Ph.D. in Electronic and Computer Engineering
Dept. of Electrical and Electronic Engineering
University of Cagliari

Textural Features for Fingerprint Liveness Detection

Luca Ghiani

Advisor: Gian Luca Marcialis
Curriculum: ING-INF/05 SISTEMI DI ELABORAZIONE DELLE INFORMAZIONI

XXVII Cycle
April 2015

Ph.D. in Electronic and Computer Engineering
Dept. of Electrical and Electronic Engineering
University of Cagliari

Textural Features for Fingerprint Liveness Detection

Luca Ghiani

Advisor: Gian Luca Marcialis
Curriculum: ING-INF/05 SISTEMI DI ELABORAZIONE DELLE INFORMAZIONI

XXVII Cycle
April 2015

Dedicated to my family

Abstract

The main topic of my research during these three years concerned biometrics and in particular the Fingerprint Liveness Detection (FLD), namely the recognition of fake fingerprints. Fingerprints spoofing is a topical issue as evidenced by the release of the latest iPhone and Samsung Galaxy models with an embedded fingerprint reader as an alternative to passwords. Several videos posted on YouTube show how to violate these devices by using fake fingerprints which demonstrated how the problem of vulnerability to spoofing constitutes a threat to the existing fingerprint recognition systems.

Despite the fact that many algorithms have been proposed so far, none of them showed the ability to clearly discriminate between real and fake fingertips. In my work, after a study of the state-of-the-art I paid a special attention on the so called textural algorithms. I first used the LBP (Local Binary Pattern) algorithm and then I worked on the introduction of the LPQ (Local Phase Quantization) and the BSIF (Binarized Statistical Image Features) algorithms in the FLD field.

In the last two years I worked especially on what we called the “user specific” problem. In the extracted features we noticed the presence of characteristic related not only to the liveness but also to the different users. We have been able to improve the obtained results identifying and removing, at least partially, this user specific characteristic.

Since 2009 the Department of Electrical and Electronic Engineering of the University of Cagliari and the Department of Electrical and Computer Engineering of the Clarkson University have organized the Fingerprint Liveness Detection Competition (LivDet). I have been involved in the organization of both second and third editions of the Fingerprint Liveness Detection Competition (LivDet 2011 and LivDet 2013) and I am currently involved in the acquisition of live and fake fingerprint that will be inserted in three of the LivDet 2015 datasets.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Among the biometric measurements [26], fingerprints are probably the best known and popular due to their properties: universality (all we have it), duration (they do not change during time), individuality (there aren't two identical fingerprints) [27]. They are simple to use because we do not need to remember any password or use any type of smartcard and, moreover, their scan is not invasive as that of other biometric measures (such as the the iris scan). Unfortunately it has been shown that fingerprint scanners can be fooled by an artificial replica of a fingertip. An attacker who present to the electronic sensor a spoofed fingerprint which reflects the pattern of ridges and valleys (respectively the dark and the bright lines in a fingerprint image) of an authorized user could pass the system check. The first papers on the topic of the so-called "Fingerprint Liveness Detection" [38, 37] only appeared in 2002. The "liveness detection" or "fakes recognition" is used to determine whether a biometric trait is true (it comes from a living person) or false (it comes from an artificial replica).

In the case of fingerprints, a FLD module, independent from the recognition system, is based on the principle that additional information is used to verify if a fingertip image is authentic. Hardware based systems use additional sensors to gain measurements outside of the fingerprint image itself to detect liveness (biometric measurements as that of the heartbeat or the blood pressure on the fingertip). Software-based systems use image processing methods to gather information directly from the collected fingerprint image to detect liveness. These methods can be divided into two main categories, static if they analyze a single static image or dynamic if they analyze multiple images of the same fingerprint, captured while the subject puts his fingertip on the acquisition surface for a certain time period. This thesis is focused on the study of software-based modules with static methods.

Over the years many different algorithms have been proposed to detect liveness. Some are based on "live" characteristic as the presence of pores and others on "fake" characteristic like artifacts typical of an artificial replica. In recent years it has been paid a special attention to several textural algorithms developed in the Oulu University (Finland). Among those we studied the LBP (Local Binary Pattern) algorithm [45] and then we worked on the introduction of the LPQ [19] (Local Phase Quantization) and the BSIF [17] (Binarized Statistical Image Features) algorithms in the field of FLD. The LBP were first employed for two-dimensional textures analysis and excellent results were obtained due to his invariance with respect to grey level, orientation and rotation. It extracts certain uniform patterns corresponding to microfeatures in the image. The histogram of these uniform patterns occurrence is capable of characterize the image as it combines structural (it identify structures like lines and borders) and statistical (micro-structures distribution) approaches. The LPQ [56] is a blur insensitive texture classification method that can be used successfully in Liveness Detection, because it is able to represent all spectrum characteristics of the images in a very compact feature representation. The use of this algorithm represents a step ahead with respect to others previous work in which a simple spectrum analysis showed some benefits but was not effective enough to detect the fingerprint liveness. Therefore, the main reason of proposing this approach is to point out the spectrum differences between a "live" fingerprint and a "fake" one. The BSIF [31] is a local image descriptor constructed by binarizing the responses to linear filters but, in contrast to previous binary descriptors, the filters are learnt from natural images using Independent Component Analysis (ICA). The BSIF descriptor has two parameters: the filter size and the number of features extracted. Our experiments proved that, with a sufficient number of features, this algorithm clearly outperformed both LBP and LPQ.

Unfortunately, despite the fact that many algorithms have been proposed so far [32], none of them showed the ability to clearly discriminate between real and fake fingertips. The main problem is due to the difficulty of training appropriately a liveness detector, since fake fingerprint images can derive from replicas made up of a wide spectrum of materials, and it is practically impossible to cover this range; moreover, each algorithm has its own rationale. To the state-of-the-art, we can consider some liveness measurements based on the live fingerprint characteristics, as the perspiration or the ridge-valley consistency. On the other hand, other liveness measurements are based on the hypothesis that the fabrication process leads to significant modifications, due to the elastic deformation of the replica, the presence of artifacts, the loss of details. From this point of view, we tried to exploit these basic differences by concatenating the feature vectors provided by each algorithm [18]. We pointed out that current FLD algorithms cannot be adopted individually, but their combination, carefully handled, can help improving the performance, thus allowing their integration in current fingerprint verification systems.

