle(\mathbf{x};\sigma_{\mathrm{I}}))^2/2c^2} \tag{A.9}$$

As a conclusion, to compute ridge and valley lines from an image the following steps are performed:

1. Computation of the gradient vector field $\mathbf{w}$ and the structural tensor field $\mathbf{M}$. A neighborhood following a Gaussian distribution is used to

compute $\mathbf{M}$:

$$\mathbf{M}(\mathbf{x};\sigma_{\mathrm{I}}) = \begin{pmatrix} s_{11}(\mathbf{x};\sigma_{\mathrm{I}};\sigma_{\mathrm{D}}) & s_{12}(\mathbf{x}\ \sigma_{\mathrm{I}};\sigma_{\mathrm{D}}) \\ s_{12}(\mathbf{x};\sigma_{\mathrm{I}};\sigma_{\mathrm{D}}) & s_{22}(\mathbf{x};\sigma_{\mathrm{I}};\sigma_{\mathrm{D}}) \end{pmatrix} \tag{A.10}$$

$$\begin{aligned} s_{11}(\mathbf{x};\sigma_{\mathrm{I}};\sigma_{\mathrm{D}}) &= G(\mathbf{x};\sigma_{\mathrm{I}}) * (L_x(\mathbf{x};\sigma_{\mathrm{D}})L_x(\mathbf{x};\sigma_{\mathrm{D}})) \\ s_{12}(\mathbf{x};\sigma_{\mathrm{I}};\sigma_{\mathrm{D}}) &= G(\mathbf{x};\sigma_{\mathrm{I}}) * (L_x(\mathbf{x};\sigma_{\mathrm{D}})L_y(\mathbf{x};\sigma_{\mathrm{D}})) \\ s_{22}(\mathbf{x};\sigma_{\mathrm{I}};\sigma_{\mathrm{D}}) &= G(\mathbf{x};\sigma_{\mathrm{I}}) * (L_y(\mathbf{x};\sigma_{\mathrm{D}})L_y(\mathbf{x};\sigma_{\mathrm{D}})) \end{aligned}$$

The new parameter σ_D denotes the standard deviation of the Gaussian kernel for the differentiation process, needed to compute $\mathbf{w}$. The parameter σ_D is called *differentiation scale*, opposite to σ_I, called *integration scale*. The differentiation scale adjusts to the structure size whose orientations are looked for, while the integration scale is set to the neighborhood size in which the orientation is dominant.

2. Analysis of the eigenvalues of $\mathbf{M}$. The normalized eigenvector $\mathbf{w}'$ corresponding to the greater eigenvalue gives the predominant orientation in the gradient. In the analysis with structural tensor, opposite directions are equally dealt with. Therefore, to recover direction, $\mathbf{w}'$ is set in the same quadrant $2D$ as $\mathbf{w}$. In this way, we obtain a new vector field $\tilde{\mathbf{w}}$:

$$\tilde{\mathbf{w}} = \mathrm{sign}(\mathbf{w'}^{\mathrm{t}} \cdot \mathbf{w})\mathbf{w}' \tag{A.11}$$

where

$$sign(x) = \begin{cases} +1 & if\ x > 0 \\ -1 & if\ x < 0 \\ 0 & if\ x = 0 \end{cases} \tag{A.12}$$

Therefore, attraction/repulsion vectors are enforced.

3. Computation of the new measure of ridges or valleys using the MLSEC-

ST operator defined as:

$$\tilde{\kappa}_d = -\mathrm{div}(\tilde{\mathbf{w}}) \tag{A.13}$$

4. Computation of the confidence measure C allowing to discard ridges and valleys appearing in isotropic zones. Thus, $\tilde{\kappa}_d\mathsf{C}$ has a lesser response than $\tilde{\kappa}_d$ in isotropic regions and it will be the final measure of *creaseness*.

Figure A.4 illustrates the process to obtain creases and valleys in two different images, belonging to two different individuals.

