Capillary-patterns for Biometric Authentication

Jordi Paloma Benedicto, Gary Garcia Molina, Fons Bruekers, and Cristian Presura
Philips Research Eindhoven

© Koninklijke Philips Electronics N.V. 2007
Concerns: Report
Period of Work: June 2006

Notebooks:

Authors’ address  J. Paloma  jordi.paloma@gmail.com
                 G. Garcia  gary.garcia@philips.com
                 F. Bruekers  fons.bruekers@philips.com
                 C. Presura  cristian.presura@philips.com
Title: Capillary-patterns for Biometric Authentication

Author(s): Jordi Paloma Benedicto, Gary Garcia Molina, Fons Bruekers, and Cristian Presura

Reviewer(s): 

Technical Note: TN-PR 2007/00403

Additional Numbers: 

Subcategory: 

Project: Private and Secure Person Identification (2002-453)

Customer: Philips Research

Keywords: Authentication, Biometric modalities, Blood vessels, Capillary-patterns, Fingers, Privacy, Secure Biometrics

Abstract: In this report, we present a method using the capillary structures under the distal interphalangeal joint (DIP joint), which is located between the second and third (distal) phalanges of the finger, for achieving secure biometric authentication. Images of the DIP joint are acquired using a microscope and they are subsequently processed to enhance the visibility of the blood vessels. DIP joint images cannot be easily faked and are not left behind. During enrollment, a set of images of the DIP joint are recorded and a region of interest that contains a large concentration of blood vessels is selected to serve as template for the enrolled subject. During authentication, the acquired image is compared to the template of the claimed identity by means of a score derived from the 2D cross-correlation. We have recorded a database containing images of sixty fingers (from nineteen subjects) on which our experiments resulted in an optimal equal-error rate (EER) of 0%.
Conclusions: The present study answers some questions that came up at the very beginning about authentication and performance. The main goal was to find out the potential of this new biometric modality, and try to get reliable results which let us evaluate its possibilities in front of the current modalities as fingerprints or palm-veins.

As a first approach, the conclusion is that capillary-patterns in the DIP joint can be used for authentication, because under certain conditions (1) it is possible to obtain a 0% EER, which is the ideal performance. But it is necessary to keep in mind the reduced size of the database. Nevertheless, this result indicates that this modality can meet the requirements to be considered at the same level that, i.e. the fingerprint authentication. When we obtain the EER using a large database, which means to take into consideration the whole set of images and 60 fingers, the ratio is 0.22% (templates only from high quality images). This result is good enough to motivate new research on the digital image processing algorithms and putting more effort on the template selection algorithms in order to be able to use all images in a more efficient way, instead of rejecting some of them due to their quality, etc.

The red images can be very useful from the point of view of correcting green images, but never for helping the authentication itself (the EER authenticating with red images is about 4.41%). They have information about the subjects, which comes from the background of the fingers, but not enough to reach low EERs (this information is also found in the green images). Although we demonstrated that the use of red images for the decision does not improve the EER, we could consider these images for other purposes. Some of the images have low quality because of the presence of dust particles or bubbles. Since these artifacts are present in both green and red images, it is interesting to consider the option of digitally remove these undesired patterns by subtracting the images. Then, an image with a low qualification could be restored and improved. This process could be considered as a processing improvement.

Which finger we use for the authentication seems to be another consideration. It is not the same to authenticate people imaging the capillary-structures of the fingers from their most-used hand, than using the structures of the fingers from the other hand. Even though we found that might be a relation between the EER and the hands, the results are considering a small database. It is something that might be submitted to further research.

The next step could be to assess whether it is worth to develop an algorithm which considers the frequency content of the image; it means that when the frequency is higher in a specific part of the image, the number of blood vessels in that part is also larger and thus, it could make the authentication easier (1).
Acknowledgments

The presented work was carried out from October 2006 to May 2007 at Philips Research, based in Eindhoven, Netherlands, under the supervision of the Universitat Politècnica de Catalunya (UPC) in Barcelona.

I especially want to thank Gary Garcia, my thesis supervisor at Philips, who gave me an important opportunity to start my career at this company, taught me so many things and made this thesis a great experience. I also want to thank him for making the time here enjoyable, sharing the office and playing in the same football team. I would like to thank Fons Bruekers for his advises that will be very useful for both my live and for my career. I will try to remember them in case we meet in Barcelona in the future. Thank you for the fruitful discussions during my work here. Thanks to Cristian Presura for being available every time I required his help and for explaining and showing me every little thing I asked during my thesis about the system.

I also want to thank my supervisor at UPC, Ferran Marquès for his support and for offering me the possibility to discover a great country in a good company.

Finally I thank the students Eric, Ton, Louis, Anka, Rick, Serap, Julien, Peter, Rami, Séverine and Mike for the early lunches and for the discussions at the coffee corner.

Special thanks go to my parents, who greatly supported me during my entire studies.

## Contents

**Acknowledgments**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2 Objectives</td>
<td>3</td>
</tr>
<tr>
<td>3 Major Contributions</td>
<td>3</td>
</tr>
<tr>
<td>4 Outline</td>
<td>3</td>
</tr>
<tr>
<td>2 Capillary-patterns and Acquisition</td>
<td>5</td>
</tr>
<tr>
<td>2.1 Skin and Subcutaneous Vasculature</td>
<td>5</td>
</tr>
<tr>
<td>2.1.1 Skin Structure</td>
<td>5</td>
</tr>
<tr>
<td>2.1.2 Organization of the Subcutaneous Vasculature</td>
<td>6</td>
</tr>
<tr>
<td>2.2 Imaging of the Capillary-patterns</td>
<td>6</td>
</tr>
<tr>
<td>2.2.1 Optical Principles</td>
<td>7</td>
</tr>
<tr>
<td>2.2.2 Optical System Description</td>
<td>8</td>
</tr>
<tr>
<td>2.2.3 Software-based Controlling Means</td>
<td>12</td>
</tr>
<tr>
<td>3 Acquired Images and Quality-based Classification</td>
<td>13</td>
</tr>
<tr>
<td>3.1 Poor-quality images</td>
<td>14</td>
</tr>
<tr>
<td>3.2 Quality-based Classification (QbC)</td>
<td>14</td>
</tr>
<tr>
<td>3 Capillary-patterns as a Biometric Feature</td>
<td>17</td>
</tr>
<tr>
<td>3.1 State of the Art</td>
<td>17</td>
</tr>
<tr>
<td>3.2 General Processing Diagram</td>
<td>19</td>
</tr>
<tr>
<td>3.3 Template Selection</td>
<td>20</td>
</tr>
<tr>
<td>3.3.1 Template Quality</td>
<td>20</td>
</tr>
<tr>
<td>3.3.2 Template Extraction</td>
<td>20</td>
</tr>
<tr>
<td>4 Pre-processing</td>
<td>21</td>
</tr>
<tr>
<td>4.1 Image Filtering</td>
<td>21</td>
</tr>
<tr>
<td>4.2 Interpolation and Subsampling</td>
<td>22</td>
</tr>
<tr>
<td>4.3 Contrast-limited Adaptive Histogram Equalization (CLAHE)</td>
<td>24</td>
</tr>
<tr>
<td>5 Template Matching and Feature Extraction</td>
<td>29</td>
</tr>
<tr>
<td>5.1 Normalized 2D Cross-correlation</td>
<td>29</td>
</tr>
<tr>
<td>5.2 Thresholding</td>
<td>31</td>
</tr>
</tbody>
</table>
# Data Acquisition and Experimental Results

1. Characterization of Biometric Systems ........................................ 34
2. Acquisition and Data Set ......................................................... 36
3. Parameter Specification .......................................................... 37
4. Experimental Results ............................................................. 39
   4.1 Detailed Procedure Explanation .......................................... 39
   4.2 Template Origin .............................................................. 40
   4.3 Magnification .................................................................. 41
   4.4 Subsampling Factor ......................................................... 42
   4.5 LPF .............................................................................. 44
   4.6 Quality ........................................................................... 45

5. Final System Performance .......................................................... 46

6. Additional Parameters ............................................................. 54
   6.1 Green and Red $score$ Distributions .................................... 54
   6.2 Left and Right Fingers ....................................................... 56

## Conclusions

1. Summary of Achievements ......................................................... 58
2. Future Work ........................................................................ 59

### Appendix

3. Mathematical Approach to Bilinear Interpolation .................... 63

### Bibliography

© Koninklijke Philips Electronics N.V. 2007
## List of Figures

2.1 Phalanges and interphalangeal joints ........................................ 6
2.2 Melanin and hemoglobin absorption spectrum ................................. 7
2.3 Penetration depth of UV and visible light through the skin ................. 8
2.4 Cutaneous vasculature and radiation ........................................ 8
2.5 Acquisition device .................................................................... 9
2.6 System magnification and image resolution .................................... 10
2.7 Beam splitter ........................................................................... 11
2.8 Finger related aspects ............................................................... 12
2.9 Scattering effect ...................................................................... 13
2.10 LabView software .................................................................... 13
2.11 Acquired images ...................................................................... 14
2.12 Undesired images .................................................................... 16
3.1 Enrollment diagram .................................................................... 19
3.2 Authentication diagram ................................................................ 20
3.3 Hamming windows ..................................................................... 23
3.4 Image resizing using interpolation ............................................. 24
3.5 Histogram equalization ................................................................ 25
3.6 CLAHE ................................................................................... 28
3.7 Normalized 2D cross-correlation example .................................... 31
3.8 Thresholding ............................................................................ 33
4.1 Impostor curves ......................................................................... 35
4.2 Client curves ............................................................................ 36
4.3 Client and impostor curves .......................................................... 36
4.4 Correlation between templates and images .................................... 39
4.5 EER performance - Magnification ............................................... 42
4.6 EER performance - Subsampling factor ........................................ 43
4.7 EER performance - Hamming window size .................................... 45
4.8 EER performance - QbC ............................................................. 45
4.9 EER final performance ............................................................... 46
4.10 FAR and FRR curves ............................................................... 47
4.11 Distribution of scores ................................................................ 48
4.12 FAR and FRR curves for Ts250/Ts300 ........................................ 48
4.13 Edge effect ............................................................................. 49
4.14 Correction of the edge effect ..................................................... 50
4.15 EER performance - Edge effect ................................................ 51
4.16 FAR and FRR curves for cropped images .................................... 52
4.17 EER performance - Zero-padded images .............................. 53
4.18 FAR and FRR curves for zero-padded images (Ts=600) .............. 53
4.19 FAR and FRR curves for zero-padded images (Ts=400) .............. 54
4.20 EER performance - Red images ...................................... 55
4.21 Green/Red distribution ................................................. 56
4.22 EER performance - Green and red images ............................ 56
4.23 Left/Right finger distribution ......................................... 57
4.24 Right/Left finger performance ........................................ 57

5.1 Bilinear interpolation ....................................................... 63
List of Tables

1.1 Empirical qualitative scoring of biometric modalities .............................. 2
2.1 Magnification and resolution ................................................................. 10
4.1 Parameter specification ........................................................................... 38
4.2 EER performance - Template origin ....................................................... 40
4.3 Wrong correlation decisions ................................................................. 52
Section 1

Introduction

1 Biometric Authentication

Person identification/authentication (I/A) has traditionally been achieved by means of passwords, tokes or PINs, i.e. something that the person possess. Biometrics adds a new dimension to person I/A by using inherent individual characteristics such as something that the person is or produces. Common biometric modalities include fingerprints, voice, facial geometry or iris-patterns.

Security, convenience and privacy protection factors are fundamental to the deployment of biometric systems for I/A. These factors, which are differently traded depending on the application, determine the choice of the biometric modality. In our application scenario, we consider the following factors:

- **SECURITY FACTORS**
  - **False acceptance** i.e. the error incurred by the system when impostors are authenticated by impersonating the biometric data of a client (enrolled user).
  - **Spoof proof** i.e. the robustness against counterfeiting of biometric templates using biometric-faking devices, e.g. fake fingerprints, contact lenses with fake iris patterns, face masks, etc.

- **CONVENIENCE FACTORS**
  - **False rejection** i.e. the error incurred when authentication of enrolled users fails.
  - **Universality** i.e. the degree to which the biometric data is present in everyone.
  - **Usability** i.e. the extent to which the acquisition of the biometric data is accepted by the users in terms of comfort, speed, easiness, etc.

- **PRIVACY FACTORS**
  - **Protection of private information (PPI)** i.e. the extent to which a biometric template prevents revealing personal information, e.g. gender, health condition, etc.

False rejection and false acceptance are mutually dependent and can be traded depending on the application. Indeed, lower false rejection leads to higher false acceptance. The performance of a biometric system is in general indicated in terms of the equal-error rate, which indicates the value at which the false acceptance is equal to the false rejection.
The most common biometric modality is fingerprint, which is used world-wide for many applications with an EER of about 1%. However, there are some other biometric modalities which have been proposed as alternative techniques. After a brainstorming session involving a multi-disciplinary team (2), a number of possible new body-measurements was considered. Hence, one of the priorities before starting the research was to choose which one of new proposed technique seemed to be more promising to develop. Thus, a brief study of the main properties, advantages and disadvantages was made in order to describe and assess every option. The selection was made based on feasibility and novelty as well as considering the above mentioned factors. In Table 1 we report a qualitative scoring of some biometric modalities.

Table 1.1: Empirical qualitative scoring of biometric modalities.

<table>
<thead>
<tr>
<th>Biometric modality</th>
<th>EER</th>
<th>Spoof proof</th>
<th>Universality</th>
<th>Usability</th>
<th>PPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fingerprint</td>
<td>good</td>
<td>poor</td>
<td>good</td>
<td>good</td>
<td>good</td>
</tr>
<tr>
<td>Iris recognition</td>
<td>excellent</td>
<td>good</td>
<td>good</td>
<td>poor</td>
<td>good</td>
</tr>
<tr>
<td>Facial scan</td>
<td>good</td>
<td>fair</td>
<td>excellent</td>
<td>fair</td>
<td>poor</td>
</tr>
<tr>
<td>Hand scan</td>
<td>fair</td>
<td>fair</td>
<td>excellent</td>
<td>fair</td>
<td>fair</td>
</tr>
<tr>
<td>Voice recognition</td>
<td>fair</td>
<td>poor</td>
<td>excellent</td>
<td>good</td>
<td>fair</td>
</tr>
<tr>
<td>capillary-patterns</td>
<td>excellent</td>
<td>excellent</td>
<td>good</td>
<td>fair</td>
<td>excellent</td>
</tr>
</tbody>
</table>

In this study we aim at enabling secure authentication that in addition of requiring excellent biometric performance and ease of use, are concerned by robustness to spoofing attempts, and privacy protection. As we said, fingerprint is the most extended modality, even though it is clear that it has a weak point: it is possible to counterfeit an identity. For the other modalities, we can find weaknesses in other aspects, because it is difficult to meet all the conditions for one modality. However, Table 1 shows that in the case of capillary-patterns (recognition based on blood vessels, more exactly, the blood vessels in the finger), the preliminary evaluation is very good, and as a first approach, only the usability seems to be compromised. Its properties make this modality to be a worthy candidate for further research.