However the most important part of my research was the analysis of the extracted features and what we called the "user specific" problem. As a matter of fact, as previously stated, the FLD algorithms performances are not yet good enough and, in our opinion, one of the reasons of such a low performance is due to a "user specific" component that reduces their ability in discriminating between live and fake fingerprint images. We provided a model for this "user specific" effect and we are the first to propose an approach to user specific FLD. Until now that of the FLD has always been dealt with as a "general purpose" problem with a certain amount of live and fake fingerprint images to be analyzed regardless of the users to which they belonged. But our research suggests that some of the extracted features also retain information related to the user as they describe the texture of that particular fingerprint. We noticed indeed that the live and fake samples of the same user are closer between them than to live and fake samples of other users. This user specific characteristic, with regard to the liveness detection, is a bias that should be removed and we successfully did it further improving the obtained results.

Since 2009, in order to assess the achievements of the state of the art in FLD, the Department of Electrical and Electronic Engineering of the University of Cagliari, and the Department of Electrical and Computer Engineering of the Clarkson University, have organized three editions of the Fingerprint Liveness Detection Competition (LivDet) [35, 59, 20]. The goal of the competition is to compare software-based FLD methodologies, using a standardized testing protocol and large quantities of spoof and live fingerprint images. The participants have to provide an algorithm capable of processing each image and output a liveness score normalized in the range between 0 and 100 (100 is the maximum degree of liveness, 0 means that the image is fake). In three different editions of the competition, eleven datasets were created using different sensors thanks to hundreds of volunteers, whose fingerprints were acquired several times. Some of those fingerprints were also “spoofed” with the cooperative or the non-cooperative method. In the cooperative method the volunteer pushes his finger into a plasticine like material creating a negative impression of the fingerprint as a mold. The mold is then filled with a material, such as gelatin, PlayDoh or silicone and the fake fingertip is created. In the non-cooperative method a latent fingerprint left on a surface is enhanced, is digitized through the use of a photograph, and finally the negative image is printed on a transparency sheet. This printed image can then be made into a mold, for example, by etching the image onto a printed circuit board which can be used to create the spoof cast.

The First International Fingerprint Liveness Detection Competition, LivDet 2009 [35], provided an initial assessment of software systems. The second [59] and third [20] Liveness Detection Competition were created in order to ascertain the current state of the art in liveness detection. Besides the work on the algorithms and the analysis of the extracted features, I have been involved in the organization of both second and third editions of the competition (LivDet2011 and LivDet2013) that were both open to all academic and industrial institutions and I’m currently involved in the acquisition of live and fake fingerprints that will be inserted in three of the LivDet 2015 datasets.

In the second chapter of this thesis the biometric systems, measures and vulnerability are introduced paying particular attention to fingerprints. A general overview of Fingerprint Liveness Detection methods and state-of-the-art is presented in chapter three. Textural algorithms are dealt in chapter four together with the “user specific” problem. Experimental results are shown in chapter five. Limitation and contributions of this work are described in the sixth chapter.

Chapter 2

Biometric systems

Biometry allows to perform a person identification based on his physical (fingerprints, face, iris) or behavioral (gait, signature) attributes and to establish an identity based on who you are, rather than what you possess (e.g. a card that can be lost or stolen) or what you remember (e.g. a password that can be forgotten) [26]. Biometric measures were first used by Alphonse Bertillon during the nineteenth century. He created a method for the convicts personal identification based on 11 measurements of different parts of the body such as head size, height or expanse of the arms. When it became clear that those measures were non unique, police departments and prisons in both the United States and Europe switched to fingerprint identification. Since then we have seen a significant growth in biometric research with the development of new sensors, novel feature extraction and many different matching algorithms.

A biometric system is a pattern recognition system. Biometric data is acquired from a person, a features set is extracted from the data, these features are compared against those stored in the database and an action is executed based on the comparison result. It can be schematized in having four main modules (see Figure 2.1):

- a sensor module (a biometric reader or scanner) required to acquire the raw biometric data of an individual;
- a quality assessment and feature extraction module that assesses the quality of the biometric data acquired by the sensor using, if necessary, a signal enhancement algorithm and extracts a set of features;
- a matching module that compares the features with the stored templates to generate match scores;
- a database module that is the repository of the biometric information saved during the enrollment process.

The template of a user can be extracted from a single biometric sample or it can be generated by processing multiple samples. In order to account for the intra-class variations associated with a single user, some systems store multiple templates. A biometric system may