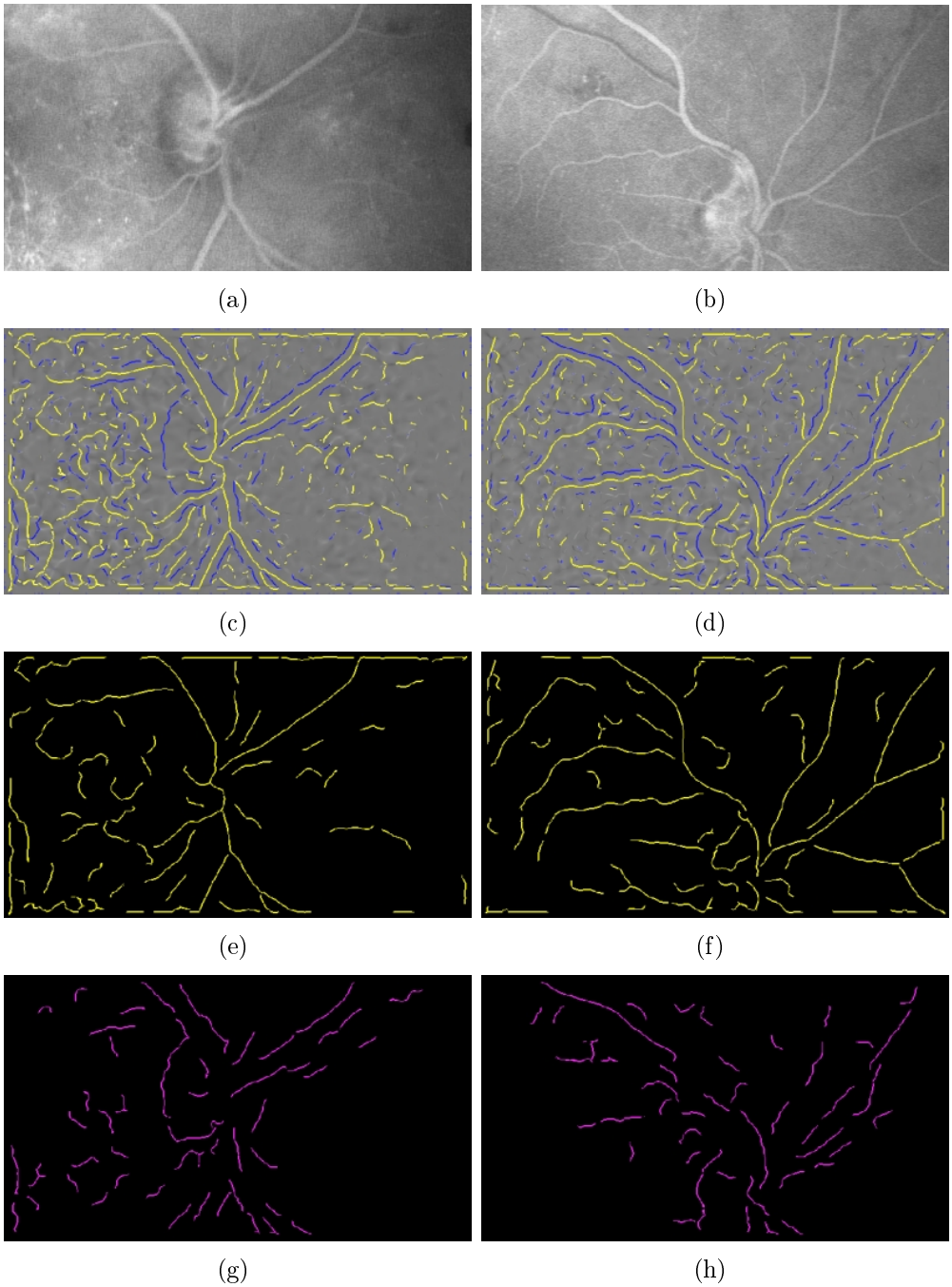


Figure A.4: Creases extraction process to two different images. (a) and (b) show original images. In (c) and (d) it is showed the obtained result using LSEC operator. Then, obtained ridges ((e), (f)) and valleys ((g) y (h)) for MLSEC-St operator are shown.