In section 4 we present the results that confirm the excellent behavior in terms of EER, although we have to consider it under specific conditions. When we talk about spoofing attempts, the capillary-recognition has very good intrinsic properties which make it suitable for our purposes. Whereas for the rest of the modalities we are measuring external features, the capillaries are placed within the body. This characteristic is the basis of the robustness to counterfeiting: it is extremely difficult to obtain a reproduction, a recording or something that permits to impersonate an identity; only the original source is valid. Since we are measuring properties of the blood vessels, blood flow is required and consequently aliveness, which is a considerable advantage with respect to other modalities. For the other techniques it may be easier to obtain a fake reproduction or left samples behind with which are possible to counterfeit an identity. The universality is a very important factor to consider, since we aim at people authentication. Hence, we need to measure something that is present in everybody in order to be able to enroll anyone. Moreover, if we want to develop a new modality which can be implemented for secure applications, it has to meet usability conditions. In the case of the the technique we are considering, the subject only has to place the finger on a specific device (as in the fingerprint case) and the system captures an image. This is the expected procedure for an application which is based on capillary-recognition. It is easy and quick to obtain the sample, and it does not suppose an effort for the subject to take part in the authentication process. Finally, we are looking for a new way of authenticating people, which means that we are not allowed to obtain more information than.
what is strictly required. Since we only image the capillary-structures and focus our effort to the blood vessel-patterns, no personal relevant information is revealed; we do not analyze the intrinsic characteristics of the subject, only the distribution and location of the capillaries.

Due to all the characteristics and properties, capillary-pattern recognition is the modality that better suits the goal of finding an alternative to the existent authentication methods. We center the study in the authentication process, where an identity has to be accepted or rejected. The identification problem is not considered in the study.

2 Objectives

The goal of this thesis is to test a new biometric modality, considering the advantages and disadvantages that it may entail. We study a new authentication method based on the capillary structures in the fingers, since as it has been demonstrated, it has enough potential to be considered as a solution in terms of security. The following goals are pursuit:

- To record a database of DIP joint images.
- To assess the biometric performance in terms of EER.
- To identify the key factors influencing that performance.

3 Major Contributions

The major contributions of the research presented in this thesis are summarized as follows.

- Creation of a capillary-pattern database.
- Implementation of
  - Algorithms for template extraction.
  - Algorithms for image processing.
- Demonstration of the fact that capillary-patterns can be used for authentication.
- Study of the effect of
  - Influencing factors,
  - Illumination,
  - Right and left-handed subjects.

4 Outline

Section 2 briefly explains the morphology of the skin and the subcutaneous vasculature. The acquisition system and the basic principles for the capillary imaging are also detailed in this section. At the end, we present our criterion to judge the quality of the obtained images in order to be able to perform the simulations.

Section 3 contains the state of the art and details the main structure of the system. It also presents the algorithms and image processing techniques used in the study, giving the fundamentals of each one and the mathematical approach which let us deeply understand these concepts.
Section 4 deals with the results obtained. In the first part, there is a discussion about the optimum parameters that have an effect on the behavior of the system, always supported by plots and images. The final performance is discussed in this section, according to the chosen parameters and the conditions detailed in the other sections.

Section 5 discusses the conclusions that have been achieved at the end of this thesis. The suggestions for future work are also contained at the end of this section.
Section 2

Capillary-patterns and Acquisition

In this section, the methods, principles, and devices that are used to acquire the capillary images are presented. In Section 1, we briefly introduce the basic concepts of the skin anatomy and the capillary structures under the skin. Section 2 explains the imaging device and its components. Section 3 includes the main properties of the acquired images, as well as a discussion about issues responsible for degrading the quality of the captured images. A criterion that makes possible to judge their quality is also detailed here.

1 Skin and Subcutaneous Vasculature

In this section we provide a brief summary about some basics of the morphology of the skin and subcutaneous vasculature. A more detailed description is presented in (3) and (4).

1.1 Skin Structure

The skin is a constantly changing organ which is made of many specialized cells and structures. It is a protective barrier against external agents and it is basically composed of three layers: epidermis, dermis, and subcutaneous tissue.

A) Epidermis. The epidermis (epi- = upon) is the outer layer of skin. The thickness of the epidermis varies in different types of skin: it is the thinnest on the eyelids at 0.05 mm and the thickest on the palms and soles at 1.5 mm. The epidermis contains 5 layers. From outside to inside, the layers are named: stratum corneum (8–15 µm), stratum lucidum (highly refractive sublayer), stratum granulosum (~3 µm), stratum spinosum (50–150 µm) and stratum basale (the thickest sublayer).

There are four types of specialized cells in the epidermis:

- The keratinocytes produce keratin, the fibrous protein that gives the epidermis its protective properties.
- The melanocytes synthesize the pigment melanin.
- The Langerhans’ cells (or epidermal dendritic cells) are phagocytes that ingest foreign substances and help to activate the immune system.
- The Merkel’s cell’s function is not clearly known.
B) Dermis. The dermis also varies in thickness depending on the location of the skin. It is 0.3 mm on the eyelid and 3.0 mm on the back. The dermis is composed of two layers where it is possible to find three types of tissue that are present throughout: collagen, elastic tissue and reticular fibers. The *papillary layer* (upper) and *reticular layer* compose the dermis.

C) Subcutaneous Tissue. The subcutaneous tissue, also named hypodermis, is a layer of fat and connective tissue that houses larger blood vessels and nerves. The size of this layer varies throughout the body and from person to person.

A more detailed explanation can be found in (4; 3; 5).

1.2 Organization of the Subcutaneous Vasculature

The subcutaneous vasculature of the skin (located throughout the three mentioned layers) is mostly composed by three kinds of blood vessels: arterioles, capillaries and venules. The arterioles and venules are located on the adipose tissue of the subcutaneous tissue, and lying parallel to the skin surface. These vessels have several ramifications, some of which supply body tissues, and some of which rise vertically to interconnect with the epidermic tissue. The latter are the capillaries, and are the smallest blood vessels in the body. The capillaries are the ones that provide direct access to nearly every cell in the body. The average capillary length is 1 mm and its average diameter is about 8–10 µm. They lie in the papillary layer of the dermis, approximately 1–2 µm below the epidermis, and they end into dermal *papillary loops* (6; 7).

2 Imaging of the Capillary-patterns

Because of the optical properties of the skin and blood, it is possible to make a system with which we can obtain clear and detailed images of the capillary structures for biometric purposes. The measuring principle is based on the reflectivity of the skin and blood components.

We focus this study on a specific area of the finger that has some optical and practical advantages. This interesting area is named *Distal Interphalangeal joint* (DIP joint) and, as Fig. 2.1 shows, it is located between the second and third (distal) phalanges of the finger. The main reason for choosing the DIP joint for extracting biometric features is due to its convenience and ease of acquisition: the epidermis is thinner at the DIP joint than the rest of the hand (8), and its proximity to the fingerprint makes it a better candidate than the other joints (see section 2).

![Figure 2.1: Phalanges and interphalangeal joints of the finger. This figure has been adapted from (9).](image-url)
2.1 Optical Principles

As mentioned in section 1.1, the epidermis contains four different cell types. They are mostly keratinocytes, which produces keratin, but from the optical point of view, the melanocytes have more interesting properties. The latter produce melanin, which is a very complex absorbing ultraviolet radiation material. Because melanin is an aggregate of smaller component molecules, it falls into two general classes (10; 11):

- Eumelanin is found in gray, black, yellow, and brown hair and skin. In humans, it is more abundant in people with dark skin.

- Pheomelanin is also found in hair and skin, and is more abundant in lighter skinned humans.

Except from the melanin (see Fig. 2.2a), the main absorber of visible light in normal skin is blood, or more precisely, the hemoglobin in the blood. The hemoglobin, the protein that makes erythrocytes red, joins easily and reversibly with oxygen carried in the blood, and it can be found in two different states. When the oxygen binds to the hemoglobin protein, is called \textit{oxyhemoglobin}. When oxygen detaches from the cell, hemoglobin recovers its original shape, and the resulting \textit{deoxyhemoglobin} becomes dark. These two states of hemoglobin are important due to their absorption spectrum shape as it is depicted in Fig. 2.2b (12). Their curves differ slightly (see Fig. 2.2b), but both have absorption peaks at similar positions (8). Mostly, due to these two absorbers, the penetration depth of the light in human skin will be limited.

\[\lambda (\text{nm})\]

\[\text{Extinction Coefficient} \left(\text{cm}^{-1}\right)\]

\(\leftarrow\) Eumelanin

\(\leftarrow\) Pheomelanin

![Graph (a) Melanin absorption spectrum.](image)

![Graph (b) Hemoglobin absorption spectrum.](image)

Figure 2.2: (a) Absorption spectrum of the two basic components of melanin. (b) Absorption spectrum of Hb (deoxyhemoglobin, with peaks at 272, 434 and 556 nm) and HbO\(_2\) (oxyhemoglobin which has peaks at 274, 344, 414, 542 and 576 nm).

The penetration depth is defined as the distance into the tissue at which the incident energy density of a wide, parallel beam of radiation is reduced to \(e^{-1}\) of its value close to (below) the surface (see Fig. 2.3). As it is discussed in section 2.2, it is an important parameter to consider, because the penetration depth is strictly dependent on the wavelength (\(\lambda\)).

Not all “lost” radiation is due to hemoglobin absorption, though. About 5% of the radiation is reflected from the outer surface of the skin. The radiation coming back from the skin is composed of reflected radiation and radiation scattered in the epidermis. Figure 2.4 illustrates another scattering radiation which is reflected by the surface of the skin. We need to consider this scattered energy since the irregularities of the skin are responsible for the loss of an important amount of radiation energy.
Figure 2.3: The penetration spectrum of light and UV radiation into fair Caucasian skin. The scattering increases with decreasing wavelength (13).

2.2 Optical System Description

The optical system consists basically of a microscope, a pair of LEDs, a camera, some lenses, a beam splitter, a positioning structure and a metallic support where to fix all the components on (see Fig. 2.5).

Microscope. The first component is an optical microscope, which offers nine different objectives (element No. 1 in Fig. 2.5): M0.57x, M0.8x, M1.0x, M1.25x, M1.6x, M2.0x, M2.5x, M3.2x and M3.6x. In addition to the microscope amplification, a lens (element No. 2 in Fig. 2.5) doubles the magnification of the microscope (2.0x). The magnification choice is critical for the visualization of the blood vessels, because the higher the magnification, the larger the size of the vessels in pixels (see Table 2.1), and consequently, more suitable for our purposes (see section 4.3). Through experimentation with different magnifications, we calculated the exact resolution of the microscope for each objective (to have an idea about the covered area for

Figure 2.4: Schematic diagram of skin, cutaneous vasculature and optical properties. This image has been adapted from (7).
a given magnification), and we also conclude that the average blood vessel size in the DIP joint is 25–50 µm, which means that, for a M0.57x objective, the capillaries have an average size of 5–10 pixels. Despite of the fact that the diameter of the capillaries is about 8–10 µm, from the point of view of the microscope, they look like short, thin, black strips, because they are curved over themselves. All details concerning magnifications and image resolutions can be found in Table 2.1 and in Fig. 2.6.

The microscope offers the possibility of focusing the image using the adjustment knob (element No.3 in Fig. 2.5), placed at the lower part of the device.

**LEDs.** The acquisition device is equipped with two LED’s which have been chosen following specific considerations (blood light absorption, penetration depth. . . ). The aim of the sensor is to image the capillary structures under the DIP joint by receiving the light from the LED’s that has been reflected by the skin and blood. So, we should have a look at Fig. 2.2 and Fig. 2.3 and understand what is going on in every particular case. As it was mentioned before, melanin hardly contributes to light absorption due to its monotonically decreasing characteristic curve. But both oxyhemoglobin and deoxyhemoglobin curves are the basis of the system, and their
Table 2.1: Magnification and resolution. There is a trade-off between the chosen objective and the convenience of the system. It is explained in section 4.

<table>
<thead>
<tr>
<th>Objective</th>
<th>Real Magnification</th>
<th>Resolution</th>
<th>Mean Vessel Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>M0.5x</td>
<td>1.14x</td>
<td>200 pix/mm</td>
<td>5–10 pix</td>
</tr>
<tr>
<td>M0.8x</td>
<td>1.6x</td>
<td>280 pix/mm</td>
<td>7–14 pix</td>
</tr>
<tr>
<td>M1.0x</td>
<td>2.0x</td>
<td>350 pix/mm</td>
<td>8–18 pix</td>
</tr>
<tr>
<td>M1.25x</td>
<td>2.5x</td>
<td>430 pix/mm</td>
<td>10–22 pix</td>
</tr>
<tr>
<td>M1.6x</td>
<td>3.2x</td>
<td>550 pix/mm</td>
<td>13–28 pix</td>
</tr>
<tr>
<td>M2.0x</td>
<td>4.0x</td>
<td>700 pix/mm</td>
<td>17–35 pix</td>
</tr>
<tr>
<td>M2.5x</td>
<td>5.0x</td>
<td>870 pix/mm</td>
<td>24–44 pix</td>
</tr>
<tr>
<td>M3.2x</td>
<td>6.4x</td>
<td>1120 pix/mm</td>
<td>28–56 pix</td>
</tr>
<tr>
<td>M3.6x</td>
<td>7.2x</td>
<td>1280 pix/mm</td>
<td>32–64 pix</td>
</tr>
</tbody>
</table>

peaks determine what kind of LED is optimum. In order to acquire capillary images with good visible capillaries, it is convenient to use light with a wavelength which corresponds to any of the absorption peaks of hemoglobin. Thus, is possible to get an image where capillaries have

Figure 2.6: Resolution of the instrument: 100 µm; (a) 1.14x-200 pix/mm; (b) 1.6x-280 pix/mm; (c) 2.0x-350 pix/mm; (d) 2.5x-430 pix/mm; (e) 3.2x-550 pix/mm; (f) 4.0x-700 pix/mm; (g) 5.0x-870 pix/mm; (h) 6.4x-1120 pix/mm; (i) 7.2x-1280 pix/mm. These images have been processed to make the visualization better.
absorbed part of the emitted energy and that it will not be received by the CCD (charge-coupled device) of the camera. The detected radiation comes from the reflected light of the skin.

If we try to figure out useful wavelengths for our purpose, the first interval to consider is thus 272–344 nm, where oxyhemoglobin has two relative maxima and deoxyhemoglobin only one. Taking into account the information of Fig. 2.3, though, this wavelengths can penetrate skin tissues at best till 100 µm, with which is possible to reach only the shallower papillary loops. The next interval to think about is 414–434 µm, which contains the first high absorption peaks of hemoglobin. According to the penetration depth curve, light at this wavelength has lost about 75% of its intensity when reaching the capillaries located at 500 µm depth. Considering that the light has to return up through the skin again, almost all intensity is then lost. The last interval, 542–576 µm, although it has a lower absorption coefficient, the penetration depth is greater than 1 mm. It means that the emitted light has energy to reach the deepest capillary loops and return up to the receiver with enough energy to obtain an image. As a result, the installed LED, which displays the capillary structures, has a dominant wavelength of 530 nm, which corresponds to the green wavelength interval (see element No. 4 in Fig 2.5).

The larger the wavelength is, the deeper the penetration and lower the absorption. It means that using a wavelength larger than 600 µm it would be possible to get images with little light absorption. Applying this idea, we can acquire images with the background information only. To get this kind of images, we use a LED with a dominant wavelength of 627 nm–red wavelength interval (element No. 5 in Fig. 2.5): the penetration depth is more than 1.2 mm and there is practically no hemoglobin absorption. The final version of the optical system is equipped with green and red LEDs.

**Beam splitter.** A special optical element is needed to focus the two light beams coming from different directions (forming a 90 degrees angle) into one point within the field of view of the microscope. This element is a beam splitter (element No. 6 in Fig. 2.5), placed at half way between the LEDs and the microscope. This arrangement is detailed in Fig. 2.7.

![Figure 2.7: The light goes through the optical device, which split the initial beam in two perpendicular beams. This case is simulating the red LED working period.](image)

**Camera.** Just below the microscope, the camera is the element (element No. 7 in Fig. 2.5) that let us capture the images through it, giving different resolution options: 640 × 480, 800 × 600, 1024 × 768, 1280 × 960, and 1600 × 1200 pixels. For each one, there are basically three possible frame rates: 3.75, 7.5 and 15.00 fps. The camera receives the radiation of the LEDs once it has passed through the splitter, reflected by the skin layers and returned through the microscope objective.
Positioning structure. The finger is positioned on a support which consists of a mechanism that can move backward and forward precisely with a thread, on a glass-made round plate (Ø 6 cm). The moveable mechanism has a stainless steel piece with a “になりました” shape to fix the finger on the glass plate—being in touch with the two sides of the metallic structure—and let it move accurately using a graduated rule on an adjusting screw (element No. 8 in Fig. 2.5 and Fig. 2.8b). The objective of this structure is to fix the finger in order to localize the interesting area from where we want to take the picture. We define region of interest (ROI) the area of the DIP joint that is imaged and stored in the database, and it is not the same for every subject: it depends on how the finger is put to image it.