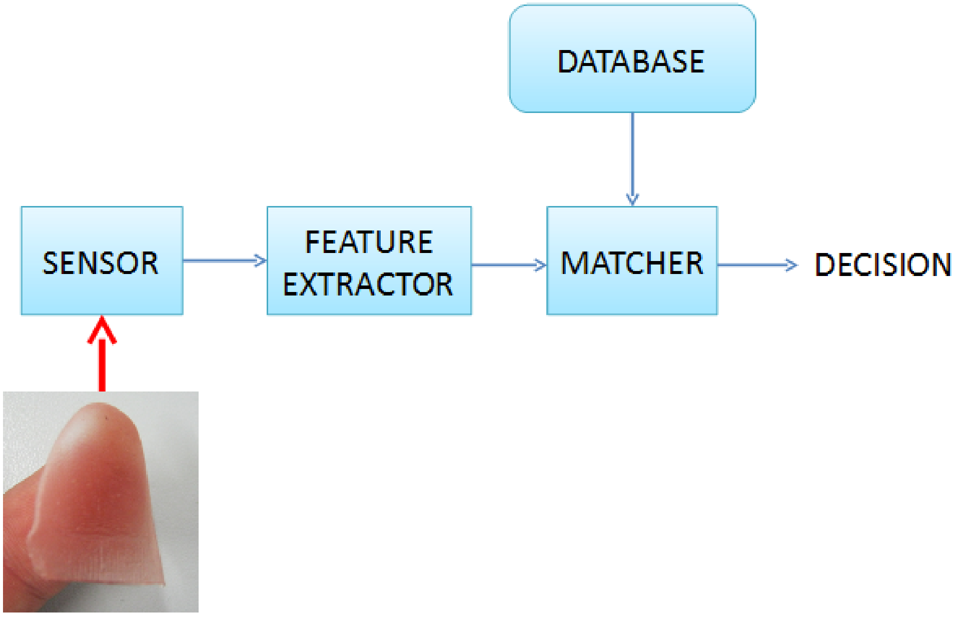


Figure 2.1: Biometric system scheme.

function either in verification or identification mode. In the verification mode, the system performs a one-to-one comparison to determine whether the identity is the one claimed by the individual. It answers to the question “am I whom I claim I am?”. In the identification mode, the system analyzes the templates of all the users in the database looking for a match in order to establish a person identity. It answers to the question “who am I?”.

Some of the characteristic of a biometric system should be [26, 8]:

- Universality: each person is expected to have this biometric.
- Distinctiveness: the data collected from any two persons should be sufficiently different.
- Permanence: over the time the biometric should be sufficiently invariant.
- Collectability: the biometric can be measured quantitatively.

However, there are other properties that should be considered in practice:

- Performance: the recognition of an individual is expected to be achieved with sufficient accuracy and speed.
- Acceptability: an indicator of how much a biometric system is accepted by the people.
- Circumvention: it shows how easy (or difficult) it is to fool the system.

2.1 Biometric measures

There is a great number of biometric measures that can be divided into two main groups: physiological and behavioral. Some of the principal biometrics are:

- DNA: it is a code unique for every person except for identical twins that share the same DNA. Its use is limited for different reasons: is easy to steal a piece of DNA from an unaware subject; a real-time recognition is impossible due to the long lasting matching process; the privacy of an individual can be violated obtaining information like susceptibilities to certain diseases.

- Ear: the shape of the ear is distinctive although the features obtained are not unique. It is based on the measurement of the distance between distinctive points.

- Face: it is the biometric that mankind has always used. The pro is that it is a non intrusive method while the con is the difficulty in ignoring the effects of ageing, the changes in facial expressions, the variations in the imaging environment and in the pose of the face.

- Infrared thermogram: with an infrared camera it is possible to capture the heat radiated by a body and it is possible to distinguish even identical twins. It is a non-invasive technology but heat sources near the body may affect the acquisition.

- Gait: it is a measurement of the way a person walks but it is not very distinctive. Moreover, it may not be invariant during a period of time due to variations of weight, injuries, inebriety and mood.

- Hand and finger geometry: it uses features of the human hand, like size and length of the fingers but it is not very distinctive.

- Iris: it is distinctive for each person and each eye and its recognition is extremely accurate and fast. The iris image is captured with the cooperation of the user.

- Keystroke dynamics: it is a behavioral biometric based on the hypothesis that each person types on a keyboard in a characteristic way. It is not unique and large variations of patterns can be observed for some subjects.

- Odor: the odor of an object is characteristic of its chemical composition. An array of chemical sensors can detect the odor emitted by a body that is distinctive of that person.

- Retinal scan: it is considered one of the most secure biometric because the retinal vasculature is supposed to be different for each individual and each eye and it is difficult to replicate. Unfortunately the high quality of the capture devices makes them very expensive and it is required a great cooperation by the users.

- Signature: it is a behavioral biometric that has legal value even though it changes through time and it is influenced by physical and emotional conditions.

- Voice: it is easy to obtain but is not distinctive enough because it can be influenced by the physical and psychological condition of the subject and it can be imitated. If a person must be recognized over the phone it is the only biometric measure that can be used.

- Fingerprints: they are the most used, oldest and well-known biometric measurements. Everybody has them and each one is different from the other. However, as recently shown, they can be forged and that is the subject of this thesis.

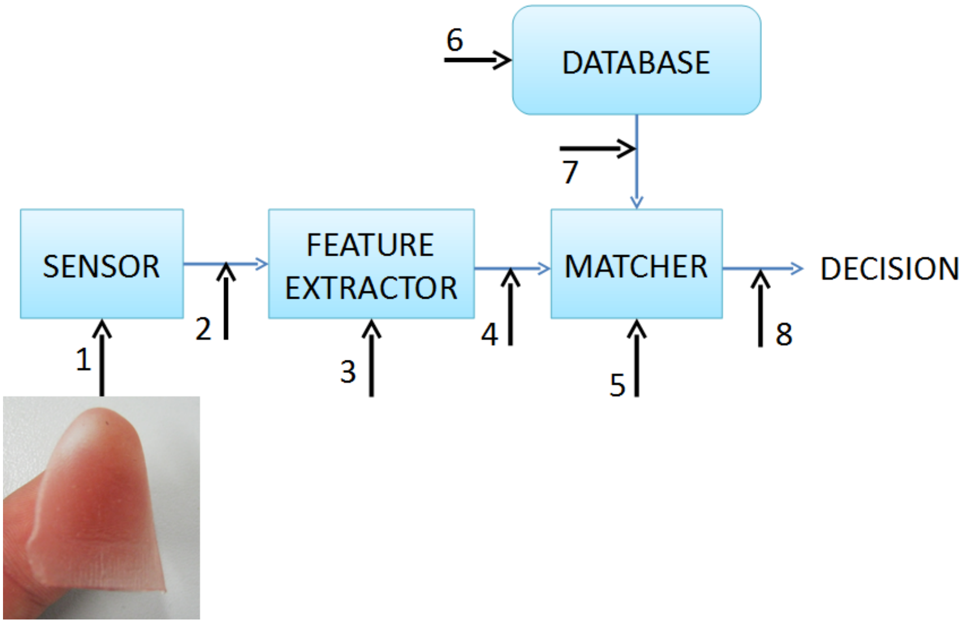


Figure 2.2: Different attacks to a biometric system.