Appendix B

Stentiford thinning method

Thinning algorithms based on templates, like the Stentiford method, are widely used due to its adaptability and effectiveness. These algorithms use templates to delete pixels in the original image if a coincidence with the template takes place. They run iteratively, eroding the most external layers of objects in the processed image until no more pixels can be removed. To properly define Stentiford algorithm, it is necessary to introduce two prior definitions:

- *Final Pixel.* A pixel will be final if it is only connected to another pixel, i.e. in binary mode, a black pixel only has another black pixel in its 8-neighborhood.

- *Connectivity.* It is a measure of how many objects are connected to a particular pixel. Connectivity number (C_n) is defined as:

$$C_n = \sum_{k \in S} N_k - (N_k * N_{k+1} * N_{k+2}) \tag{B.1}$$

  where N_k is the pixel value of the k neighbor of a given pixel, $S = \{1, 3, 5, 7\}$.

Neighbors are taken in a counterclockwise fashion.

Figure B.1 illustrates the connectivity number. Figure B.1 (a) represents a connectivity value of 0 as the central pixel is not connected to any object. Case (b) represents a connectivity value of 1 as the central pixel can be removed without altering connectivity between left and right part. In case (c), the connectivity is 2 because as we remove the central pixel, both side parts will be separated. Analogously, cases (c) and (d) represent connectivity values of 3 and 4 respectively indicating the different parts in the neighborhood that will remain unconnected by removing central pixel.

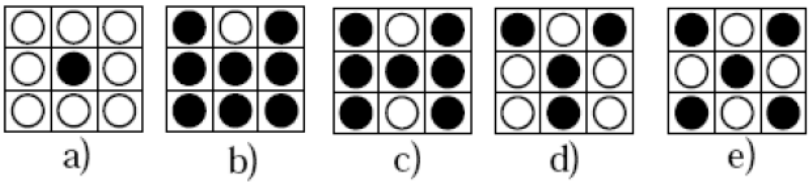


Figure B.1: Examples of different connectivities. a) $C_n = 0$ b) $C_n = 1$ c) $C_n = 2$ d) $C_n = 3$ e) $C_n = 4$.

The algorithm uses a set of four templates of 3×3 to track the image. Figure B.2 shows these arrays where black points represent object pixels, white points represent background pixels and empty squares are locations where it is not needed to check the image value to perform the operation.

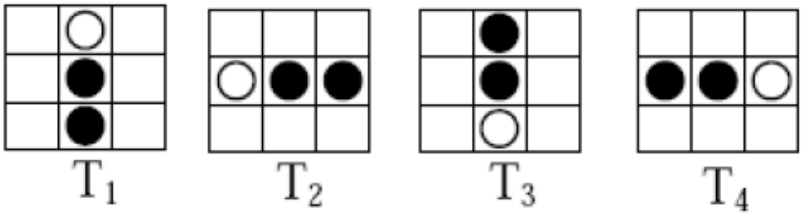


Figure B.2: The four templates used to perform the thinning on the segmented image.

Each template is aimed to test a particular subset of pixels in the image:

- T1 template removes pixels from the upper edge of objects.
- T2 template removes pixels from the left side of objects.
- T3 template removes pixels from the lower edge of objects.
- T4 template removes pixels from the right side of objects.

The thinning of the segmented structure is performed by tracking the resulting image of the previous step with each of the four templates in FigureB.2.

1. For each pixel in the image, its connectivity number, C_n, is computed following Equation B.1. Only pixels with a connectivity value equal to one will be processed. Therefore, only pixels that their removal will not alter connections between objects connected to it will be processed.
2. Each pixel with $C_n = 1$ is evaluated to check if it is an endpoint, this is, if only a black pixel is present in its neighborhood. This constraint is important to avoid erosion of objects in their direction.
3. For the pixels obtained from previous stages ($C_n = 1$ and not being end points), each of the four templates is compared centered in it. If the template matches with the analyzed window, the pixel is removed from the image. Otherwise, there will not be any modification.

All of these steps are repeated for each template, alternating their application and ending the iterative algorithm when none of the templates could match any pixel, i.e. no pixels were removed in the previous iteration.