![Positioning structure diagram](image)

(a) The oil is located between the finger and the glass plate in order to reduce the scattering effect at the surface.  
(b) The finger has to be positioned in this way.

Figure 2.8: Finger related aspects.

Fluid interface. Because of the roughness of the skin surface, and the optical properties of its layers, two scattering effect might happen. When the light travels through the skin tissues, it is diverted and cannot be received by the camera; this effect is solved with the proper selection of the LED wavelength taking into account the penetration depth and the absorption coefficient. The second scattering effect happens when big portion of the light beams which are not locally perpendicular to the skin surface is reflected. The reflection is due to the difference in the refraction index of the stratum corneum \((n_{\text{skin}} \approx 1.55)\) and the air \((n_{\text{air}} \approx 1.00)\) (see Fig. 2.9a). In order to soften this scattering effect, we use a fluid interface (oil) to match the refraction index of the skin with that of the surrounding environment \((n_{\text{oil}} \approx n_{\text{skin}})\). Since the border between the oil and the air is very smooth, all light beams can be perpendicular to that border and almost no reflection occurs. Thus, the light beams can go straight through both the oil and the skin surface (Fig 2.9b). Besides, the oil is a highly convenient for smoothing the skin, since it fills all irregularities of its surface, but it reduces the usability of the system (in terms of user comfort). Figure 2.8a shows where the oil is used. Some alternatives are discussed in section 2.

2.3 Software-based Controlling Means

All the components of the system need to be controlled, and a specific program (designed and implemented using LabView) let us acquire the images just by configuring a few parameters in
Figure 2.9: Scattering effect on the skin surface.

Figure 2.10: LabView screen shot.

3 Acquired Images and Quality-based Classification

The acquisition device provides grayscale images with 256 intensity levels (8 bits) and they are stored as bitmap files (.bmp) with a resolution of $1200 \times 1600$ pixels.

As it has explained in the previous section, depending on the wavelength of the light, some blood components absorb the radiation and some do not. When we illuminate the skin with green light, we obtain an image where is possible to distinguish black dots or stripes (papillary loops) on a gray background (where we can see some structural shapes). So, green images provide information about the capillary structures and the background. On the other hand, red light, with a larger wavelength, cannot penetrate so deep and it is not absorbed by hemoglobin either. Therefore, in red images there is only information about the background (see Fig. 2.11).
3.1 Poor-quality images

In addition to the parameters that can be adjusted during the enrollment, there are some other factors that might decrease the quality of the captured images. Some of these factors depend on the subject itself. The situations that make an image to be of low quality for our purposes are basically five:

1. The subject clearly moves the finger just when the device is capturing the image. The result is a blurred image (see Fig. 2.12a and Fig. 2.12b).

2. If the subject removes the finger before both the green and the red pictures are taken, the system acquires a good green image, but a useless red one (totally black) (see Fig. 2.12c and Fig. 2.12d).

3. It comes up when the subject exerises a high pressure with his/her finger. Then, the blood goes away from the capillary vessels of the finger, and only some blood vessels appear on the image (see Fig. 2.12e and Fig. 2.12f).

4. If the ROI is not completely covered by the oil, or if the quantity of fluid is not enough, few air bubbles may appear on both green and red images (see Fig. 2.12g and Fig. 2.12h).

5. Dust particles sometimes appear on the images. Similarly to the bubbles, they appear as black artifacts in both images (see Fig. 2.12i and Fig. 2.12j).

Besides the mentioned problems, the properties of the images might be degraded because of the acquiring software parameters like light intensity, shot delay, etc. and because of the subject behavior (finger position). Despite of the fixing structure of the sensor, the finger is not totally subjected, so it is possible to have slight variations between the registered images and the new acquired images: translations along \( x \)-axis and \( y \)-axis and rotations. A big translation or a remarkable rotation might cause to get an image where the ROI is not on it, making the matching and the recognition definitely impossible.

3.2 Quality-based Classification (QbC)

The final results of the system strictly depends on how good the images are, in terms of definition, contrast and brightness. But it also might change due to the above commented problems.
that sometimes come up. Taking it into account, it would be helpful for the final results to create different sets of images considering some quality rules. This classification is required since the developed work is only a first approach of the final system. It means that we must use the stored images and run experiments with them in order to figure out whether is proper to use all of them regardless the quality, or if it is better to consider some parameters that evaluate how good they are.

The first approach is to qualify every image manually following a predefined criteria. The objective of defining rules to classify the images is to try to make the process as objective as possible. An improvement related with the image classification is detailed in section 2. The rules are:

1. All the images in the data set have a qualification
2. An image receive a qualification in the 1, 2, 3 range.
3. An image with the highest quality ($Q_3$) meets the following criteria:
   - It must be clear.
   - The capillary structures are not blurred.
   - It must be bright enough to distinguish easily the vessels.
   - No artifacts (bubbles, dust . . . ) on the image.
   - The enrolled images from the same subject must show approximately the same area of the finger.
4. If the image does not meet one or more of the previous rules, it could belong to the second group of images: 2-qualified images ($Q_2$):
   - The image is clear, but some small artifacts are present on both the green and red image (in the same place).
   - A bit blurred image.
   - Dark images.
5. When the acquired image cannot be considered in the previous two groups, usually it is because:
   - The image does not have enough dots due to the pressure applied while capturing.
   - The image is very blurred.
   - Particles on the image with such a remarkable size that does not let the information be shown.
   - The images do not correspond with the main selected area for that subject.

In this case, the images have a qualification of 1 ($Q_1$).

Applying the rules to the whole set of training images available, three groups of images are ready for the experiments: $Q_3$, $Q_2$ and $Q_1$. From now on, when we talk about quality, we consider that we use images/templates of the mentioned quality or superior. That is, if we use $Q_2$ images, it means that the experiment runs using $Q_2$ and $Q_3$ images (not $Q_1$). The same with the templates. A sample of some of these images is depicted in Fig. 2.12.
Figure 2.12: (a) and (b) are blurred images due to the movement of the finger during the capture; In (c) and (d) the subject removed the finger before the acquisition device took the red picture; (e) and (f) come from a session where the subject was pushing the glass plate while testing; (g) and (h) show the effect of a bubble; In (i) and (j) is clear the presence of dust particles.
Section 3

Capillary-patterns as a Biometric Feature

The first section of the section puts the reader in context, mentioning the last advances related with our biometric modality and the techniques they use for the authentication. Thus, we will be in position to start with the main aim of the section: the discussion of the methods that makes the images suitable for a later use, let us say pre-processing, the presentation of the basic processing algorithm as well as the mathematical explanation lying behind. Moreover, in order to make the different processing steps more understandable, a general diagram will be presented at the beginning, and we will refer to it several times along the section.

1 State of the Art

Traditionally, person authentication has been accomplished by associating to the person’s identity something that he/she possesses (e.g., a key, a card, etc.) or knows (e.g., a password, a PIN). Biometric recognition adds a new dimension by associating a person’s identity with something that he/she is (or produces). Something that a person is indicates a physiological characteristic inherently associated with the person.

Infrared images of some parts of the human body (face, palm, etc.) have been also used for biometrics in (14; 15; 16). One of the latest research study, related with our modality, uses infrared images for detecting palm vein-patterns within inside the body. The palm is a suitable part of the body for this technology; since it does not have hair which can be an obstacle for photographing the blood vessel pattern, and it is less susceptible to a change in skin color, unlike the back of the hand. It uses, as capillary-pattern detection does, a specific optical principle of the skin and blood: the light absorption. The palm vein sensor captures an infrared ray image of the user’s palm. The lighting of the infrared ray is controlled depending on the illumination around the sensor, without special consideration in the position of the palm. When the infrared ray image is captured, only the blood vessel pattern is visible as a series of dark lines. The software then computes the position and the orientation of the acquired image, and the vein authentication device translates the black lines of the infrared ray image as the blood vessel pattern of the palm. Finally, the system matches this pattern with registered pattern contained in the database (14).

The main advantage of this technique resides in the fact that it can unobtrusively acquire the relevant data using contactless authentication technology. Because of the properties of the system, it is already being used in financial solutions. The illegal withdrawal of bank funds using
stolen or fake bank cards has been reduced since palm vein authentication has been utilized for costumer confirmation of transactions at bank windows or ATMs. The smart card from the customer’s bank account contains the customer’s palm-vein-pattern and the matching software of the palm veins pattern. A palm vein authentication device at the ATM scans de customer’s palm vein pattern and transfers it into the smart card. The blood vessel pattern is then matched with the one registered in the smart card. Since the registered customer’s palm vein pattern is not released from the smart card, the security of the customer’s vein-pattern is preserved.

The palm vein detection technology is also used as an access control unit. A special device controls access to rooms or buildings that are for restricted personnel. These devices consists of the palm vein sensor plus the control unit that executes the processing part (15). This technique is interesting because it deters forgery, because its biometric source is internal, which makes it less sensitive to external factors. The palm vein authentication offers a high level of accuracy: a false rejection rate (FRR) of 0.01% and a false acceptance rate (FAR) of 0.00008% or lower, based on Fujitsu research using 140000 palms as a database (16).

The capillary structures are also used for clinical purposes in (8). Skin microcirculation, especially the superficial network, can be assessed by a computer capillary video microscope system. Capillary microscopy is a method to visualize the top papillary loops, or actually the red blood cells in these loops located immediately under the skin. Using this technique is possible to measure parameters such as diameter and blood flow in the individual capillaries in order to characterize and diagnose some diseases like hypertension, diabetes, inflammation, ischemia, connective tissue disease and erythromelalgia. However, not only the information of the capillary patterns themselves is interesting. When diagnosing skin diseases, it should be taken into account some other parameters like capillary distribution or blood flow direction. Some projects were aiming for developing an application which made an automatic capillary distribution analysis, computation and storage in a database. The capillaries were identified in microscope images and some experiments were run in order to investigate how polarization filters affected those images and also how various light sources could change the image properties. In that case also, was fundamental to know the properties of the skin and the blood components since they were working with optical parameters like wavelengths and polarization.

The acquired images in (8) resemble the ones that we use in this project: grayscale images with black dots on a gray background in case of acquiring the image with a green light. The study also considered the images taken with red and blue light. As a pre-processing, they were using Gaussian filters for noise reduction, high-pass filters to enhance the capillary edges of the images, background subtraction, thresholding, capillary segmentation and triangulation methods. In order to improve the performance of the system and to make the blood vessels more visible, a pilot experiment was run: they studied the effect of applying a local anaesthetic cream on the skin.

The most important difference between this method and capillary analyzing methods is that this one focuses on global capillary parameters, while others usually analyze the parameters of the single capillaries, such as diameter and shape. The pros of computerized methods are that they are much faster than manual measurements, and they are also more objective due to the automatic analysis. This particular method is however still subjective, since some parameters need to be adjusted manually. Moreover, it is still the human eye that makes the final conclusion about the capillary network, and that conclusion might differ from person to person and over time (8).

On the same direction of the last research work, some studies about geometrical capillary network analysis have been developed at some biologic labs and hospitals of France (17; 18). The microvasculature of the skin and the microcirculation were investigated by combining
videocapillaroscopy (VCP) and image processing techniques based on computational geometry and graph theory. Different geometric methods were developed, based on proximity parameters (distance and surface) in order to define and construct the capillary network. Different algorithms were developed and were implemented in an image processing software (17). The same authors published another work about recognition of blood vessel networks by means of neural algorithms. A detection system was tested by combining videocapillaroscopy and principal component analysis (PCA). Their goal was to build a generic detector of capillary associated with a connected neural network filter. The filter examines small windows of an image, and decides with this detector whether each window contains a capillary or not. Comparisons with manual detections showed that the system has a detection rate of 82% on test set A containing 100 good-quality images and 65% was obtained on test set B containing 50 images with noisy background and large artifacts. The performance was increased by a color detector with a detection rate of 71% on the last test. These results correspond to a false detection rate lower than or equal to 10% (18).

Among the relative few studies in the field of capillary pattern analysis, only the vein-patterns in the palm considers biometric applications.

2 General Processing Diagram

The acquired images are processed according to the scheme represented in Fig. 3.1.

When we talk about the general diagram, we must distinguish two different situations: enrollment and authentication/verification. The enrollment is the process of extracting templates from the acquired data, that are relevant for authenticating a subject. Any user of the system must be enrolled. One or more pictures of the ROI are taken in order to enroll the subject and to store them in the system. Then, the information extracted from those images will be used as a pattern for recognition. The enrollment is supervised by an operator in such a way that low quality measurements are rejected. As Fig. 3.1 and Fig. 3.2 illustrate, the subsampling algorithm and the equalization process are the same for both the enrollment and the authentication.

Figure 3.1: Enrollment diagram. The acquired images are \((A_x \times A_y)\), the extracted templates \((B_x \times B_y)\) and the subsampled templates \((N_x \times N_y)\). In this case, the extraction method for the templates would be taking the central part of the image. See section 3 and section 4.

The first blocks of the diagram of Fig. 3.1 are the acquisition and the classification of the images, which have been explained in section 2. The rest of the processes concerning the diagram is detailed in this section.

The goal of the authentication process is different since the main task is the authentication of a claimed identity. The acquired image is processed (like during the enrollment) but instead of storing it, it is compared with the registered patterns by means of a matching process. Thus, the system obtains a score from which it will decide if the claimed identity belongs to it or not.
In Fig. 3.2 the complete diagram for the authentication process is detailed.

Figure 3.2: Authentication diagram. The acquired images are \((A_x \times A_y)\), the subsampled images \((M_x \times M_y)\) and the subsampled templates \((N_x \times N_y)\). The main block decides whether the subject is successfully authenticated. See section 4.

The classification block makes reference to the subjective QbC of the acquired images, strictly attending to the described criteria (see section 3.2). The template extraction block is only present in the enrollment diagram. Then, we find the filtering and subsampling and the CLAHE (Contrast-limited adaptive histogram equalization) block. These ones are common in both diagrams and are detailed in the following sections. From this block on, the diagrams differ basically in two blocks: the matching and the decision block, which are explained at the end of the section.

### 3 Template Selection

The main reason to design an image-QbC method, apart from knowing how good is an image, is to decide which images are suitable for processing. The study of their properties is basic in authentication/identification systems, because they are the main tool from which the rest of the processes will arise.

#### 3.1 Template Quality

The template is selected from each acquired image during the enrollment, and only those having an acceptable level of quality are used for processing (it is possible to use more than one template per subject).

Since the enrollment is always done under supervision, it is fair to consider that all enrolled images have good quality. This basic principle is carried out by extracting only the templates from \(Q_3\) images. Thus, it is for sure that all the templates have very good quality as if they came from a supervised enrollment process. A template is never selected from images which have a qualification lower than 3 (so templates \(\in Q_3\)), because then, the results will not be representative of a real biometric system. Section 4 explains deeper the use of the templates and the images, as well as how they interact for the feature extraction.

#### 3.2 Template Extraction

The next step is to figure out what part of the image is the best for correct authentication. Many different methods can be applied in order to find the best solution. The first trial: manual extraction. For every image of the database, we select a region by hand, approximately the same for each image of the same subject and having the same size each one. The main problem of these selection is that if the data set is too large, it is not possible to do the selection manually.
because of time considerations. So an automatic template extraction is then considered: we use the central part of the image as a template; this is the easiest automatic selection that we can do.