- Palmprint: it is based on the observation of the ridges flow and the minutiae extraction of the entire hand.

2.2 Vulnerability of biometric systems

There are no systems 100% secure. A biometric system can be subject to several kind of threats [26]:

- Circumvention: the attacker gains access to protected resources.
- Covert acquisition (contamination): the attacker access the system using biometric information taken from a legitimate user.
- Collusion and coercion: the attacker is a legitimate user that acts freely (collusion) or under some kind of threat (coercion).
- Denial of service (DoS): the system performances are slowed down or stopped so that the legitimate users are no longer able to access the system.
- Repudiation: the attacker, a corrupt user, denies accessing the system claiming that his biometric data were stolen or faked.

According to Ratha et al. there are several kinds of attack of a biometric system [48] (see Figure 2.2):

1. Presenting fake biometrics at the sensor: a reproduction (in our case a fake fingerprint) is presented as input to the system.

2. Resubmitting illegally intercepted data (replay): a signal is stored and then it is replayed to access the system.

3. Overriding the feature extraction process: The feature extractor is attacked using a Trojan horse, so that it produces feature sets preselected by the intruder.

4. Replacing legitimate features: The features extracted from the input signal are replaced with a different feature set.

5. Corrupting the matcher: The matcher is attacked and corrupted so that it produces preselected match scores.

6. Tampering with stored templates: if the attacker is able to modify a template in the database, he could authorize a fraudulent individual or he could deny service to a legitimate user.

7. Attacking the channel between the stored templates and the matcher: the data sent to the matcher through a communication channel could be intercepted and modified.

8. Overriding the final decision: if an attacker is able to override the final match decision, then the authentication system has been disabled.

2.3 Errors in a biometric system

The output of a biometric identification system is a matching score s (usually a value in the range $[0,1]$) that is a measure of the similarity between the data received in input and a template in a database. A value closest to 1 represents a highest similarity. Once a threshold t is selected, if the score s is higher than t, then we have a match: the input data and the template belong to the same subject. There are two kind of errors that can be made: a positive match between sets of measures belonging to different persons, called false match and a negative match between sets of measures belonging to the same person, called false rejection. An evaluation of those errors is given by two measures:

- FAR (False Acceptance Rate): percentage of impostors accepted as genuine users. It is the fraction of impostor scores that exceed the threshold t:

$$FAR(t) = \int_t^1 P(s|impostor)ds \tag{2.1}$$

- FRR (False Rejection Rate): percentage of users rejected as impostors. It is the fraction of genuine scores below the threshold t:

$$FRR(t) = \int_0^t P(s|genuine)ds \tag{2.2}$$

As we vary the threshold t the value of both FAR and FRR varies: if t grows also the FRR grows, while the FAR decreases and, conversely, if t decreases also the FRR decreases, while the FAR grows (see Figure 2.3).

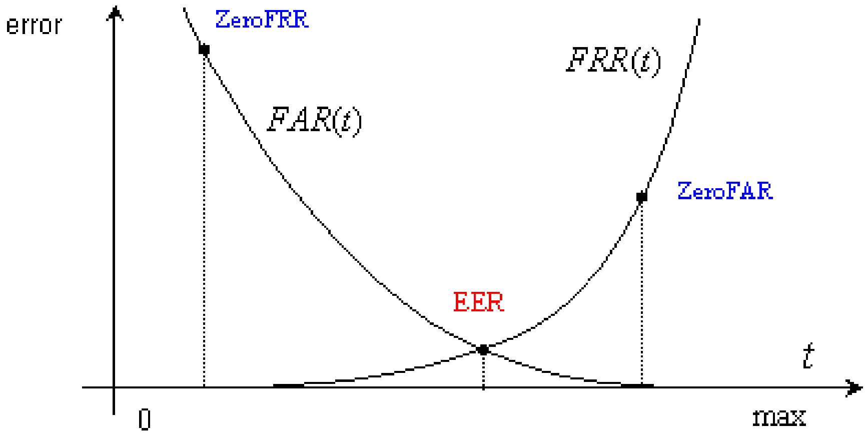


Figure 2.3: False acceptance rate (FAR) and false rejection rate (FRR) as functions of the threshold t.

There are other important values that we can consider:

- GAR (Genuine Acceptance Rate): it is the percentage of genuine users accepted by the system. It is defined as:
$$GAR = 1 - FRR \tag{2.3}$$
- EER (Equal Error Rate): it is the point where $FAR(t) = FRR(t)$.
- HTER (Half Total Error Rate): it is defined as $(FAR + FRR)/2$.
- 1%FAR (1%FRR): It is the FRR (FAR) that correspond to the threshold for which the FAR (FRR) is fixed to 1%.
- ZeroFAR (ZeroFRR): It is the FRR (FAR) that correspond to threshold for which the FAR (FRR) is fixed to 0%.

The different values of FAR and FRR (or GAR), obtained at the variation of the threshold, can be show with a curve called ROC (Receiver Operating Characteristic) (see Figure 2.4).

2.4 Fingerprints

Fingerprint ridges formation is a combination of factors both environmental and genetic [27]. As a matter of fact, DNA gives general instructions on the skin formation in a fetus but the specific way in which it happens it is a result of random events (and that is why even the fingerprints of identical twins are different). There are so many variations during the formation of a fingerprints that it would be impossible for two fingerprints to be exactly alike (but this uniqueness has never been demonstrated).