Appendix C

Selected publications

M. Ortega, C. Marino, M.G. Penedo, M. Blanco, F. González, "Biometric Authentication using digital retinal images", *International Conference on Applied Computer Science*, pp. 422-427, China, April 2006.

M. Ortega, C. Marino, M.G. Penedo, M. Blanco, F. González, "Personal Authentication Based on Feature Extraction and Optica Nerve Location in Digital Retinal Images", *WSEAS Transactions on Computers*, 5(6), pp. 1169-1176, 2006.

M. Ortega, M.G. Penedo, L. Espona, F. González, "Biometric Authentication Based on Feature Extraction in Digital Retinal Images.", *1st Spanish Workshop on Biometrics (SWB 07)*, Girona (Spain), CD proceedings, June 2007.

L. Espona, M.J. Carreira, M. Ortega, M.G. Penedo, "A Snake for Retinal Vessel Segmentation", *Lecture Notes in Computer Science: Pattern Recognition and Image Analysis*, 4478, pp. 178-185, 2007.

M. Ortega, M.G. Penedo, C. Marino, M.J. Carreira, "Similarity Metrics Analysis for Feature Point Based Retinal Authentication", *Lecture Notes in Computer Science: Image Analysis and Recognition*, 5112, pp. 1023-1032, 2008.

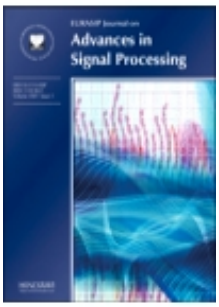
L. Espona, M.J. Carreira, M.G. Penedo, M. Ortega, "Comparison of Pixel and Subpixel Retinal Vessel Tree Segmentation using a Deformable Contour Model", *Lecture Notes in Computer Science: Pattern Recognition*, 5197, pp. 683-690, 2008.

L. Espona, M.J. Carreira, M.G. Penedo, M. Ortega, "Retinal Vessel Tree Segmentation using a Deformable Contour Model", *19th International Conference on Pattern Recognition (ICPR 2008)*, DOI 10.1109/ICPR.2008.4761762 , 4 pp., Tampa Bay, USA, December 2008.

C. Alonso-Montes, M. Ortega, M.G. Penedo, D.L. Vilarino, "Pixel Parallel Vessel Tree Extraction for a Personal Authentication System", *IEEE International Symposium on Circuits and Systems ISCAS*, pp. 1596-1599, Seattle (USA), May 2008.

M. Ortega, M.G. Penedo, C. Marino, M.J. Carreira, "A Novel Similarity Metric for Retinal Images Based Authentication", *International Conference on Bio-inspired Systems and Signal Processing (Biosignals 2009)*, pp. 249-253, Porto, Portugal, January 2009.

M. Ortega, J. Rouco, J. Novo, M.G. Penedo, "Vascular landmark detection in retinal images", *Eurocast 2009*, pp. 71-72, Las Palmas, Spain, February 2009.



M. Ortega, M.G. Penedo, J. Rouco, N. Barreira, M.J. Carreira, "Retinal verification using a feature points based biometric pattern", *EURASIP Journal on Advances in Signal Processing. Special Issue on Recent Advances in Biometrics Systems*, Article ID 235746, 13 pp., DOI 10.1155/2009/235746, 2009.

M. Ortega, M.G. Penedo, J. Rouco, N. Barreira, M.J. Carreira, "Personal verification based on extraction and characterization of retinal feature points",*Journal of Visual Languages and Computing*, 20 (2), 80-90, 2009.

D. Calvo, M. Ortega, M.G. Penedo, J. Rouco, B. Remeseiro, "Characterisation of retinal feature points applied to a biometric system", International Conference on Image Analysis and Processing (ICIAP), Salerno, Italy, September 2009. (To Appear)

Bibliography

[1] S. Arivazhagan, T. ArulFlora, and L. Ganesan. Fingerprint verification using gabor co-occurrence features. In *International Conference on Computational Intelligence and Multimedia Applications*, pages 281–285, 2007.

[2] C. Beavan. *Fingerprints : The Origins of Crime Detection and the Murder Case That Launched Forensic Science*. Hyperion Press, 2001.