Every time a template has to be selected, the next step and, at the same time, obvious, is to think about its size. In other words, we have to figure out how big needs to be the area that we select from the original images as a template. It is logic to think that there is a strict relation between the number of pixels of the template and the final results (section 4). Indeed, it exists an optimal size. Intuitively, the larger the number of pixels, the larger amount of information that is contained in the template. Then, we could think that if we use the biggest template we can afford, the result will be the best. Because of the processing technique and the size of the images, it is not possible, not efficient and not fast to work with the complete images as a template. Working with smaller templates makes the processing much faster. So, there is an important trade-off to consider.

4 Pre-processing

Even though an image-QbC has been done in order to work with suitable images, not all of them are ready to be used in the processing. As the diagrams show (Fig. 3.1 and Fig. 3.2), some previous transformations and changes are applied to them before the matching process itself (explained in section 5). These pre-processing techniques are the filtering, subsampling and a histogram equalization. The pre-processing has to be applied in both enrollment and authentication processes right before the storage and the matching, respectively. If we are talking about the enrollment, the pre-processing will be applied to the extracted templates. Thus, the database only contains data about pre-processed templates, linked with the identity of the subjects. On the other hand, during the authentication process, the system pre-processes the complete acquired images directly, and no data is stored.

4.1 Image Filtering

This section briefly introduce some basic points of the sampling theory and filtering related aspects. A more detailed presentation of these concepts can be found in (19; 20; 21).

In this study, subsampling is used to diminish the processing time. This technique allows the system to work and make calculations by accessing to a minimum amount of data. Therefore, the system can get the results faster, depending on the filter complexity, the subsampling factor, etc.

Generating a subsampled image is simply a sampling process, so it is subjected to the same sampling rules that apply to signal processing. The sampling theorem says that if we sample a band-limited, periodic signal at a rate that is at least twice its band-limiting frequency, then we can exactly reconstruct the original signal. Sampling at a rate that is at least twice the band limiting frequency is referred to as the Nyquist criteria. Violating the Nyquist criteria will result in a digital signal that is distorted, that can not be corrected within the computer, and that is undetectable. These effects are called aliasing, and they are caused because of frequency overlapping since the high frequencies are masquerading as low frequencies. In order to avoid the aliasing effect, the signal must be band limited prior to sampling; in signal processing it is done by using low pass filters (LPF), which limit the frequency range and avoid the high frequencies in the original signal. Translating this sampling theory to digital images, we should think about how to band limit the images before subsampling. If the subsampling is done without limiting the frequency range, it is called decimation. In decimation we simply select the pixels,
in a uniform manner, from the source image at a rate that is needed to produce the desired resolution. This simple decimation will produce an image that contains artifacts, which the most common manifestation is the “blocky” appearance in the subsampled image. The artifacts are a result of aliasing high spatial frequencies into the lower spatial frequency of the subsampled image.

In order to avoid the aliasing effect, we have to spatially filter the image. By filtering the images, we also reduce the noise and smooth them, avoiding sharp intensity level transitions (due to high frequencies). Spatial filters in image processing are typically carried out as convolution kernels. A convolution kernel is a rectangular (usually square) window that is passed over the source image. The window contains weights that are multiplied by the pixel value of the image which is lying on. These values are then summed and divided by the sum of the weights. The resultant value replaces the pixel value over which the window is centered. This process is repeated for every pixel in the source image. The type (low pass, high pass, etc.) and band limits of the filter are determined by the sign and magnitude of the individual weights. After the application of the filter, we have an image with the same size as the original but with new spatial frequency characteristics determined by the filter. It is desirable to have a LPF with a very sharp cut-off frequency for the anti-aliasing effect. For instance, a $8 \times 8$ kernel will have a sharper cut-off than a $4 \times 4$ kernel, than a $2 \times 2$ kernel and so on. So the sharpness of the cut-off of a convolution filter is proportional to the size of the kernel (order of the filter). Because of that, the number of coefficients of the filter has to be carefully selected, and moreover, a trade-off exists: sharper filters require more processing operations and hence a longer processing time. The experiments related with the influence of the filter size are presented in section 4.5.

Another important parameter of the filter is the window function on which the design is based. In this project we use a 2D Hamming window, which coefficients are computed from the following equation, taking into consideration that the filter is constructed by applying this definition in both $x$-direction and $y$-direction (so it is a symmetric LPF) (22; 23):

$$w(n) = 0.53836 - 0.46164 \cos\left(\frac{2\pi n}{N - 1}\right),$$

where $w(n)$ are the value of the coefficients of the filter, $n$ is the position of the coefficient within the filter, and $N$ is number of coefficients.

Being the order of the filter the number of coefficients plus one, we consider the following filters: order 4 ($LPF3$), order 6 ($LPF5$), order 9 ($LPF8$) and order 12 ($LPF11$). Figure 3.3 shows a 3-D representation of the Hamming window for each order using a cut-off frequency of $0.25\pi$ rad/sample (normalized frequency).

As we see in the representation of the filters, for a higher order, we have a sharper cut-off frequency. The kernel is shifted all over the image to get the new value of the pixels, and for Hamming windows, the pixel where the kernel is located on, has a major importance (higher coefficient) when calculating its new value.

### 4.2 Interpolation and Subsampling

Once the anti-aliasing filter has been applied, the next step is to subsample it. An interpolation might be necessary depending on the desired final size of the images, as much for enlarging an image as for reducing it. On the one hand, when enlarging images new pixels are created, so by interpolation is possible to estimate their value, taking into consideration the value of the neighbor pixels. On the other hand, downsizing is performed by replacement of a group of pixel
values by one arbitrarily chosen pixel value from within this group. The first point is that the selected pixel must be representative of its surroundings. The second point is that sometimes it is not possible to select a pixel from a group because of the desired final size of the image. For instance, if we have an image with $15 \times 15$ pixels and we downsize the image by a factor of $\frac{1}{3}$, we select one pixel out of 3 (factor $F_{0.333}$), and the selected samples will coincide with the samples that already exist: no interpolation is needed in this case (see Fig. 3.4a). Then, the final resolution would be $5 \times 5$ pixels. But when we try to resize an image to $\frac{5}{6}$ for example, we cannot select 5 pixels out of 6 directly, because $\frac{5}{6}$ is not multiple of $15 \times 15$ pixels (factor $F_{0.833}$) (see Fig. 3.4b). The lower the factor, the smaller the resolution and the shorter the processing time. When the factor is higher than 1, the resulting image has higher resolution than the original (enlargement). So the factor is again another parameter to consider for the experiments (see section 4.4).

Three main interpolation methods are used in digital image processing (24):

- Nearest Neighbor interpolation: It is the most basic and requires the shortest processing time of all the interpolation algorithms because it only considers one pixel. This method fits a piecewise constant surface through the data values. The value of an interpolated point is the value of the nearest point. This has the effect of simply making each pixel bigger.

- Bilinear interpolation: This method fits a bilinear surface through existing data points. It considers the closest $2 \times 2$ neighborhood pixels. It then takes a weighted average of these 4 pixels to arrive at its final interpolated value.

- Bicubic interpolation: It goes one step beyond bilinear, that is, it fits a bicubic surface by considering the closest $4 \times 4$ neighborhood of pixels—for a total of 16 pixels. Since these
Figure 3.4: The solid lines represent the pixels of the filtered (LPF) image. The dotted lines indicate new pixels with a new value, due to the interpolation. The values between brackets shows the estimated pixels and their intensity value. Taking the new samples, we obtain the subsampled version of the original one. In (a) no new pixels are estimated since the scaling factor is multiple of the original image: $F_{0.333}$. In (b), new pixels are created in order to get a new image with a factor $F_{0.833}$.

are at various distances from the central pixel, closer pixels are given a higher weighting in the calculation.

The selection of the interpolation method is very important, as it determines the quality of the images. Because of the results explained in section 4, the interpolation used when subsampling is bilinear. Bilinear interpolation is performed by making three linear interpolations: two in one direction and the other one in the other direction (see Fig. 5.1 in section 3 of the appendix).

4.3 Contrast-limited Adaptive Histogram Equalization (CLAHE)

The information of this section has been adapted from (19; 20).

The histogram in the context of image processing is the operation by which the occurrences of each intensity value in the image is shown. Normally, the histogram is a graph showing the number of pixels in an image at each different intensity value found in that image. For an 8-bit grayscale image there are 256 different possible intensities, and so the histogram will graphically display 256 numbers showing the distribution of pixels among those grayscale values.

The usual histogram equalization is the technique by which the dynamic range of the histogram of an image is increased. Histogram equalization assigns the intensity values of pixels in the input image in such a way that the output image contains a uniform distribution of inten-
sities. This technique can be used on a whole image or just on a part of an image, and its aim is to redistribute intensity distributions. If the histogram of any image has many peaks and valleys, it will still have peaks and valley after equalization, but peaks and valley will be shifted. In histogram equalization, each pixel gets a new intensity value based on the its previous intensity level. See Fig. 3.5 as an example.

Figure 3.5: (a) Typical capillary image; (b) Histogram of image (a); (c) Equalized capillary image; (d) Histogram of image (c); (e) Equalized capillary image using CLAHE technique; (f) Histogram of image (e).

The histogram equalization automatically determines a transformation function that produces an output image that has a uniform histogram. But sometimes, because of the content of the images, the uniform histogram of the output image is not the best approach. In our case, the images have more high-value pixels (background) than black pixels (blood vessels). Figure 3.5b shows that the mode is around 180, and the values are not distributed along all the possible range. In order to get a better contrast maintaining the properties of the image, we use what is called Contrast-limited adaptive histogram equalization (CLAHE), which take advantage of an
CLAHE is a special class of histogram equalization. CLAHE maximizes the contrast of an image by adaptively enhancing the contrast of each pixel relative to its local neighborhood. For adaptive histogram equalization to enhance local contrast, histograms are calculated for small regional areas of pixels called tiles, producing local histograms. Each tile’s histogram is equalized from the (commonly) narrow range of intensity values to the full available range, so that the histogram of the output region matches the specified histogram. The neighboring tiles are then combined using bilinear interpolation to eliminate artificially induced boundaries. The contrast, especially in homogeneous areas, can be limited to avoid amplifying any noise that might be present in the image.

In order to understand how CLAHE works, we need first to understand the normal histogram equalization. Consider the variable \( r \) represent the gray levels of the image to be enhanced, assuming that it has been normalized to the interval \([0, 1]\) and being \( r = 0 \) white and \( r = 1 \) black. It is possible to define a transformation of the form

\[
s = T(r) \quad 0 \leq r \leq 1
\]

that produces a level \( s \) for every pixel with value \( r \) in the original image. Then, \( T(r) \) satisfies

\[
0 \leq T(r) \leq 1 \quad 0 \leq r \leq 1.
\]

We denote the back transformation as

\[
r = T^{-1}(s) \quad 0 \leq s \leq 1.
\]

The gray levels in an image can be considered as random variables in the range \([0, 1]\). So it is possible to describe it with its probability density function (PDF). Then, we define \( p_r(r) \) and \( p_s(s) \) as the probability density functions of random variables \( r \) and \( s \), respectively. From the probability theory, if \( p_r(r) \) and \( T(r) \) are known and \( T^{-1}(s) \) satisfies that is single-valued and monotonically increasing in the interval \( 0 \leq r \leq 1 \), then \( p_s(s) \) can be written as:

\[
p_s(s) = p_r(r) \left| \frac{dr}{ds} \right|.
\]

Thus, the probability density function depends on the PDF of the input image and on the transformation (3.1).

An important transformation in digital image processing is:

\[
s = T(r) = \int_{0}^{r} p_r(w)dw.
\]

The right side of (3.4) is the cumulative distribution function (CDF) of random variable \( r \). So given the transformation \( T(r) \), we find \( p_s(s) \) by applying (3.3). We apply that the derivative of a definite integral with respect to its upper limit is the integrand evaluated at that limit:

\[
\frac{ds}{dr} = \frac{dT(r)}{dr} = \frac{d}{dr} \left[ \int_{0}^{r} p_r(w)dw \right] = p_r(r).
\]

Using this result and (3.3), we obtain:

\[
p_s(s) = p_r(r) \left| \frac{dr}{ds} \right| = p_r(r) \left| \frac{1}{p_r(r)} \right| = 1 \quad 0 \leq s \leq 1.
\]
From (3.6) we know that \( p_s(s) \) is a **uniform** probability density function since all the possible values of \( s \) have the same probability. So we have demonstrated that using the transformation function described in (3.4) we obtain a histogram that tends to have a **uniform** distribution of the values. (See Fig. 3.5c and Fig. 3.5d)

But CLAHE goes one step further. We have to consider now the gray levels \( r \) and \( z \), and \( p_r(r) \) and \( p_z(z) \) their corresponding probability functions. The function \( r \) still denotes the gray levels of the input image, and \( z \) denotes the values of the output image. As we know from the histogram equalization, we can obtain \( p_r(r) \) from the input image, while \( p_z(z) \) is the specified density function that we want for the output image.

Thus, we have the expression

\[
s = T(r) = \int_0^r p_r(w)dw. \quad (3.7)
\]

Then we define a new variable \( z \) in such a way that

\[
G(z) = \int_0^z p_z(t)dt = s. \quad (3.8)
\]

It is easy to deduce from (3.7) and (3.8) that

\[
G(z) = T(r),
\]

and then,

\[
z = G^{-1}(s) = G^{-1}[T(r)]. \quad (3.9)
\]

The transformation function \( T(r) \) can be obtained from (3.7) once \( p_r(r) \) has been estimated from the original image. \( G(z) \) can be obtained using (3.8) because \( p_z(z) \) is given. Then, the procedure to obtain the final histogram of an input image is:

1. Obtain \( T(r) \) from (3.7).
2. Obtain \( G(z) \) from (3.8).
3. Obtain the inverse transformation function \( G^{-1} \).
4. Obtain the output image by applying (3.9) to all pixels in the input image. The value of the pixels of the resulting image (\( z \)) will have a PDF \( p_z(z) \).

Sometimes it is not possible to get the analytical expressions for \( T(r) \) and \( G^{-1} \). In the discrete case, the problem is solved with the disadvantage that only an approach of the desired histogram is obtained.

The discrete version of (3.7) is

\[
s_k = T(r_k) = \sum_{j=0}^{k} p_r(r_j) = \sum_{j=0}^{k} \frac{n_j}{n} \quad k = 0, 1, 2, \ldots, L - 1 \quad (3.10)
\]

where \( n \) is the total number of pixels in the image, \( n_j \) is the number of pixels with gray level \( r_j \), and \( L \) is the number of discrete gray levels. Thus, (3.8) can be written as:
\[ v_k = G(z_k) = \sum_{i=0}^{k} p_z(z_i) = s_k \quad k = 0, 1, 2, \ldots, L - 1. \quad (3.11) \]

Similarly, we write (3.9) as:

\[ z_k = G^{-1}(s_k) = G^{-1}[T(r_k)] \quad k = 0, 1, 2, \ldots, L - 1. \quad (3.12) \]

The implementation in the discrete case is exactly the same that the continuous one. Figure 3.6 shows a graphical explanation of the equalization process using CLAHE, and Fig. 3.5e and Fig. 3.5f show the resulting image and its histogram.

Some parameters that have to be fixed in order to get an optimal equalization are:

**Number of tiles** The number of parts in which the image is divided in order to perform the sectioned equalization (the same for x and y-directions). Default: 8.