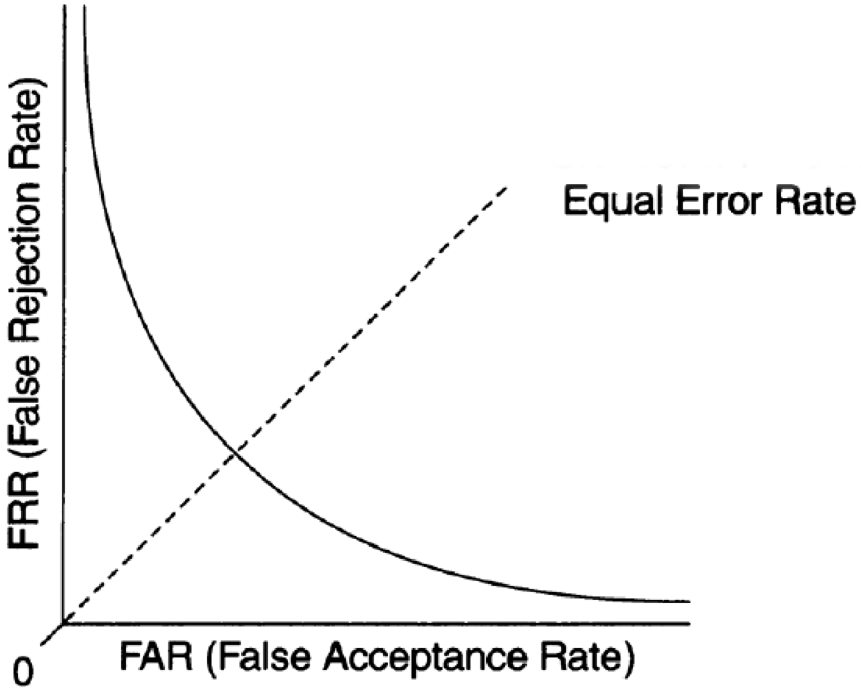


Figure 2.4: Receiver Operating Characteristic (ROC curve) and Equal Error Rate (EER).

2.4.1 A brief history

Fingerprints have been probably used since ancient times as a kind of signature but the first study has been published in 1686 by Marcello Malpighi. He was an anatomy professor at the University of Bologna who for the first time analyzed ridges, spirals and loops in fingerprints. Fingerprints individuality was suggested in 1880 by Henry Fauld. In 1888 Sir Francis Galton introduced the use of the features called minutiae for fingerprint matching and in 1899 Edward Henry devised a system of fingerprint classification. Since the beginning of the twentieth century fingerprints were used for criminal identification and in the early 1960s, thanks to the development of computer processing, the AFIS (Automatic Fingerprint Identification Systems) was created. Nowadays, thanks to less expensive devices and reliable matching algorithms, fingerprints are used not just in forensic applications but also for personal identification.

2.4.2 Fingerprint structure

We asserted that fingerprints are unique but that uniqueness depends on the type and number of features extracted (basically less features means less details and therefore less information obtained). Fingerprints are composed of epidermic ridges and valleys that usually run in parallel. On the images obtained through sensors dark lines are the ridges while the bright lines are the valleys. The extracted images details can be divided in three different levels: on the first level the global pattern of the fingerprint is analyzed, on the second there are the so called minutiae (such as ridge bifurcations and endings) and on the third level some

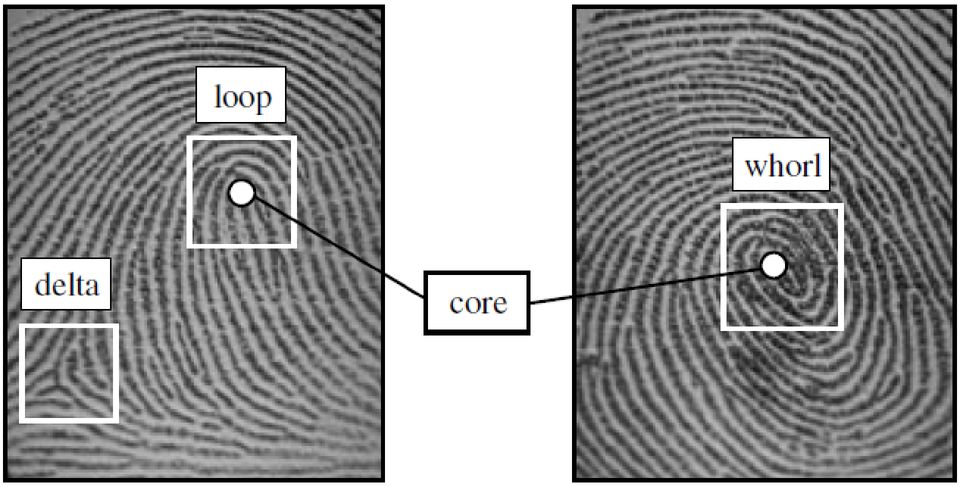


Figure 2.5: Core and delta singular points in fingerprint images.

fingerprint micro-characteristics are analyzed such as pores and ridges contour.