[3] V. Bevilacqua, S. Cambó, L. Cariello, and G. Mastronardi. A combined method to detect retinal fundus features. *Proceedings of IEEE European Conference on Emergent Aspects in Clinical Data Analysis*, 2005.

[4] W. Bicz, Z. Gurnienny, and M. Pluta. Ultrasound sensors for fingerprints recognition. *Proc. of SPIE, Vol. 2634, Optoelectric and electronic sensors*, pages 104–111, 1995.

[5] J. Bigun, G. Granlund, and J. Wiklund. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:775–790, 1991.

[6] H. Blum. A Transformation for Extracting New Descriptors of Shape. In W. Wathen-Dunn, editor, *Models for the Perception of Speech and Visual Form*, pages 362–380. MIT Press, Cambridge, 1967.

[7] R. Bolle, J. Connell, S. Pankanti, N. Ratha, and A. Senior. *Guide to Biometrics*. SpringerVerlag, 2003.

[8] R. M. Bolle, J. H. Connell, and N. K. Ratha. Biometric perils and patches. *Pattern Recognition*, 35(12):2727–2738, 2002.

[9] R. M. Bolle, A. W. Senior, N. K. Ratha, and S. Pankanti. Fingerprint minutiae: A constructive definition. In *Biometric Authentication*, pages 58–66, 2002.

[10] R. Brunelli and T. Poggio. Face recognition: Features versus templates. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(10):1042–1052, 1993.

[11] M. Burge and W. Burger. Ear biometrics. In *BIOMETRICS:Personal Identification in Networked Society*. Kluwert Academic Publishers, 1999.

[12] M. Burge and W. Burger. Ear biometrics in computer vision. In *ICPR*, pages 2822–2826, 2000.

[13] I. G. Caderno, M. G. Penedo, C. Mariño, M. J. Carreira, F. Gómez-Ulla, and F. González. Automatic extraction of the retina av index. In *ICIAR (2)*, pages 132–140, 2004.

[14] A. Cerdá Barberá. *La huella digital: expresión genética de la tipología psicológica*. Ediciones ACB, 2000.

[15] S. Chen, B. Ma, and K. Zhang. The normalized similarity metric and its applications. In *BIBM '07: Proceedings of the 2007 IEEE International Conference on Bioinformatics and Biomedicine*, pages 172–180. IEEE Computer Society, 2007.

[16] C.-T. Chou, S.-W. Shih, and D.-Y. Chen. Design of gabor filter banks for iris recognition. In *IIH-MSP*, pages 403–406, 2006.

[17] R. Clarke. Human identification in information systems: management challenges and public policy issues. *Information Technology and People*, 7(4):6–37, 1994.

[18] C.Mariño, M.G.Penedo, M.J.Carreira, and F.Gonzalez. Retinal angiography based authentication. *Lecture Notes in Computer Science*, 2905:306–313, 2003.

[19] A. P. Condurache and T. Aach. Vessel segmentation in angiograms using hysteresis thresholding. *Proceedings of the Ninth IAPR conference on Machine Vision Applications 2005*, 1:269–272, 2005.

[20] R. Curbelo. Noisy fingerprint identification by artificial neural networks. In *9th Workshop on virtual intelligence / dynamic neural networks*, pages 432–449, 1999.

[21] DARPA. Human ID at a distance (humanID). http://infowar.net/tia/www.darpa.mil/iao/HID.htm.

[22] J. Daugman. Biometric personal identification system based on iris analysis. United States Patent No.5,291.560, 1994.

[23] L. De Schaepdrijver, L. Simoens, H. Lauwers, and J. DeGesst. Retinal vascular patterns in domestic animals. *Res. Vet. Sci.*, 47:34–42, 1989.

[24] P. den Elsen, J. Maintz, D. E.J.D.Pol, and M.A.Viergever. Automatic registration of CT and MR brain images using correlation of geometrical features. *IEEE Trans. on Medical Imaging*, 14(2):384–396, 1995.

[25] S. Ehrenfeld and S. Littauer. *Introduction to Statistical Method.* McGraw-Hill, 1964.

[26] V. Espinosa-Duro. Fingerprints thinning algorithm. *IEEE AES Transactions on Aerospace and Electronics Systems*, 18:28–30, 2003.