**Range** The possible value of the output image data. Default: 256 grayscale intensity levels (8 bits).

**Distribution** Specification of the desired histogram shape for the image tiles. Default: exponential

Attending to the appearance of the images, the selection of the distribution parameter is clear. As we explained in section 3, they mostly have a gray background with some black dots on it. So the final distribution is supposed to contain lots of bright dots and few black dots. In this case, the histogram that better suits this consideration is the exponential distribution.
5 Template Matching and Feature Extraction

Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. When performing analysis of complex data one of the major problems comes from the number of variables involved. For this reason, is crucial to simplify all useful information contained in the images into some variables that let us work faster, more efficiently and easily. The approach that offers this compression of information is template matching. In digital image processing, template matching consists on finding small parts of an image which match a template image (in our particular case would be to find the capillary structures of the template within the acquired images). The mathematic procedure that performs this search is the correlation. Specifically, we use the Normalized 2D cross-correlation (N2DCC) algorithm, and the reasons and a mathematical approach are explained in the next section.

5.1 Normalized 2D Cross-correlation

A common problem in digital image processing is to localize a given pattern in an image, regardless if the conditions change from one realization to an other. It means that we can distinguish two different situations when authenticating: (1) The acquired image does not contain the pattern; then, the used technique should ensure that there is no possible match, or at least, that the possible match is not reliable enough (2) The image contains the pattern, but it has a different position and different light conditions; then, a match is completely necessary, but it should be unsensitive to the intensity level of the pixels.

N2DCC has these properties, and moreover, the algorithm is not high time consuming. In order to match the pattern in the image, the position of the given pattern is determined by image pixel comparisons with the given template that contains the desired pattern. At the end, we get an image of correlation where the pixels with the highest value represent the position where the template better suits, and it means that the searched structure is found in that position.

Let \( f(x, y) \) denote the intensity value of the image \( f \) of the size \( M_x \times M_y \) at the point \( (x, y) \), where \( x \in [0, \ldots, M_x - 1] \) and \( y \in [0, \ldots, M_y - 1] \). The pattern is represented by a given template \( t \) of the size \( N_x \times N_y \). The cross-correlation has its origin in the measure of the Euclidean distance:

\[
d^2_{f,t}(u, v) = \sum_{x,y} [f(x,y) - t(x-u, y-v)]^2, \tag{3.13}
\]

where the sum is over \( x, y \) under the window containing the template \( t \) positioned at \( (u, v) \). If we expand (3.13), we get

\[
d^2_{f,t}(u, v) = \sum_{x,y} [f^2(x,y) - 2f(x,y)t(x-u, y-v) + t^2(x-u, y-v)],
\]

where the term \( \sum t^2(x-u, y-v) \) is constant, since the template is always the same. If the term \( \sum f^2(x, y) \) is approximately constant, we can consider then that the remaining cross-correlation term is

\[
\gamma(u, v) = \sum_{x,y} f(x,y)t(x-u, y-v), \tag{3.14}
\]

which is a measure of how similar are the image and the template. Some disadvantages may cause to look for an improvement over this method.
• If the image energy \( \sum f^2(x, y) \) varies with position, the match using (3.13) can fail. For example, if we correlate the template with an exacting matching region in the image, the matching might be lower than the correlation between the same template and a bright part of the image.

• The size of \( \gamma(u,v) \) depends on the size of the template.

• Equation (3.13) varies with changes in image amplitude (lighting conditions).

A common way to calculate the position \((u_{pos}, v_{pos})\) of the pattern in the image \(f\), solving the previous problems, is to normalize (3.14), evaluating the N2DCC value at each point \((u,v)\) for \(f\) and the template \(t\) which has been shifted by \(u\) steps in the \(x\)-direction and by \(v\) steps in the \(y\)-direction. Equation (3.15) gives the basic definition for the N2DCC:

\[
\gamma(u,v) = \frac{\sum_{x,y} [f(x,y) - \bar{f}(u,v)][t(x-u,y-v) - \bar{t}]^2}{\sqrt{\sum_{x,y} [f(x,y) - \bar{f}(u,v)]^2 \sum_{x,y} [t(x-u,y-v) - \bar{t}]^2}}.
\]  

(3.15)

In (3.15) \(\bar{f}(u,v)\) denotes the mean value of \(f(x,y)\) within the area of the template \(t\) shifted to \((u,v)\), which is calculated by:

\[
\bar{f}(u,v) = \frac{1}{N_x N_y} \sum_{x=u}^{u+N_x-1} \sum_{y=v}^{v+N_y-1} f(x,y).
\]

With similar notation, we express \(\bar{t}\) as follows:

\[
\bar{t} = \frac{1}{N_x N_y} \sum_{x=u}^{u+N_x-1} \sum_{y=v}^{v+N_y-1} t(x-u,y-v).
\]

The denominator in (3.15) is the variance of the zero mean image function \(f(x,y) - \bar{f}(u,v)\) and the shifted zero mean template function \(t(x-u,y-v) - \bar{t}\). Due to this normalization \(\gamma(u,v)\) is independent to changes in brightness or contrast of the image. The desired position \((u_{pos}, v_{pos})\) of the pattern which is represented by \(\gamma(u,v)\) is equivalent to the position \((v_{max}, v_{max})\) of the maximum value \(\gamma_{max}\) of \(\gamma(u,v)\), where \(\gamma(u,v)\) belongs to \([-1, 1]\). When \(\gamma(u,v)\) is equal to 1, it means that the pattern in the image is exactly the same as the one in the template. On the contrary, when \(\gamma(u,v)\) is equal to -1, the pattern is exactly equal to the template, with inverse values. The algorithm used to evaluate the N2DCC computes the correlation for every position of the template over the image, regardless if it is within the image or not (if the template does not fit completely within the image, it is padded with zeros). Thus, the resulting image is of size \((M_x + N_x - 1) \times (M_y + N_y - 1)\). In our study, we are not interested in the complete image of correlation, so we only use the part of it that corresponds to the correlation positions where the template totally fits within the image; the resulting image is actually, \(M_x \times M_y\).

An example of the N2DCC result is illustrated in Fig. 3.7: an image containing the template that will be correlated. In the resulting image, is possible to see that bright spots do not hide the maximum of the correlation \((u_{max},v_{max})\).

Considering the above definitions, all N2DCCs are computed:

\[
C_{kr,sp} \quad k = 0,1, \ldots, K \quad r = 0,1, \ldots, R \quad s = 0,1, \ldots, K \quad p = 0,1, \ldots, P
\]
where $k$ indicates the subject from which the template comes from ($K$ subjects), $r$ the template belonging to this subject that will be used for the correlation ($R$ templates per subject), $s$ the subject from which the image for correlating comes from ($K$ subjects), $p$ the image belonging to this subject that will be used for the correlation ($P$ images per subject) and $C$ the N2DCC.

For instance, if we have $K = 2$ subjects, $P = 2$ images per subject and considering all the images belonging to $Q3 (R = P)$, we compute all possible cross-correlations: $C_{11,11}, C_{11,12}, C_{11,21}, C_{11,22}, C_{12,11}, C_{12,12}, C_{12,21}, C_{12,22}, C_{21,11}, C_{21,12}, C_{21,21}, C_{21,22}, C_{22,11}, C_{22,12}, C_{22,21}, C_{22,22}$.

More details about the N2DCC can be found in (25; 26; 27).

## 5.2 Thresholding

Suppose that we have the grayscale image depicted in Fig. 3.7c, where the value of its pixels are defined by $\gamma(u, v)$. The aim of computing the N2DCC is to find a match between the image and the template; if there is a match (from what it is called *intra-scores* or *client-scores*), the image should have some pixels with high intensity level, so the image should ideally have an absolute peak. If not (*inter-scores* or *impostor-scores*), the value of its pixels should be approximately equal. Figure 3.8 shows two different surfaces of two correlated images. The first one does not have a match, therefore there is not a high peak. The second one corresponds to a matching
image, and therefore a peak is easily identified. Thus, what defines if there is a match is the value of the highest pixel in the image obtained after applying the N2DCC algorithm. One obvious way to extract the required information, is to select a value $T$ (threshold). Then, if any point $\gamma(u, v)$ is higher than $T$, the image of correlation comes from the correlation between an image and a template extracted from the same subject (the searched pattern is contained in the image—Fig. 3.8b). On the other hand, if there are no values exceeding the threshold $T$, there is no match (Fig. 3.8a). That is

\[
\text{If } \exists (u, v) / \gamma(u, v) > T \Rightarrow \exists \text{ match at } (u, v)
\]

Otherwise $\nexists$ match.

The value of the highest pixel of the correlation $\gamma_{max}$ is the extracted feature from which the final decision is made. From now on, when we talk about threshold or decision threshold we refers to this $T$, and the value of $\gamma_{max}$ will be quoted as score. It has to be considered that varying the value $T$, an image can be accepted as a match or not. Due to the range of $\gamma(u, v)$, $T$ can only take values in the same range.
Figure 3.8: In (a) all values $\gamma(u, v)$ are below $T$; in (b) there is at least one value exceeding $T$. 
Section 4

Data Acquisition and Experimental Results

This chapter describes the data acquisition process, the experimental results in terms of the usual biometric indicators, that is FAR, FRR and EER, and shows the influence of important factors in the biometric performance. In section 1, we explain how to characterize a biometric system, regardless of the biometric data source. At the end of section 2 and section 3 the reader will have a further knowledge about the parameters used in the experiments and their effect on the mentioned terms. Section 5 discusses the performance of the system applying the optimum configuration. All of it always supported by plots and curves to make easier the understanding.

1 Characterization of Biometric Systems

In order to evaluate the results of the experiments, we need some “tools” to compare them in an objective way. As we commented before, the criteria we use are of standard use in biometric systems. Biometric authentication can be seen as a detection task where we have to consider two types of errors:

- **false-rejection rate** (FRR), false nonmatch or miss detection, which occurs when a subject who belongs to the system is rejected by the system (henceforth these are referred to as clients).

- **false-acceptance rate** (FAR), false match or false alarm, that occurs when a subject who does not belong to the system (impostor) is accepted as if he/she was a client.

When we are interested in security, it is convenient to have low FAs, even if the FRRs increase. On the contrary, if we aim for a convenient system, the lower ratios should be the FRRs.

In a biometric verification system. The test data consists of both authorized users (clients) and non-authorized subjects (impostors) patterns. Let us first take a look at the impostor patterns example. This is depicted in Fig. 4.1a. Then, FAR is defined as

$$FAR = \frac{N_{ai}}{N_i},$$

where:

- $N_{ai}$ total number of falsely accepted impostors
- $N_i$ total number of impostors tested
According to Fig. 4.1a, varying the threshold, the FAR can also be computed as

\[ FAR(T) = \frac{\int_T^{score_{max}} I(t)dt}{\int_{score_{min}}^{score_{max}} I(t)dt} \]  

(4.1)

being

- \( I(score) \) the impostor score distribution
- \( T \) the chosen threshold
- \( score_{min}, score_{max} \) the minimum and maximum value that the N2DCC can achieve

and we get the results shown in Fig. 4.1b. For instance, if \( T \) is equal to 0, everybody is accepted, so the FAR is equal to 1. Looking at the same figure, we see that the impostor-scores have low values, since this parameter represents the resemblance between an image and the template (see section 5.1).

We now consider the client-scores. In this case, the scores are obviously higher than in the preceding case, since if there is a match, the score will be high at the position \((u, v)\) (Fig. 4.2). If a decision threshold that is too high is applied to the client-scores, some of them are falsely rejected. The fraction of the number of rejected client divided by the total number of client is the FRR. According to the FAR, its value lies in between zero and one.

\[ FRR = \frac{N_{ru}}{N_u}, \]

where:

- \( N_{ru} \) total number of falsely rejected authorized users
- \( N_u \) total number of authorized users

The FRR can be computed similarly to the FAR, i.e.:

\[ FRR(T) = \frac{\int_T^{score_{min}} C(t)dt}{\int_{score_{min}}^{score_{max}} C(t)dt} \]

where:

- \( C(score) \) is the client-score distribution

Usually, in biometrics, the performance is given in terms of the ROC (receiver- or relative-operating characteristic) curves. This plots show the probability of FA versus FR, when varying
the decision threshold \( (T) \). Although this curve is enough to characterize the behavior of a system, it is desirable to report performance using a single number. Usually, this number is the equal-error rate (EER), the point at which the FAR and FRR are equal. In a perfect biometric system, these two rates would be zero. But actually, biometric systems are not perfect, so we need to accept a trade-off between the FAR and FRR. By denying access to everybody (FRR equal to 1), the system will reach a FAR of zero. On the other hand, if everybody is accepted (FAR equal to 1), then the system will not reject any authorized subject (FRR equal to 0) (see Fig. 4.3). Clearly, biometric systems operate between these extremes. For systems that provide secure access to an area, the aim is a low FAR at the expense of a higher FRR, since the security is the most important consideration. On the other hand, for a domotic system, where the objective is to avoid irritating legitimate customers, the FRR has to be lower than the FAR.

**Figure 4.2: Client curves.**

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>score_{C_{min}}</td>
<td>FRR(T)</td>
</tr>
<tr>
<td>T</td>
<td>score_{C_{max}}</td>
</tr>
</tbody>
</table>

(a) Client-scores

(b) FRR curve.

**Figure 4.3: Client and impostor curves. An example of \( EER \neq 0 \) (due to distribution overlapping).**

### 2 Acquisition and Data Set

The captured images are the basic information in the database, so in order to run different experiments and get results, we enroll every subject following these criteria:

- We take five images of the ROI of the same finger (five with green illumination and five with red illumination).
Some subjects have more than one finger enrolled.

Every time that someone has to be authenticated, the system takes a picture of the ROI, and by comparing with the database the decision is taken. But the image stored in the database is not exactly the same that has been taken for authentication; as translations, changes in light intensity, etc. may occur, the current acquisition setting is built in such a way so as to avoid rotations as much as possible. Small rotations are neglected in this study. We simulate these changes by taking the pictures proceeding as follows, and repeating five times for each subject:

- The subject puts his/her finger on glass plate smeared with oil.
- The experimenter localizes the ROI (the same in each acquired picture from the same subject).
- An image of the ROI is taken. The subject cooperates in this process.
- The subject removes the finger from the plate.

Thus, we get five images from each subject that must be qualified using the QbC method (section 3.2). The details of the database are: 60 fingers (5 images per subject), 236 images qualified as Q3, 46 images qualified as Q2 and 23 images qualified as Q1. In total, 305 images.

Fingerprint formation is similar to the growth of capillaries and blood vessels in angiogenesis, so we can consider some principles when we talk about capillary-pattern recognition. Due to the creation process of the fetus, the cells of the fingerprint (blood vessels) grow in a microenvironment within the uterus that is slightly different from hand to hand and finger to finger. The finer details of the fingerprints are determined by this changing microenvironment. Then, whether we talk about a single subject, his/her fingerprints are not totally random patterns either. Thus, in order to make the study more robust, we took pictures of different fingers from the same subject (for some of them). It means that the content of the database is sixty fingers, which does not mean sixty subjects (the database actually contains nineteen different subjects). Then, according to (28), if we obtain results using images of fingers that belong to the same subject (similar capillary-patterns) and we can still authenticate, then, the performance is considering a worse case than if we had used sixty fingers from sixty different subjects.

The main facts about this database are:

**Statement 1** Every image that belongs to the same subject contains approximately the same ROI for that subject.

**Statement 2** Every subject has, at least, one Q3 image.

**Statement 3** The results are based on closed-set authentication.

All these images have been taken using a resolution of 1200 × 1600 pixels and a camera frame rate of 7.5 fps. The magnification used is detailed in section 4.3. In order to build a good database for future work, we take and store pictures using both green and red illumination.

## 3 Parameter Specification

If we are interested in evaluating the performance of the system, the EER constitutes an important ratio. But this study is also considering another basic parameter in order to test the images
Table 4.1: Parameter specification. The parameters are specified using the symbol followed by the value (i.e. D60 in case of using a database with 60 fingers).

and the results: the size of the template. As it is exposed in this chapter, the EER depends to a large extent on the number of pixels that compose the template; so it is very important to understand its behavior while varying the size of the template, at the same time that some other important parameters are taken into consideration. The FAR and FRR curves depending on the threshold are also a matter of study. Even though there are many parameters that are object of study in this report, we have tested the most relevant ones in terms of EER.