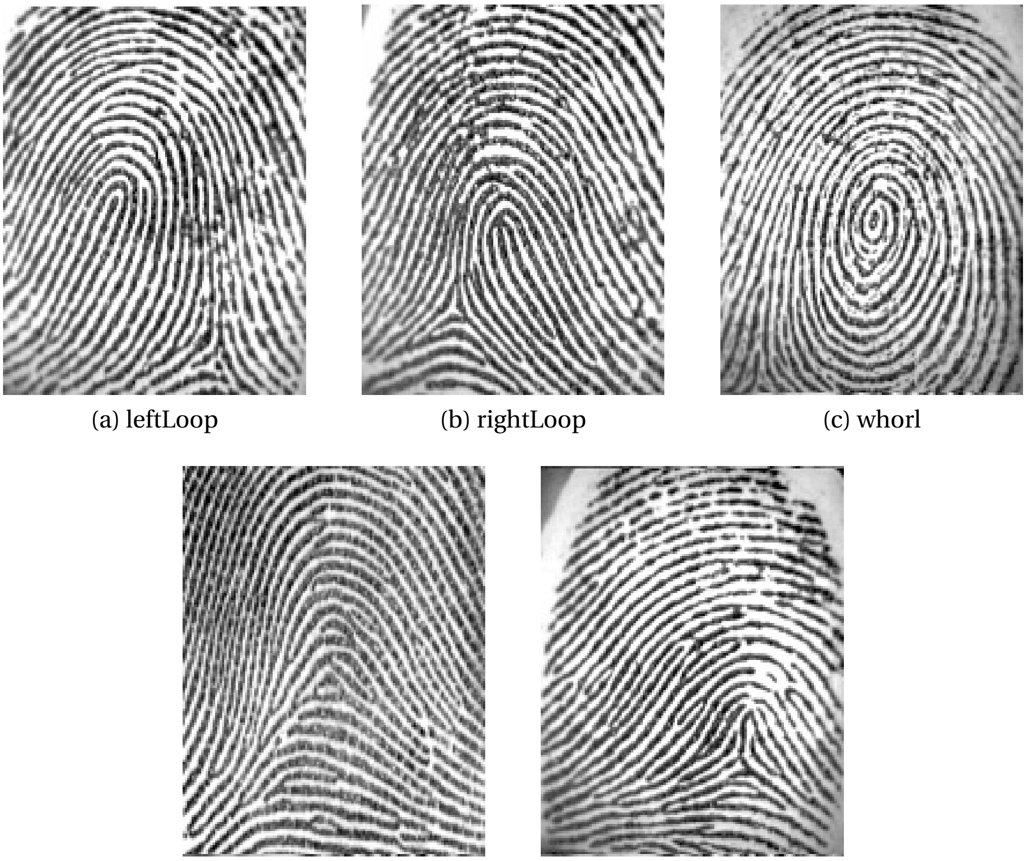
First level features

The singular points and ridges flow are the features of this level. The singular points are the centers of regions where the ridge orientation is erratic. There are two types of these points (see Figure 2.5):

- Core: it is the center of a series of marked curves of ridges and valleys. It is frequently considered the center of the fingerprint.
- Delta: it is the center of a kind of triangle described by three patterns.

Following Henry classification system and depending on the ridges flow, core and delta positions, it is possible to distinguish five major pattern categories (see Figures 2.6 (a-e)):

1. Arch
2. Tented Arch
3. Right Loop
4. Left Loop
5. Whorl



(a) leftLoop

(b) rightLoop

(c) whorl

(d) arch

(e) tentedArch

Figure 2.6: Examples of the five major fingerprint categories.

Second level features

At a local level some ridges discontinuity called minutiae can be found. There are many different kind of minutiae but they can be divided in two main groups: termination (when a ridge end) and bifurcation (when a ridge splits into two ridges). During a fingerprint features extraction each minutia is usually described by his position with respect to a common point of origin and some additional characteristic such as direction (of the ridge) or type (of the minutia). The most popular and used fingerprint matching methods are based on the matching of minutiae represented by the triplet (x, y, θ) where (x, y) is the minutia position and θ is the orientation of the ridge in the point where the minutia is (see Figure 2.7).

Third level features

The third level features are the fingerprint micro-characteristic such as the pores. A pore is formed when a sweat gland under the skin generates a sweat duct that reaches the surface of the epidermis. Once these pores are formed on the ridge, they are fixed at that location. It has been shown that pores do not disappear, neither move or spontaneously generate over time [49]. In addition to the minutiae, pores can be used in the matching process because of

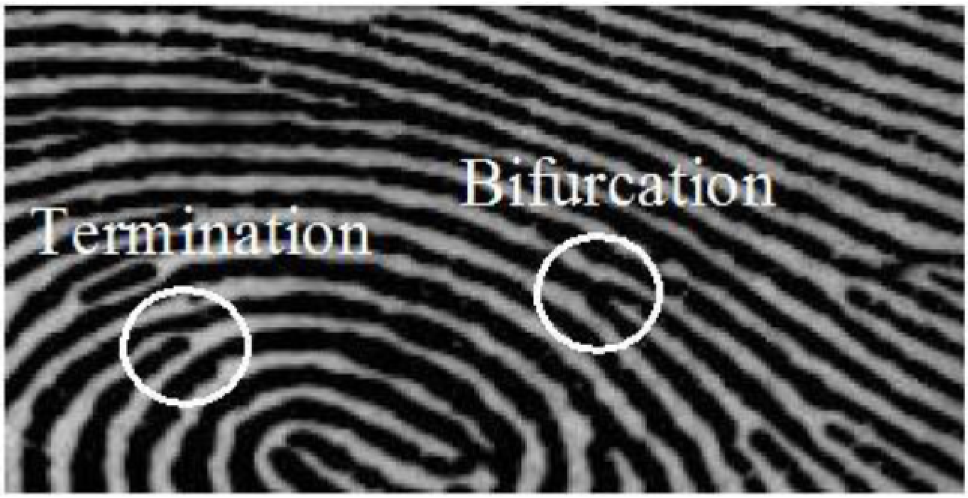


Figure 2.7: Fingerprint minutiae: termination and bifurcation.