[27] F. Family, B. Masters, and D. Platt. Fractal pattern formation in human retinal vessels. *Physica D*, 38:98–103, 1989.

[28] H. Farzin, H. Abrishami-Moghaddam, and M.-S. Moin. A novel retinal identification system. *EURASIP Journal on Advances in Signal Processing*, ID 280635:10 pp., 2008.

[29] Fingerprint Verification Competition. FVC 2006. http://bias.csr.unibo.it/fvc2006/results.asp.

[30] V. Fleury and L. Schwartz. Difussion limited aggregation from shear stress as a simple model of vasculogenesis. *Fractals*, 7:33–39, 1999.

[31] V. Fleury and L. Schwartz. Modelisation of 3-d microvasculature by interlaced diffusion limited aggregation. *Fractals*, 8:255–259, 2000.

[32] J. Foley, A. van Dam, S. Feiner, and J. Hughes. *Computer Graphics: Principles and Practice, second edition.* Addison-Wesley Professional, 1990.

[33] D. Follette, E. Hultmark, and J. Jordan. Direct optical input system for fingerprint verification. *IBM Technical Disclosure Bulletin, (74C 00989)*, 1974.

[34] E. Grisan, A. Pesce, A. Giani, M. Foracchia, and A. Ruggeri. A new tracking system for the robust extraction of retinal vessel structure. *Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE*, 1:1620–1623, 2004.

[35] Z. He, Z. Sun, T. Tan, X. Qiu, C. Zhong, and W. Dong. Boosting ordinal features for accurate and fast iris recognition. In *CVPR*, 2008.

[36] R. Hill. Retina identification. In A. Jain, R. Bolle, and S. Pankanti, editors, *Biometrics: Personal Identification in Networked Society*, pages 123–142. Kluwer Academic Press, Boston, 1999.

[37] L. Hong, Y. Wan, and A. K. Jain. Fingerprint image enhancement: Algorithm and performance evaluation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(8):777–789, 1998.

[38] B. Jähne. Spatio-temporal image processing. *Lecture Notes in Computer Science*, 751:153–152, 1993.

[39] B. Jähne. *Digital Image Processing*, chapter 13. Springer, 2002.

[40] A. K. Jain, R. Bolle, and S. Pankanti, editors. *Biometrics, Personal Identification in Networked Society: Personal Identification in Networked Society*. Kluwer Academic Publishers, Norwell, MA, USA, 1998.

[41] A. K. Jain, P. Flynn, and A. A. Ross. *Handbook of Biometrics*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007.

[42] A. K. Jain, A. Ross, and S. Pankanti. A prototype hand geometry-based verification system. In *AVBPA*, pages 166–171, 1999.

[43] J.P.Lewis. Fast template matching. *Vision Interface*, pages 120–123, 1995.

[44] H.-C. Kim, D. Kim, S. Y. Bang, and S.-Y. Lee. Face recognition using the second-order mixture-of-eigenfaces method. *Pattern Recognition*, 37(2):337–349, 2004.

[45] J. Kim, S. Cho, J. Choi, and R. J. Marks. Iris recognition using wavelet features. *VLSI Signal Processing*, 38(2):147–156, 2004.

[46] D. R. Kisku, A. Rattani, M. Tistarelli, and P. Gupta. Graph application on face for personal authentication and recognition. In *ICARCV*, pages 1150–1155, 2008.

[47] G. Landini, G. Misson, and P. Murray. Fractal analysis of the normal human retinalfluorescein angiogram. *Curr. Eye Res.*, 12:23–27, 1993.

[48] L.G.Brown. A survey of image registration techniques. *ACM Computer Surveys*, 24(4):325–376, 1992.

[49] D. Lloret, A. López, J. Serrat, and J. Villanueva. Creaseness-based CT and MR registration: comparison with the mutual information method. *Journal of Electronic Imaging*, 8(3):255–262, July 1999.

[50] D. Lloret, C. Mariño, J. Serrat, A.M.López, and J. Villanueva. Landmark-based registration of full SLO video sequences. In *Proceedings of the IX Spanish Symposium on Pattern Recognition and Image Analysis*, volume I, pages 189–194, 2001.