Before starting with these simulations, it is convenient to relate these parameters with the block diagrams in Fig. 3.1 and Fig. 3.2, cited in section 2, because they are present in every simulation and some of them need a special attention. As it was discussed in section 3, the origin from where the template is extracted seems to be a basic point (referred to the template extraction block in the diagram). Different experiments related to this fact have been run, and they are extensively described in section 4.2. The next block is the filtering and subsampling, where parameters like the size of the LPF, subsampling factor and interpolation method have to be considered.

There are four basic parameters when we talk about CLAHE (section 4.3). The number of tiles, the output range, the output desired distribution and the alpha parameter. Due to the large number of parameters to fix, we have to establish a default configuration which we can take as a reference. Making variations on some of these variables while leaving constant the rest, we manage to see how they behave and the relation among them. The default values for these parameters were set by observation of their effect on the performance, and they can be found in Table 3.

For every experiment we indicate the configuration of the values for an experiment detailing the parameter specification (Ps). Henceforth, whenever a parameter is not specified in the Ps,
it is considered to be its default value.

4 Experimental Results

4.1 Detailed Procedure Explanation

The results we present in this section are evaluated following the explained ratios, such as FAR, FRR and EER. Processing the images using the N2DCC permits to extract some features that represent the relation between an image and a template. The way in which we combine all those images and templates and how we compute the EER, is discussed in this section.

Given a set of images (i.e. \( Q_1, Q_2 \) or \( Q_3 \)), where every subject has five images, we obtain the corresponding templates in the way section 4.2 explains. Depending on the quality of those images and following the statements (section 2), we obtain a certain number of templates (usually lower than the number of images, that is, because of quality reasons we cannot extract a template from every image). Then, we apply the N2DCC between every template and the rest of images. If the template and the image correspond to the same subject, the score we obtain is an intra-score value. Otherwise, if the image corresponds to a subject and the template to a different one, we obtain an inter-score value. Figure 4.4 shows the way in which the N2DCC between images is done.

![Figure 4.4: Correlation between templates and images. The solid lines represent the correlations which give intra-scores. The dashed lines correspond to the inter-score operations.](image)

For every available template the algorithm performs all the possible cross-correlations. As Fig. 4.4 illustrates, there are images which do not have a template, but still take part into the correlation process. The empty space between some images represents the low quality images that are not considered in the simulations.
Every time we perform all the cross-correlations using the N2DCC as we explained, considering a specific $P_s$, we obtain two vectors which contain the *intra-scores* and *inter-scores* respectively. Then, applying the proper algorithm we calculate the EER for a given threshold. When we obtain the EER for a range of thresholds, then we can plot a curve representing the EER against the $T_s$. Each run leading to an EER is referred to as simulation or experiment.

### 4.2 Template Origin

Even though at the beginning is not possible to know what are the optimum parameters of the system, we can start running some experiments to test some template properties. The obvious question that we need to solve, apart from the size of the template, is its position. The first approach, considering the number of people in the database (not too large), is to create a sub-database of templates from the acquired images. The goal of the experiment is to know if we have to develop a specific method to extract the optimum template from the image, or, on the other hand, the origin does not affect the final results.

We need to compare results from different simulations where the origin of the used templates is different. To run this experiment, we consider a small database with 26 fingers, where every one of them has 5 images of the ROI. Thus, from every image, we select, by hand, a characteristic part (an area with an important number of blood vessels) which is present in the five images that belong to the same subject. The template is a square area selected by eye inspection; so the results are totally subjective, since it is unlikely to select exactly the same portion in the five images using this method. We repeat this procedure for all subjects in the database, which means that we get a subdatabase that contains only templates. By selecting by hand the part of the image that we think is the most characteristic for that user, we are indirectly taking into account the frequency content of the image (although we do not compute the 2-D Fourier Transform of the image). We could say that we are selecting the part of the image having high frequency content, because a large number of black dots on a gray background means many sharp transitions. Table 4.2 indicates the EER using the detailed ($P_s$).

<table>
<thead>
<tr>
<th>$P_s$</th>
<th>$T_s$</th>
<th>$M$</th>
<th>$Q_1$</th>
<th>$Q_1$</th>
<th>Method</th>
<th>EER (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ts200</td>
<td>M0.8x</td>
<td>TQ1</td>
<td>Q1</td>
<td>Hand-selection</td>
<td>1.73</td>
<td></td>
</tr>
<tr>
<td>Ts200</td>
<td>M0.8x</td>
<td>TQ1</td>
<td>Q1</td>
<td>Central-part</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>Ts400</td>
<td>M1.25x</td>
<td>TQ1</td>
<td>Q1</td>
<td>Hand-selection</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>Ts400</td>
<td>M1.25x</td>
<td>TQ1</td>
<td>Q1</td>
<td>Central-part</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>Ts400</td>
<td>M2.0x</td>
<td>TQ1</td>
<td>Q1</td>
<td>Hand-selection</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>Ts400</td>
<td>M2.0x</td>
<td>TQ1</td>
<td>Q1</td>
<td>Central-part</td>
<td>2.45</td>
<td></td>
</tr>
</tbody>
</table>

Maintaining the template size and magnification constants, another possible extraction is considered in this section to compare the results; the main advantage of the new extraction is that it is not so time consuming than the manual extraction, and it could be easily applied to a large database (it is an automatic method). We are talking about extracting the central part of the image as a template, assuming that for every subject there are no large translations in the images. In this case, the central point of the template coincides with the central point of the image. Table 4.2 also shows the results using automatic extraction. The templates, are
extracted (as an exceptional case) from Q1, Q2 and Q3, so we are using all images and all possible templates, regardless of the quality.

The Table 4.2 shows that we are not comparing equivalent simulations, because for \( M_{0.8x} \) we are using a template of \( Ts_{200} \) instead of using \( Ts_{400} \). The aim of these trials is to know from where we should obtain the template by comparing equivalent simulations, and it makes not much sense to compare those which do not have almost the same \( Ps \). Thus, we should extract conclusions from every pair of simulations. For \( M_{1.25x} \) and \( M_{2.0x} \), it appears to be better to take into consideration the content of the image and take those parts where the number of blood vessels are higher. In these two cases, the difference between their EERs is important, so the source of the template has a remarkable effect on the results. For higher magnifications, this difference is larger. That is because if we take a part of the image randomly, it is more probable that the template contains very few blood vessels, and this effect increases with the magnification.

For \( M_{0.8x} \), the results indicate that it is better to take the central part of the image. But we have to recall that the hand-selection is obtained in a subjective way. Then, it could happen that the hand-selected part was not the optimum one, and the central part is actually better in terms of EER.

We conclude that the content of the images is important. However, this is not considered in this study as we lack in an automatic and fast procedure. The default extraction that we consider is the central part, which can be performed automatically. Some ideas about optimum-automatic extraction of the template is found in section 2. Then, we should remark that the obtained final results could be improved using the proper extracting method.

4.3 Magnification

The microscope permits to take pictures using different magnifications, but it is important to know that there is a trade-off between the magnification and the sensitivity to translations. With low magnifications (i.e. \( M_{0.8x} \)) it is relatively easy to find the ROI in every subject (regardless of translations) but the vessels are represented with few pixels. When we consider higher magnifications (i.e. \( 3.6x \)), every capillary vessel is composed of more than 30 pixels (see Table 2.1), but it is rather difficult to find the ROI. For instance, using \( M_{0.8x} \), if the finger position varies of about 0.5 mm with respect the template stored in the database, the new image will have a translation of 140 pixels; this is relatively small compared to the size of the image (1200 \( \times \) 1600 pixels), and the authentication is still possible. But in case of using \( M_{3.6x} \) with the same variation, the translation in the image is about 640 pixels, which covers more than half of the image size. In this case, the performance would be considerably degraded, since the captured image could not contain the pattern which allows the verification.

With the results presented in this section, we determine which is the best magnification to use in the following experiments. There are several magnifications with which we could work, but since some of them magnify almost by the same factor, we selected three representative objectives to perform results. Thus, Fig. 4.5 illustrates the EER versus \( Ts \), using three different microscope objectives: \( M_{0.8x} \), \( M_{1.25x} \) and \( M_{2.0x} \). Then, we have a curve for a low magnification, one for middle-magnification and one curve for a high magnification objective.

Looking at the results, we can say that the curves do not behave as we expected. Contrary to the theory, \( M_{2.0x} \) seems to be the worst option in terms of EER. It is caused because even using big sizes for the templates, this magnification is too high to capture enough blood vessels for the subject to authenticate. Although \( M_{2.0x} \) offers high resolution for the capillaries, the total size of the captured image is sufficient only to contain few of them.
The election then is between $M_{0.8x}$ and $M_{1.25x}$. We have to consider the fact that the number of fingers in this experiment is low, so the results might not totally represent the real behavior of the system. For this result, the quality of the image was not taken into account; we are not looking for the best performance, instead we are trying to figure out the most convenient magnification. We see from the graph that between $T_s_{250}$ and $T_s_{400}$, both plots have a similar EER around 0.5%. So, considering the trade-off and the size of the database, the best conclusion is to acquire the images using $M_{1.25x}$, because despite of the higher EER in this concrete experiment, the capillary vessels have better resolution, the information in the pictures seems to be enough for achieving low EERs, and all without making too difficult to localize the same ROI for every subject.

The general behavior of the three curves shows that for small masks the EER is high, then it decreases up to an optimal value and then increases again (like a $∪$-shape). This shape appears in all graphs where the EER is plotted versus $T_s$ and it is reported and studied in section 5.

4.4 Subsampling Factor

The images that the system provides have high resolution ($1200 \times 1600$ pixels). On the one hand, a high resolution permits to have a better representation of the capillary structures. On the other hand, images with such a high resolution require large memory space storage. Considering that the images are first equalized and then submitted to a correlation process, which means that the system has to compute $12N^2\log_2 M$ multiplications and $18M^2\log_2 M$ additions/subtractions only for the N2DCC (being $N_x$ equal to $N_y$ equal to $N$ and $M_x$ equal to $M_y$ equal to $M$, and without considering fast implementations of the N2DCC algorithm), it is interesting to reduce the computation time. One way can be to reduce the size of the image and the templates, but as it is explained in this chapter (section 5), it is not convenient to reduce (crop) neither the size of the image nor the template, since then we are losing information. Subsampling is a possible solution, since by applying the proper pre-processing the relevant information is adequately preserved.

Thus, we have to study the effect of subsampling the images and templates, because in any
case (cropping or subsampling) we expect to obtain worse results than using the images with their original size. Remind that before subsampling, it is always necessary to filter the image with a LPF to avoid aliasing, and in some cases, it is required interpolate (see section 4.2). In this section we present the results considering five different subsampling factors \( F \). The goal is to find the optimum factor that can be used without degrading the performance in terms of EER. Figure 4.6 shows five curves, corresponding to the five factors, plotting EER versus the different template sizes. The rest of the parameters of the system remain constant with the default \( P_s \).

![Figure 4.6: EER performance considering different subsampling factors. Default \( P_s \).](image)

Looking at the results, we see that every factor has an optimum value between \( T_s300 \) and \( T_s400 \). All of them have the same behavior, which means that for low values of \( T_s \), for instance \( T_s100 \), the EER is around 18%. The explanation is that the template is so small that it does not contain enough information, so the system are not able to authenticate correctly the clients (high FRR) and it accepts impostors (high FAR). As \( T_s \) is increased, the EER also decreases. After some variations, every curve reaches its minimum value, being EER equal to 0\%\(^1\) for \( F0.25 \) and \( F0.5 \).

It is normal that for high reducing factors like \( F0.1 \), the performance is worse, since we are reducing the original image ten times. In theory, the better performance should be achieved by using the original images (\( F1.0 \)), and as we reduce the factor, the EER should also increase. But Fig. 4.6 shows that this is not the case. For \( F1.0 \) we have the worse performance, which means that the filtering process applied to \( F0.1, F0.25, F0.5 \) and \( F0.75 \) before downsampling has a positive effect on the EER, getting a better performance than \( F1.0 \) and \( F0.75 \) in terms of EER. The study of the filter is done in section 4.5.

As section 4.2 explains, not always an interpolation is needed (i.e. \( F0.1, F0.25 \) and \( F0.5 \)). But \( F0.75 \) needs it in order to subsample the images. As it is expected, for \( F0.75 \) the EER is almost always lower than the other curves, except for the fact that it never reaches EER equal to 0\% (as the others do). Since the LPF applied is the same for all of them (LPF11), the reason of this “high” EER for the case of \( F0.75 \) is due to the interpolation, which is introducing some errors that degrade the EER.

As a first conclusion we could say that the proper filtering is a powerful tool for improving the results that we could get by directly using the original images. The proof is that for \( F1.0 \) we

\(^1\)This value is due to the reduced number of people in the database.
have the highest EER. The second conclusion is that we should avoid as much as possible subsampling factors which need a subsequent interpolation, since we see for $F0.75$ that although the performance is better than $F1.0$ (because of the LPF), is slightly worse than the rest (its minimum is a bit higher). Thus, it is better to use a higher subsampling factor which does not require interpolation, rather than using a low reducing factor which needs it. Even though the EER seems to be better for high factors, there is not a remarkable minimum EER (the minimum for $F0.75$ is at $Ts300-330$ with an EER of 0.0023%), whereas using the other factors, we see an EER of 0% at $Ts400$ for $F0.25$ and $F0.5$. Moreover, the computational time for processing images at $F0.75$ is much higher than using lower factors, because the size of the images is bigger (N2DCC depends on $M$ and $N$) and it needs interpolation. As a first approach, the results for $F0.75$ are obtained using a bilinear interpolation, since it requires less computation time than bicubic interpolation.

Hence, the factors which seem optimum for our images are $F0.25$ and $F0.5$, because we can reach a low EER reducing the size of the image two or four times. In both cases we use a LPF, no interpolation and the EER has almost the same behavior, with a minimum at $Ts400$ with EER equal to 0%. So, considering that it is almost the same to downsize the image four times or two times, the first one requires less computation time. Therefore, the optimum factor is $F0.25$. The rest of the results achieved are obtained using this factor, which does not need any interpolation technique.

### 4.5 LPF

The filter applied to the images has an important influence on the performance that has to be understood in order to find the optimal parameters for the matching process. In the previous section, we realized that it is completely necessary to filter the images even if they do not need to be interpolated, because a LPF smooths the images and filters some noise. The goal of this section is to optimize the size of the LPF.

By default, for most of the experiments we use a LPF11, because we need to define, as a first approach, the kernel of the filter (section 4.1. But as with the rest of the parameters, it is necessary to find the size that better suits our needs. The next simulation shows the performance of the system when considering that the images have been pre-processed using a Hamming window of order 4, 6, 9 and 12 as a pre-sub-sampling LPF (see Fig. 4.7).

All curves in Fig. 4.7 have the same U-shape, with an optimal value for the EER at $Ts400$. The first impression is that at $Ts330$, all curves reach the same value (EER equal to 0.00080%), which is low considering that we are taking Q2 images. After this point, the performance is different for every curve, from an EER equal to 0.0057% for LPF11 down to an EER equal to 0.0023% for LPF3 at $Ts350$. Actually, this is not the expected trend, since for a sharper cut-off filters, the EER is higher. However, the difference between them is rather insignificant (remind that the graph has a log $y$-axis), and it might be caused by the algorithm itself. Where the size of the filter has a noticeable effect is at $Ts400$, where for LPF5, LPF8 and LPF11 curves, the EER is 0%. Hence, for larger templates sizes, is required to have a larger number of filter coefficients, since it causes a large improvement in the authentication. When the template is very large (i.e. $Ts500$ and $Ts600$), an edge effect (explained in section 5) appears, and the order of the filter does not matter. All curves reach approximately the same values at each template size.