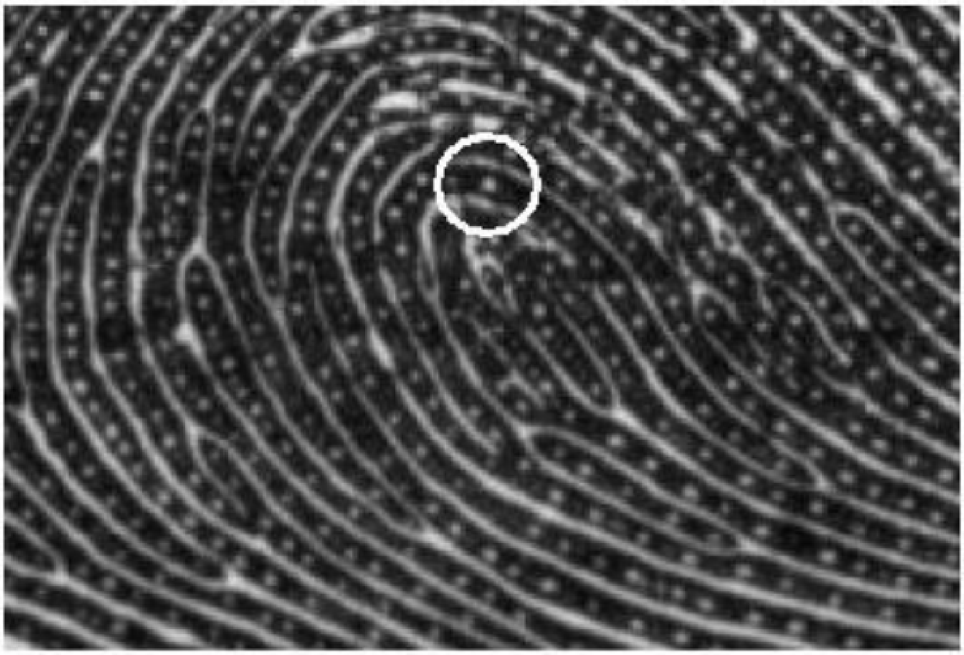


Figure 2.8: Pores in a fingerprint.

their unique configuration. In this thesis we analyze the number of pores of a fingerprint for liveness detection purposes (see Figure 2.8).

2.4.3 Fingerprint scanners

Up to recently fingerprints were acquired with the ink-technique which is a process called off-line. Another case of off-line acquisition is that of the latent fingerprints found in crime scenes. Nowadays, the acquisition is an on-line process executed by digital scanners. There are several types of sensors and their collected images have different characteristics. The main characteristics of the images obtained by a fingerprint scanner are [1]:

- Resolution: it indicates the number of dots or pixels per inch (dpi). The minimum resolution for FBI-compliant scanners is 500 dpi.
- Area: it is the size of the fingerprint portion that can be acquired by the scanner. According to the FBI it should be at least 1x1 square inches.
- Number of pixels: it is the number of pixels in the fingerprint image.
- Dynamic range (or depth): it is the number of bits used to encode the intensity value of each pixel. FBI standard requires a depth of 8 bit.
- Geometric accuracy: it is the geometric distortion introduced by the acquisition device with respect to the x and y directions.
- Image quality: it is difficult to define and especially tell how much the level of this quality depends on the image itself or on the intrinsic finger quality.

The fingerprint sensors can be:

- Optical sensors: the fingertip is illuminated by a bank of light-emitting diodes (LED) and the light is absorbed at the ridges (that is why they appear dark) and reflected at the valleys (which appear bright). The reflected image is focused onto a CCD or CMOS image sensor and acquired [27]. There are several kind of internal mechanisms:
  - Frustrated Total Internal Reflection (FTIR): when the finger touches the top side of a glass prism there is a contact between the prism surface and the ridges while the valleys remain distant. The light enters the prism, it is reflected, exits from one side and it is focused through a lens onto the sensor.
  - FTIR with a sheet prism: it uses a number of prisms instead of a single large one and that is why the size is reduced (as a matter of fact the sheet prism is nearly flat) but the quality of the images is generally lower than the one obtained with a traditional FTIR.
  - Optical fibers: the sensor size can be reduced if prism and lens are substituted with a fiber-optic platen. The CCD/CMOS is in direct contact with the platen therefore it is greater than the one in the FTIR and the cost of the device is higher.
  - Electro-optical: it is a device constituted by two layers. In the first one there is a light-emitting polymer; the potential across the surface is not the same because the ridges touch the polymer while the valleys do not and the emitted light depends on that potential. In the second layer there is a photodiode array that receive the light emitted and produce the digital image.
  - Direct reading: the finger does not touch the surface of the device as it is equipped with a mechanical support through which a fingertip picture is taken by an high-quality camera. Obviously the surface does not require any cleaning but it is difficult to obtain well-focused and high-contrast images.

- Solid-state sensors (silicon sensors): They are made up of an array of pixels, each pixel being a tiny sensor itself. There are not lens, CCDs or CMOSs and the fingertip stand directly on the silicone surface. Both Apple Iphone and Samsung Galaxy new models have solid state sensors. Depending on the way the information is converted into electrical signals, four different kind of sensors can be distinguished:
  - Capacitive: it is a two-dimensional array of micro-capacitor plates while the other plate of each capacitor is the finger skin itself. Between the finger and the silicon plate small charges are created. The charges magnitude depends on the distance between the fingerprint and the plates and consequently it is different for ridges and valleys.
  - Thermal: it is made of pyro-electric material that generates current based on temperature differential that varies from the ridges, in contact with the surface, to the valleys. The contact produces the image that is acquired.
  - Electric field: consists of a drive ring that generates a sinusoidal signal and a matrix of active antennas that receives a very small amplitude signal modulated by the subsurface of the skin. For that reason, it does not matter if fingertips are too dry, too humid or even if they have little scars.
  - Piezoelectric: a dielectric material on the surface generates a different amount of electric current with the change applied by the finger pressure. Due to a different distance from the sensor surface, the current produced by the so called piezoelectric effect is different for ridges and valleys.
- Ultrasound sensors: They are composed by a transmitter that generates acoustic signals sent toward the finger and by a receiver that detects the pulses bounced off the fingerprint surface. In this way it is possible to compute a good quality image of the subsurface of the skin. Despite the good quality obtained, this sensor is far too big and expensive. Additionally, it takes a few seconds to acquire an image.