[51] A. López, D. Lloret, J. Serrat, and J. Villanueva. Multilocal creaseness based on the level-set extrinsic curvature. *Computer Vision and Image Understanding*, 77(1):111–144, 2000.

[52] A. López, F. Lumbreras, J. Serrat, and J. Villanueva. Evaluation of methods for ridge and valley detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(4):327–335, 1999.

[53] A. M. López, F. Lumbreras, and J. Serrat. Creaseness from level set extrinsic curvature. In *ECCV*, pages 156–169, 1998.

[54] L. Ma, Y. Wang, and T. Tan. Iris recognition using circular symmetric filters. In *ICPR (2)*, pages 414–417, 2002.

[55] J. Mainguet, M. Pegulu, and J. Harris. Fingerchip^{TM}: Thermal imaging and finger sweeping in a silicon fingerprint sensor. In *Proc. of AutoID 99*, pages 91–94, 1999.

[56] M. Mainster. The fractal properties of retinal vessels: embryological and clinical implications. *Eye*, 4:235–241, 1990.

[57] D. Maio and D. Maltoni. Direct gray-scale minutiae detection in fingerprints. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(1):27–40, 1997.

[58] C. Mariño, M. G. Penedo, M. Penas, M. J. Carreira, and F. Gonzalez. Personal authentication using digital retinal images. *Pattern Analysis and Applications*, 9:21–33, 2006.

[59] B. Masters. Fractal analysis of the vascular tree in the human retina. *Annu. Rev. Biomed. Eng.*, 6:427–452, 2004.

[60] B. R. Masters and D. Platt. Development of human retinal vessels: a fractal analysis. *Invest. Ophthalmol. Vis. Sci.*, 30:391, 1989.

[61] K. Messer, J. Kittler, M. Sadeghi, S. Marcel, C. Marcel, S. Bengio, F. Cardinaux, C. Sanderson, J. Czyz, L. Vandendorpe, S. Srisuk, M. Petrou, W. Kurutach, A. Kadyrov, R. Paredes, B. Kepenekci, F. B. Tek, G. B. Akar, F. Deravi, and N. Mavity. Face verification competition on the xm2vts database. In *AVBPA*, pages 964–974, 2003.

[62] A. S. Mian, M. Bennamoun, and R. A. Owens. Keypoint detection and local feature matching for textured 3d face recognition. *International Journal of Computer Vision*, 79(1):1–12, 2008.

[63] B. Miller. Vital signs of identity. *IEEE Spectr.*, 31(2):22–30, 1994.

[64] B. Moghaddam and A. Pentland. Probabilistic visual learning for object representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(7):696–710, 1997.

[65] M.S.Markov, H.G.Rylander, and A.J.Welch. Real-time algorithm for retinal tracking. *IEEE Trans. on Biomedical Engineering*, 40(12):1269–1281, 1993.

[66] M. Nabti and A. Bouridane. An improved iris recognition system using feature extraction based on wavelet maxima moment invariants. In *ICB*, pages 988–996, 2007.

[67] W. Niessen, A. López, W. Van Enk, P. Van Roermund, B. ter Haar Romeny, and M. Viergever. In vivo analysis of trabecular bone architecture. *Lecture Notes in Computer Science*, 1230:435–440, 1997.

[68] D. Noden. Embryonic origins and assembly of blood vessels. *Am. Rev. Respir. Dis.*, 140:1097–1103, 1989.

[69] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9(1):62–66, 1979.

[70] P. Patil, S. Suralkar, and H. Abhyankar. Fingerprint verification based on fixed length square finger code. In *17th IEEE International Conference on Tools with Artificial Intelligence, ICTAI 05.*, pages 662–667, 2005.

[71] A. Pinz, S. Bernögger, P. Datlinger, and A. Kruger. Mapping the human retina. *IEEE Trans. on Medical Imaging*, 17(4):606–619, 1998.

[72] F. Prokoski and R. Riedel. Infrared identification of faces and body parts. In *BIOMETRICS:Personal Identification in Networked Society.* Kluwert Academic Publishers, 1999.