As a result of this experiment we can say that although filters with more coefficients are generally better, a Hamming filter with five coefficients it is enough for a correct authentication. In fact, using a LPF5 there are no errors in the decision when we use $Ts400$. For higher filter orders, the process is more time consuming. So, using only a 5 × 5 pixels Hamming kernel, the
aliasing is avoided, the image is smoothed and the noise reduced. Thus, the optimum order of the LPF for the final performance is 6 (LPF5).

4.6 Quality

Considering that the templates are always extracted from images of quality level Q3, we can obtain three curves that represent the relation between EER and \( T_s \) for images submitted to the QbC method (section 3.2). Figure 4.8 illustrates their behavior, using the default \( P_s \).

Figure 4.8: EER performance considering the QbC. Default \( P_s \).

From Fig. 4.8 is clear that simulating with the whole set of images (Q1), the performance in terms of EER is the worst, meaning that the EER is 0.22% for \( T_s300-Ts330 \). Although the EER is the highest for Q1, even then, the performance can be considered quite good, since the images are not controlled for quality.

When we consider only images from Q2, there is a large improvement respect to the Q1-performance. Nevertheless, comparing Q2 with Q3, the difference is minimum. Looking thor-
oughly at both curves, we can even see that for $Ts_{300}$, the EER is lower for $Q_2$ than for $Q_3$. The explanation lies on the fact that the QbC is a subjective method, made by eye inspection. From the plot, we understand that the qualification between $Q_3$ and $Q_2$ it is not so different, although the selection was made following the specified rules; in fact, the images considered as belonging to $Q_2$ could be considered as very good images ($Q_3$). The system is able to authenticate (without any error for the proper template) images containing small artifacts, low intensity level or slightly blurred blood vessels. So it makes no sense to study the rest of parameters considering the most restrictive case (which is the one that considers less number of images): from this section on, the plots show the performance for both $Q_2$ and $Q_1$.

5 Final System Performance

After defining the optimum parameters of the system, we can plot the final EER curves of the system. This section presents the best results that, by now, is possible to get, as well as a discussion about the behavior of these curves.

Along this chapter we presented, one by one, the effect of the parameters that have repercussion on the results. We fixed the part of the image that we should use for the extraction of the template, although for convenience we are using another method, we have determined the optimal subsampling and filtering parameters, the optimum magnification of the microscope and we know the set of images we have to use for obtaining reliable results. Thus, the plot of the EER versus $Ts$ (Fig. 4.9) allows us, finally, to find the best size of the template. Therefore, Fig. 4.9 shows the best performance we can get using the acquired database, making the QbC and obtaining the respective curves.

![Figure 4.9: EER performance considering the optimal configuration. $Ps$: LPF5.](image)

Again, the templates are extracted only from $Q_3$. The graph shows an excellent behavior as long as we use $Q_2$ images. In that case, the results are much better than considering all the images from the database. Thus, we can get very low EERs for $Ts_{330}$ or even 0% for $Ts_{400}$, which means that the system is able to decide, without any error, if a subject is an impostor or a client. We have to remark that even though the EER is equal to 0%, in the best case, it is due to the reduced number of fingers in the database. So, if more people is enrolled, this value is expected to increase. But as a first approach, with 60 fingers and $Q_2$), it is a promising result. In
this case, though, the number of images that are used for the simulation is 282, from where we obtain 236 templates.

Figure 4.10 illustrates the FAR and FRR curves for $T_{s400}$, where the cross point (EER) is at

$$T \in [0.386, 0.393].$$

![Figure 4.10: Performance for $P_s$: Q2.](image)

This range is quite narrow (see Fig. 4.11) yet sufficiently to get an EER equal to 0%, since there is no overlap of the distributions. In this plot we see that the inter-scores have less variance than the intra-scores, which is an advantage when we need to fix a decision threshold: usually, the tricky distribution is the one with the intra-scores. So the improvement in the Q2 case would be to try expand the range of $T$ where the EER is equal to 0%, because for a larger database, the intra-scores and inter-scores distributions might be more spread and then, it might be more difficult to find a $T$ that let us completely separate both distributions (0% of EER).

The most interesting curve in Fig. 4.9 is the one that represents $Q_1$, because in this case, we use all images for the cross-correlation process. The curve presents a $\cup$-shape, which minimum EER is 0.22% at $T_{s250}$-$T_{s300}$. As it is explained in section 1, the EER is a representative parameter of the system, but it might not indicate the best $T$ depending on the final aim (convenience or security). We should have a look at the FAR and FRR curves corresponding to $T_{s250}$-$T_{s300}$ in order to know the best threshold and the real FAR and FRR values. Figure 4.12 illustrates this fact. If we are looking for a low acceptance of impostors (Fig. 4.9), instead of selecting the $T$ at the EER point, we can decrease the FAR in favor of the FRR. If the aim is to avoid rejecting clients, the FRR can be decreased at expenses of the FAR. In the $T_{s250}$ case, although the EER is equal to 0.22%, is possible to consider to decrease the FAR till 0% loosing convenience (1.08% FRR), which means a big improvement in terms of security. On the other hand, it is not worth trying to make the system more convenient; we could have 0.88% of FAR and 0.11% of FRR, which is not a big decrease with respect the EER. If we consider the plot for $T_{s300}$, is very similar, but the security that we can achieve is much better: 0% of FAR when the FRR is 0.43% (no impostors accepted). The convenience of the system cannot be improved without making the FAR very high. So, the behavior of the system using $Q_1$ and $T_{s300}$ is better than the one using $T_{s250}$ (although they have the same EER) when the system is aimed to security applications. The rest of $T_s$ in Fig. 4.9 for $Q_1$ give higher EER values, and
studying the FAR and FRR curves, we conclude that there is no reason to use templates with sizes which are different from $T_{s330}$.

Along this chapter we have noticed that all plots have a common characteristic: a $\cup$-shape. This behavior indicates that this system, regardless of the parameter that is being considered, always has an optimum value. In general, the first part of the curves decrease with $T_s$ till reaching the optimal $T_s$ and then increasing again. The explanation of this behavior is that for small templates, there is not enough information contained in them to recognize all subjects enrolled in the system (which means an increase of the FRR), and to reject impostors (increase of the FAR); as a result, the EER is higher. The second part of the curves, that we could consider is between $T_{s300}$ and $T_{s400}$, contains the optimum values (with some ripples). Finally, when the template is larger than $T_{s400}$, the EER performance always increases. This increase is due to the edge effect. As we said in section 5.1, the size of the image of correlation is $(M_x \times M_y)$,
because the N2DCC is computed only in those positions where the template completely lays on the image (no padding is required). Using this procedure, it might happen that some templates do not match when they should. The templates stored in the database contain an image of a specific pattern of the ROI of the subject. It is supposed that every time a client has to be authenticated, approximately, the same ROI of the finger has to be imaged. This new image might suffer a translation, but in any case, the characteristic pattern should be present in the image. This is what is defined in statement 1 (section 2). But we do not consider the fact that the pattern can be located at the edge of the acquired image. In this case, N2DCC algorithm cannot match the stored pattern with the one in the image, simply because it is located extremely on the border of the image. Given the image in Fig. 4.13a, we extract a template, which contains a specific pattern. In Fig. 4.13b, we have an other image where there is a translation in the ROI, and the stored pattern is placed close to the border. Then, the matching is not possible since the template cannot be placed as the Fig. 4.13b shows.

![Figure 4.13: Edge effect.](image)

The result of the edge effect is that images belonging to clients, are considered as impostors, since the score is too low to be considered intra-score: the FRR increases. There is important information on the edges of the images that should be taken into consideration for the decision. Otherwise, the U-shape in the plots should not exist; it should be simply constant and equal to the minimum value of the performance, because if the size of the template is larger than what is needed, it should not make the result worse.

The translations in the images obtained for the authentication have an important role in this system. In order to know if this effect is actually the cause of the increasing behavior in the plots for large template sizes, we run some simulations to find it out. If there are some patterns that are located on the edges of the images (because of translations) and they cannot be matched. A
way to test it, would be to acquire larger images. Then, we would not have the problem with the N2DCC, because the pattern would be more in the center, the correlation would be possible and the FRR would decrease, making the EER decrease as well (Fig. 4.14).

Figure 4.14: Correction of the edge effect. Acquiring larger images (dashed line), the template can be placed on the searched pattern: the match is possible.

But the problem is that the system cannot acquire images with higher resolution than 1200 × 1600 pixels. Another way to test the influence of the edges, is doing the other way around: if we get smaller images by cropping them, then some patterns located at the edges would be removed, and consequently the templates will never match when they should. It means that the FRR should increase due to those cases where the translations are important, and the EER should also increase. So the next simulation shows the EER versus $T_s$ using different sizes for the original images. These images are obtained just by removing some pixels from every side (cropping method). Specifically, we consider three curves where the images are of 1200 × 1600, 1100 × 1500 (50 pixels less per side) and 1000 × 1400 pixels (100 pixels less per side). Remark that the templates are obtained in the same way as before, i.e. extracted from the central part of the images. With this simulation we are not trying to obtain results to compare with any other plot; we just want to see the effect of the fact that some important information (patterns) at the edges, increase the EER. To run the experiment, we use the default $T_s$ (see Fig. 4.15). Now we see that by making the images smaller by removing pixels from the edges, the EER increases significantly, which means that an edge effect exists. For images with the original size, the performance is very good (EER equal to 0% at $T_s$). But by removing 50 pixels from each side of the images, the EER increases up to 0.34%, maintaining (as it was expected) the $U$-behavior. When the images are 1000 × 1400 pixels, the performance decreases (EER equal to 0.92% at $T_s$). For small templates, the edge effect has no importance, because of the insufficient number of blood vessels contained in the templates; then, the EER is the same regardless of the size of the image.

But we have to check if this increase is due to FRR, that is, clients that could not be authenticated. Figure 4.16 illustrates the difference in EER considering $T_s$ for every cropped image. Clearly, the three plots represents the FAR and FRR curves for different sizes of the images after cropping them. Using original images, the EER is equal to 0.032%. When the images have 50 pixels cross on each side, the FAR curve is exactly the same (no more impostors accepted), but the FRR increases considerably, yielding an EER of 0.23%. In the third plot, the FAR is still the same, but again the FRR has increased due to the loss of information in the edges (the images are 100 pixels smaller). In that case, the EER is 0.79%, which is 3.5 times worse
than the second case, and almost 250 times worse than the original case. Thus, we demonstrate that the information contained on the edges of the image increase the EER of the system, and moreover, the number of removed pixels does not have a linear relation with the EER.

It is also interesting to try to do the same the other way around; if cropping the images, the performance is worse, we could try to make the images larger by padding them. Thus, we would let the templates correlate more positions at the edges, and then, it is supposed the EER will decrease (following the same reasoning than cropping images). Therefore, we have to pad the images before applying the N2DCC algorithm. The first approach is using a zero-padding, with which we should get a better result unless the pad with this value introduces some errors. It is proved that the resulting image after using the N2DCC between a template and an image that only contains zeros, is an other image which only contains zeros (black image). In Fig. 4.17a we can depict the aspect of the new images after adding 50 pixels per side with zero value applying CLAHE and subsampling by $F_{0.25}$. Figure 4.17b shows the outcome of the simulation.

The simulation in Fig. 4.17b shows two facts. The first one is that for large $Ts$, the performance is slightly better, which means that the zero-pad contributes to authenticate some more clients than before. Figure 4.18 shows the FAR and FRR curves for $Ts=600$, and the difference in EER comes from the FRR curve, which has lower values for that template size. It proves that padding with zeros improves the FRR, while the FAR has the same behavior in both cases. With large templates, adding zeros to the image does not affect the number of false accepted impostors, and as a result the EER improves from 0.81% (no padding) to 0.23% (zero-padding).

The second fact, and the most important one, is what happens when $Ts=400$. Although the results were expected to be better when there is a padding, for this case is much worse. While using the original images the EER is 0%, under the new approach the EER increases up to 0.0039%. The conclusion then is that for some $Ts$, the zeros are introducing some artifacts, contrary to what was expected. In fact, the FAR increases because of the zeros. Some impostors are considered as clients when we use the zero-padded images. This effect is depicted in Fig. 4.19, where the two plots represent the curves for $Ts=400$ with and without padding images. In those cases, the FRR is almost the same (there is no effect over the client decisions).

With these simulations using cropping and padded images we have explained that the U-shape in all curves is due to the pattern-information located in the edges of the images, and it
Figure 4.16: FAR and FRR curves for cropped images (Ts370). From (a) to (c) the FAR does not change its monotonically decreasing behavior, whereas the FRR becomes worse.

sometimes affects the decision over the clients, who cannot be properly authenticated.

Table 4.3: Number of wrong intra- and inter-correlations decisions for different Ts.

<table>
<thead>
<tr>
<th>Ts</th>
<th>intra-correlations</th>
<th>inter-correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ts100</td>
<td>152</td>
<td>10793</td>
</tr>
<tr>
<td>Ts200</td>
<td>4</td>
<td>316</td>
</tr>
<tr>
<td>Ts300</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Ts330</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ts350</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Ts370</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Ts400</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ts500</td>
<td>2</td>
<td>142</td>
</tr>
<tr>
<td>Ts600</td>
<td>7</td>
<td>1318</td>
</tr>
</tbody>
</table>

In all the plots, there is an unexpected behavior of the EER in the range Ts300-Ts400 (see Fig. 4.9). From Ts100 it is expected that the EER monotonically decreases till a minimum and then it increases again. But in the mentioned range, the curve takes higher values before reaching the absolute minimum. This shape is understandable if we look in detail at every point. After some simulations, we get the number of wrong decisions in which the system incurs for some of
(a) Zero padded image (50 pixels per side). \( Ps: 1300 \times 1700 \)

(b) EER performance using zero-padded images.

Figure 4.17: Typical zero-padded image and EER performance.

(a) FAR and FRR for \( Ts600 \). EER=0.81%.
\( Ps: 1200 \times 1600 \) (no padding).

(b) FAR and FRR for \( Ts600 \). EER=0.23%.
\( Ps: 1300 \times 1700 \) (zero-padding).

Figure 4.18: FAR and FRR curves for zero-padded images (\( Ts600 \)). The FAR is the same in (a) and (b). For large templates, the FRR decreases, making the EER better.
(a) FAR and FRR for Ts400. EER=0%. Ps: I1200 × 1600 (no padding).
(b) FAR and FRR for Ts400. EER=0.0039%. Ps: I1300 × 1700 (zero-padding).

Figure 4.19: FAR and FRR curves for zero-padded images (Ts400). The FRR is the same in (a) and (b), but the FAR is worse because of the padding.

the critical sizes of the templates. Table 4.3 illustrates this information. If we consider Ts100 and Ts200, the number of mistakes are quite high, and thus, the EER is also high. Basically, in these two cases, the FAs (inter-correlations) cause the low performance. For Ts300 there are four cross-correlations that should be considered as inter-score (impostor), but due to its high value, it is considered as intra: in this case it is a FA. Even then, the EER is low. When we look at Ts330, there are less wrong inter-correlation decisions than before, but a new wrong intra-correlation decision. The explanation is that from Ts300 to Ts330 we use 15 pixels more per side to make the match. Thus, the new pixels are introducing noise (in terms of FRR) and useful information (in terms of FAR), and as a result, the FA goes down, but a new FR appears. When we use 20 pixels more for the authentication (respect of Ts330), the decision over the clients is perfect, but again we have the same problem than before: even though all clients are accepted, the noise introduced for the new pixels is reducing the capacity of the system to reject impostors (three more than before). From Ts350 to Ts400, the fact of adding new pixels helps to improve the performance, reducing the number of false acceptances. The number of false rejections keeps on being constant till Ts500 and Ts600, where we can see the edge effect (increase of FRs). Although Table 4.3 shows that for large templates the FAR also increases, it is no contradictory with the explanation about the edge effect found in this section. We said that the value of the EER increases because the FRR is higher, and the FAR remains invariable. The shape of the FAR does not change, but the EER is different for each Ts because the FAR curve changes, which means that the T at the EER point is also different. Then, at every EER the number of wrong decisions can increase even though the FAR curve is the same.