[73] N. Ryan, C. Heneghan, and P. de Chazal. Registration of digital retinal images using landmark correspondence by expectation maximization. *Image and Vision Computing*, 22:883–898, 2004.

[74] H. M. Schey. *DIV,GRAD,CURL and all that.* Norton, New York, 1973.

[75] L. Seung-Hyun, Y. Sang-Yi, and K. Eun-Soo. Fingerprint identification by use of a volume holographic optical correlator. *Proc. SPIE Vol. 3715, Optical Pattern Recognition*, pages 321–325, 1995.

[76] D. Sidlauskas. 3d hand profile identification apparatus. United States Patent No.4,736.203, 1988.

[77] J. A. Siguenza Pizarro and M. Tapiador Mateos. *Tecnologias Biometricas Aplicadas a Seguridad.* Ra–Ma, 2005.

[78] C. Simon and I. Goldstein. A new scientific method of identification. *J. Medicine*, 35(18):901–906, 1935.

[79] F. W. M. Stentiford and R. G. Mortimer. Some new heuristics for thinning binary handprinted characters for ocr. *IEEE Transactions on Systems, Man, and Cybernetics*, pages 13(1)81–84, 1983.

[80] B. systems Lab. Hasis, a hand shape identification system.

[81] B. ter Haar Romeny and L. Florack. *The Perception of Visual Information (W.Hendee and P Well,Eds.)*, pages 73–114. Springer-Verlag, Berlin/New York, 1993.

[82] M. Tico and P. Kuosmanen. Fingerprint matching using an orientation-based minutia descriptor. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8):1009–1014, 2003.

[83] P. Tower. The fundus oculi in monozygotic twins: Report of six pairs of identical twins. *Arch. Ophthalmol.*, 54:225–239, 1955.

[84] VARIA. VARPA Retinal Images for Authentication. http://www.varpa.es/varia.html.

[85] S. Venkataraman, M. A. Sohoni, and V. Kulkarni. A graph-based framework for feature recognition. In *Symposium on Solid Modeling and Applications*, pages 194–205, 2001.

[86] K. Venkataramani and V. Kumar. Fingerprint verification using correlation filters. In *AVBPA*, pages 886–894, 2003.

[87] B. Victor, K. W. Bowyer, and S. Sarkar. An evaluation of face and ear biometrics. In *ICPR*, pages 429–432, 2002.

[88] P. A. Viola and M. J. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154, 2004.

[89] J. Wayman. A scientific approach to evaluating biometric systems using mathematical methodology. *Card Tech/Secure Tech*, pages 477–492, 1997.

[90] J. C. Whittier, J. Doubet, D. Henrickson, J. Cobb, J. Shadduck, and B. Golden. Biological considerations pertaining to use of the retinal vascular pattern for permanent identification of livestock. *J. Anim. Sci*, 81:1–79, 2003.

[91] T. A. Witten and L. Sander. Diffusion-limited aggregation, a kinetic phenomena. *Phys. Rev. Lett.*, 47:1400–1403, 1981.

[92] J. Woodward. Biometrics: Privacy's foe or privacy's friend? *Proc. IEEE Special Issue on Automated Biometrics*, 85(9):1480–1492, 1997.

[93] R. Xiao, H. Zhu, H. Sun, and X. Tang. Dynamic cascades for face detection. In *ICCV*, pages 1–8, 2007.

[94] M.-H. Yang, N. Ahuja, and D. J. Kriegman. Face recognition using kernel eigenfaces. In *ICIP*, 2000.

[95] H. Zhang, W. Gao, X. Chen, S. Shan, and D. Zhao. Robust multi-view face detection using error correcting output codes. In *ECCV*, pages 1–12, 2006.

[96] W.-Y. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Comput. Surv.*, 35(4):399–458, 2003.

[97] B. Zitová and J. Flusser. Image registration methods: a survey. *Image Vision and Computing*, 21(11):977–1000, 2003.

[98] R. Zunkel. Hand geometry based verification. In *BIOMETRICS:Personal Identification in Networked Society*. Kluwert Academic Publishers, 1999.