6 Additional Parameters

6.1 Green and Red score Distributions

So far, we have presented the performance of the system using images acquired under green illumination. Recall that the system has a second LED, which let us image the background without the information of the blood vessels. This section considers red images in the authentication process, and discusses the corresponding performance.

First of all, in order to be able to compare the EER, we show in Fig. 4.20 the equivalent plot
to Fig. 4.6, using red illumination and the default $Ps$, except from the distribution, which is uniform. Although green and red images are similar, the latter only contains the background. In this case, the final desired distribution after applying CLAHE does not have to be exponential, since there are no black dots in the image, only bright intensity levels.

Looking at the plot, we see that red images contain important discriminative information about the subjects. The EER is, as in green images, dependant on $Ts$, but in this case there are no fluctuations around $Ts_{300}$ and $Ts_{400}$. If $Ts$ increases, EER decreases monotonically. For red images, there is not an optimal $Ts$: the bigger the template, the better the EER. Although red images also contain information about the subjects, it is not enough to reach low EER values. Hence, we can conclude that even though with the red images is possible to have an EER around 3%, the blood vessels (found in the green images) are the principal features that enable the authentication.

In order to illustrate this, it is interesting to plot the scores (maximum of the correlated images) of the green images versus the scores of the red images in the same graph. Thus, we get a distribution of intra-scores and inter-scores from which is easy to extract a conclusion. The aim is to be able to separate the two types of scores by using either a single threshold (on the green or red images) or a linear separation plane as in linear discriminant analysis (29). Figure 4.21 shows the distribution of scores using $Q2$ images.

Figure 4.21 shows that the inter-scores are more grouped than intra-scores. Considering only the threshold for red images (red score), it is not possible to make a clear division between inter and intra-scores, so the EER will be relatively high (Fig. 4.22b). On the other hand, by taking into consideration the threshold from green images, we can perfectly decide if a point belongs to an impostor or to a client (Figure 4.22a).

For the green images, the EER reaches 0%, whereas for red images the EER is 4.41%. These plots have been obtained using the default $Ps$. If we compare Fig. 4.21 and Fig. 4.22 considering the value of the thresholds at the EER point, we can see that for green images, there is a $T$ for which there is no error in the decision. For the red images, we always have errors in the decision since it is impossible to define a region with only intra-scores or only inter-scores by using one unique decision threshold. Thus, using the red images does not improve the system performance. Yet red images can help in improving the quality of its corresponding green images.

Figure 4.20: EER performance. $Ps$: uniform distribution.
Figure 4.21: Distribution of scores for red and green images. P\(s\): uniform distribution for red images.

Figure 4.22: EER performance for green and red images. Default P\(s\).

### 6.2 Left and Right Fingers

The performance of the system has always been limited to the use of fingers, regardless which finger was used for authentication. Since we do not use both hands equally, there might be some differences between the performance of the right-fingers and the left-fingers. With this purpose, we run some experiments that conclude in an interesting result.

For this simulation, we need, for each subject, to have images of the same finger from both hands. Specifically, we took pictures of the right and left index finger from 19 subjects. The configuration of the parameters is the default one, and the distribution of points is represented in Fig. 4.23.

The distribution of inter-scores and intra-scores are clearly separated, which means that for these 19 subjects, the EER is 0\% regardless of the imaged finger. The distribution of inter-scores is also more grouped than the intra-scores one. Moreover, if we study individually the intra-score points that lie at the outer part of the distribution, there is a common characteristic in some of them: they belong to subjects who are about 50 years old. So, if the database was
larger, we could say that it might exist a relation between the age of the subjects and the score of the matches.

In Fig. 4.24 there is a representation of the FAR and FRR curves for $T \approx 400$.

As we see, for the right finger, the FRR curve is constant in the interval $T \in [0.390, 0.517]$, due to only one point located at $(0.651, 0.393)$. Surprisingly, if it was not for this subject (who seems to be a special case out of the distribution of intra-scores), the range for an EER equal to zero would be much larger for the right finger than for the left one, instead of being the same. It can be possible that the performance is linked to the fact that the subject is right or left-handed. So for authentication, it would be better to use the right finger as long as the subject is right-handed. Otherwise, for left-handed subjects, the authentication will be better with the left one.
Section 5

Conclusions

This section contains the final conclusions of the thesis as well as a brief resume of the results achieved during the study. In section 2 we present the directions for further research.

1 Summary of Achievements

The study of a new biometric modality always carries on some risks from the point of view of performance; that is, the first results might encourage to start a research about a specific topic, but actually, nobody can preview if the study will succeed with useful and interesting conclusions. The present study answers some questions that came up at the very beginning about authentication and performance, which are resumed with the objectives of the thesis. The main goal was to find out the potential of this new biometric modality, and try to get reliable results which let us evaluate its possibilities in front of the current modalities as fingerprints or palm-veins.

As a first approach, the conclusion is that capillary-patterns in the DIP joint can be used for authentication, because under certain conditions (see section 4) it is possible to obtain a 0% EER, which is the ideal performance (with \( Q_2 \)-quality images). But it is necessary to keep in mind the reduced size of the database. Nevertheless, this result indicates that this modality can meet the requirements to be considered at the same level that, i.e. the fingerprint authentication. After the simulations, we realized that the difference between \( Q_2 \) and \( Q_3 \) is so small in terms of EER that it is not worth to consider only \( Q_3 \) images: we almost obtain the same result using \( Q_2 \) (is preferable to take into consideration a larger set of images - that is, \( Q_2 \) instead of \( Q_3 \)). When we obtain the EER using a larger database (\( Q_1 \) images), which means to take into consideration the whole set of images, the ratio is 0.22\% (templates only from \( Q_3 \)). This result is good enough to motivate new research on the digital image processing algorithms and putting more effort on the template selection algorithms in order to be able to use all images in a more efficient way, instead of rejecting some of them due to their quality, etc.

An important part of the thesis has been to identify the key factors influencing the performance. After different simulations and research, we are in position of presenting the most important parameters that have an effect on the result when we talk about capillary-pattern authentication, and we also detail the value that gives the optimum performance in Table 1.

We recorded a database containing sixty fingers, with five images of the DIP joint from each one. The number of subjects is different from the number of fingers, since we recorded more than one finger per subject. In total, the experiments have been run using 236 images qualified as \( Q_3 \), 46 images qualified as \( Q_2 \) and 23 images qualified as \( Q_1 \). In total, 305 images.
Another remarkable aspect of the database is the QbC of the images. The results show that is very important to make a classification of the images, because then, depending on the set of images we use, the performance improves or drastically worsens. In our case, we subjectively classified the images in three categories, and it permitted to obtain some useful plots about the behavior of the system depending on their quality.

With the proper simulations and the convenient processing, we have demonstrated the importance that the edges of the images have when we try to authenticate a client. The edge effect might change the performance since the information contained on the edges can be as important as the information contained in the rest of the image (due to translations).

The red images can be very useful from the point of view of correcting green images, but never for helping the authentication itself (the EER authenticating with red images is about 4.41%). They have information about the subjects, which comes from the background of the fingers, but not enough to reach low EERs.

Which finger we use for the authentication seems to be another consideration. It is not the same to authenticate people imaging the capillary-structures of the fingers from their most-used hand, than using the structures of the fingers from the other hand. Even though we found that might be a relation between the EER and the hands, the results are not definitive given the size of the database. It is something that might be submitted to further research. The age may be related with the performance as well.

Although we demonstrated that the FAR can decrease at expenses of increasing the FRR, given the conditions of using Q1 images and considering the U-shape of the curve, this system should be more aimed for security applications. The simulations show that it is not completely suitable for applications which require convenience (it has a high FRR).

## 2 Future Work

The work presented in this thesis can be extended in various ways.

- First of all, the system could be improved, in terms of application, by adding an identification function, instead of only considering authentication. It would entail to modify the existing algorithm. A first prototype of the system could be build with the work done till this moment.

- The database is not large enough to consider all the results completely reliable, that is, if we could consider the ideal case which is a database with the whole population, the
final EER of the system would be different. Then, an important task to accomplish before keeping on working on this project, is to get a larger database, at least with as many subjects as fingerprint databases, for example. Following with database considerations, we have demonstrated that a QbC can be relevant for the results. Nevertheless, the subjective classification we did, was not accurate enough, even though we followed the developed criteria. In order to have an objective classification, it would be useful for the computation of new results to consider an automatic method which properly classified the acquired images. The frequency content of the images could be a matter of study.

- Considering the results related with the extraction of the template, there are some issues on which is convenient to work. In section 4 we have presented the results taking into consideration the content of the image and without doing it. They show that even with a reduced number of subjects in the database, the value of the EER is lower when we make a hand selection rather than an extraction from the center of the image. Then, the next step is to develop an algorithm which considers the frequency content of the image; it means that when the frequency is higher in a specific part of the image, the number of blood vessels in that part is also larger and thus, making easier the authentication.

- From the beginning of the document, we have assumed that all the images in the database which come from the same subject, contain the same ROI. But this is a manual task which consists on looking for the same region before taking every picture. It is an inaccurate method, and moreover, may entail errors. Since the chosen interphalangeal joint for this project has been the distal one (closer to the fingerprint), it would be interesting to design a technique which was able to localize the same ROI for every subject automatically. A proposal for that is to use the fingerprint, more precisely, its core. The ROI of the DIP joint could be selected localizing the core of the fingerprint (which is already possible) and from there, detecting the the first joint of the finger. Once the system has localized the DIP joint, it is easy to select always the same ROI (for instance, taking a perpendicular line from the core till the minutiae disappears; then we are located in the upper part of the DIP joint). Using a method like this, it could be considered that the authentication using capillary-patterns would be reinforced by fingerprint authentication. Using this method, we could use only large magnifications (less resolution, but less sensitivity to movements).

- Although in section 4 we demonstrate that the use of red images for the decision does not improve the EER, we could consider these images for other purposes. Some of the images have low quality because of the presence of dust particles or bubbles. Since these artifacts are present in both green and red images, it is interesting to consider the option of digitally remove these undesired patterns by subtracting the images. Then, an image with a low qualification could be restored and improved. This process could be considered as a processing improvement. We could also include a deeper study about the filter (shape) as well as its order, the equalization algorithm, the correlation and the thresholding procedure. Any improvement in any of these processing steps could remarkably change the final EER achieved.

- According to the results presented in section 6.2, the detection with the right finger is better than the detection using the left finger. Then, in order to get a better performance, the authentication can be done by using images from two or more fingers for every subject (from the right hand in case of right-handed subjects). Then, if the EER using Q1 and one
finger is about 0.22%, we could make it lower by using two fingers; then, the new EER, considering the independance of two fingers of the same subject, would be computed as:

\[
EER_{1\_finger} = 0.0022
\]
\[
EER_{2\_finger} = EER_{1\_finger}^2 = 0.0022^2 = 4.66 \times 10^{-6}.
\]

Adding one finger to the authentication process, the EER improves three orders of magnitude (0.00046%). It means that the computation time is higher, but depending on the application, this low EER can be more important than the time that the system needs to perform the decision.

- Finally, we have to talk about the fact of using oil as fluid interface. The oil yields good results, but it is not convenient, because of its fatty properties (the subject needs to clean the hands after the authentication). It makes the system not very useful as a common authentication system. An important forward step would be to acquire images with enough quality without using any substance as interface (only the air), as the fingerprint case.
Appendix

3 Mathematical Approach to Bilinear Interpolation

Suppose that we want to find the value of the unknown function $f$ at the point $(x, y)$. It is assumed that we know the value of $f$ at the four points $(x_1, y_1)$, $(x_1, y_2)$, $(x_2, y_1)$, and $(x_2, y_2)$ (filled black dots in Fig. 5.1).

Figure 5.1: Bilinear interpolation.
We first do linear interpolation along the $x$-direction. This yields:

\[
f(x, y_1) = \frac{f(x_2, y_1) - f(x_1, y_1)}{(x_2 - x_1)}(x - x_1) + f(x_1, y_1). \quad (A-1)
\]

\[
f(x, y_2) = \frac{f(x_2, y_2) - f(x_1, y_2)}{(x_2 - x_1)}(x - x_1) + f(x_1, y_2). \quad (A-2)
\]

We proceed then interpolating in the $y$-direction.

\[
f(x, y) = \frac{f(x, y_2) - f(x, y_1)}{(y_2 - y_1)}(y - y_1) + f(x, y_1). \quad (A-3)
\]

Using (A-1) and (A-2) in (A-3) yields:

\[
f(x, y) = \left[ \left( \frac{f(x_2, y_1) - f(x_1, y_1)}{(x_2 - x_1)}(x - x_1) + f(x_1, y_1) \right) - \left( \frac{f(x_2, y_2) - f(x_1, y_2)}{(x_2 - x_1)}(x - x_1) + f(x_1, y_2) \right) \right] \frac{(y - y_1)}{(y_2 - y_1)} + \\
\left( \frac{f(x_2, y_1) - f(x_1, y_1)}{(x_2 - x_1)}(x - x_1) + f(x_1, y_1) \right).
\]

This gives us the desired estimation of $f(x, y)$:

\[
f(x, y) = \frac{(x_2-x)(y_2-y)}{(x_2-x_1)(y_2-y_1)}f(x_1, y_1) + \frac{(x_2-x)(y-y_1)}{(x_2-x_1)(y_2-y_1)}f(x_1, y_2) + \\
\frac{(x-x_1)(y_2-y)}{(x_2-x_1)(y_2-y_1)}f(x_2, y_1) + \frac{(x-x_1)(y-y_1)}{(x_2-x_1)(y_2-y_1)}f(x_2, y_2). \quad (A-4)
\]

As we see in (A-4), the interpolated point depends on $(x_1, y_1)$, $(x_1, y_2)$, $(x_2, y_1)$, $(x_2, y_2)$ and their images. Contrary to what the name suggests, this interpolation is not linear. Instead, it is of the form:

\[
f(x, y) = a_0 + a_1x + a_2y + a_3xy \quad (A-5)
\]

where:

\[
a_0 = \frac{1}{(x_2-x_1)(y_2-y_1)} \left[ y_1 \left( f(x_1, y_2) - f(x_2, y_2) \right) + y_2 \left( f(x_2, y_1) - f(x_1, y_1) \right) \right]
\]

\[
a_1 = \frac{1}{(x_2-x_1)(y_2-y_1)} \left[ y_1 \left( f(x_2, y_1) - f(x_2, y_2) \right) + y_2 \left( f(x_1, y_2) - f(x_1, y_1) \right) \right]
\]

\[
a_2 = \frac{1}{(x_2-x_1)(y_2-y_1)} \left[ f(x_1, y_1) - f(x_1, y_2) - f(x_2, y_1) + f(x_2, y_2) \right]
\]

\[
a_3 = \frac{1}{(x_2-x_1)(y_2-y_1)} \left[ f(x_1, y_1)x_2y_2 - f(x_1, y_2)x_2y_1 - f(x_2, y_1)x_1y_2 + f(x_2, y_2)x_1y_1 \right]
\]

In both cases, the number of constants

\[
a_i \quad i = 0, 1, 2, 3,
\]
correspond to the number of data points where \( f \) is given (see (A-5)). The interpolation is linear along lines parallel to either the \( x \) or the \( y \)-direction, equivalently if \( x \) or \( y \) are set constant. If we had first performed the linear interpolation in the \( y \)-direction and then in the \( x \)-direction, the resulting approximation would have been the same.
Bibliography


© Koninklijke Philips Electronics N.V. 2007