Chapter 3

Fingerprint liveness detection state of the art

3.1 Introduction to fingerprint liveness detection

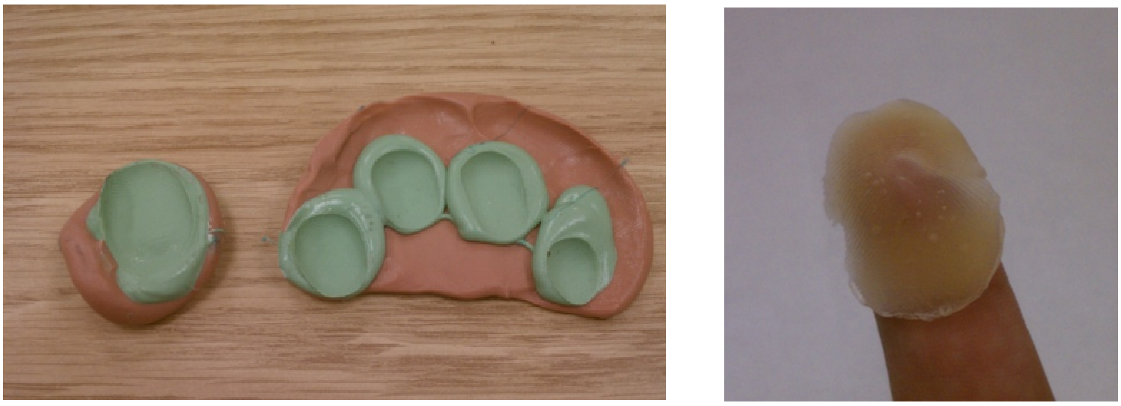
A spoof is a counterfeit biometric that is used in an attempt to circumvent a biometric sensor. In the case of fingerprints, it can be an artificial finger or a dismembered finger. The purpose of liveness detection is to distinguish “live” from “fake” biometric traits [26]. Liveness detection is based on the principle that additional information can be garnered above and beyond the data procured by a standard verification system, and this additional data can be used to verify if a biometric measure is authentic. The concept of spoofing has existed for some time now. Research into spoofing began in 1998 from research conducted by D. Willis and M. Lee where six different biometric fingerprint devices were tested against fake fingers and it was found that four out of six were susceptible to spoofing attacks [58]. This research was approached again in 2000-2002 by multiple institutions including Putte et al. [57] as well as Matsumoto et al. [38]. The presented research looked at the vulnerability of spoofing. In 2002, Schuckers delved into using software approaches for liveness detection [50].

3.2 Fingerprint reproduction process

There are two general forms of creating artificial fingers, the cooperative method and the non-cooperative method:

- Cooperative method: the subject pushes his or her finger into a malleable material such as dental impression material, plastic, or silicon gum creating a negative impression of the fingerprint as a mold (see Figure 3.1 (a)). The mold is then filled with a material, such as Gelatin, PlayDoh or Silicone. This cast is an almost perfect copy of a live subject finger (see Figure 3.1 (b)).

- Non-cooperative method: it involves enhancing a latent fingerprint left on a surface, digitizing it through the use of a photograph, and finally printing the negative image on a transparency sheet. This printed image can then be made into a cast. Another way is to etch the image onto a printed circuit board and then drip material onto the PCB to make the cast (see Figure 3.2).



(a) Mold

(b) Fake fingertip

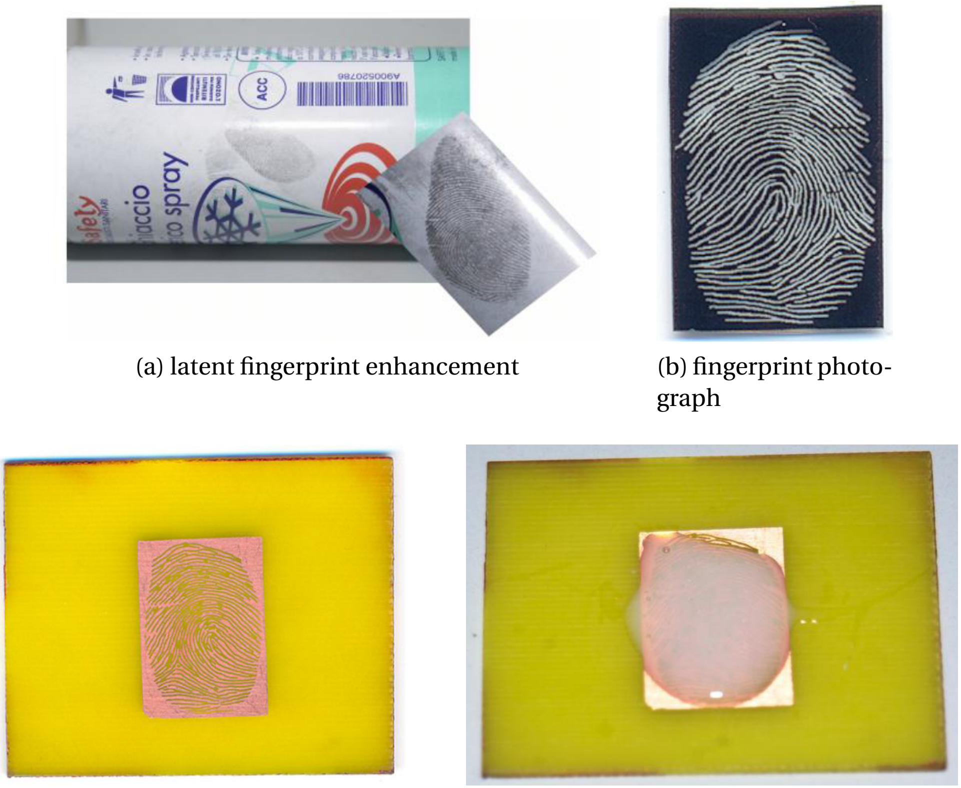
Figure 3.1: Mold that reproduce a negative impression of five fingerprints (a) and latex spoof of a fingerprint (b).

3.3 Liveness detection methods

The challenge for a fingerprint recognition systems is the ability to detect if a presented fingerprint has been acquired from a live person or an artificial finger. Systems are being upgraded to incorporate liveness detection solutions that will be able to detect if the submitted probe is a spoof or live finger [32, 52]. Following the taxonomy proposed by Coli et al. [11], those detection methods are divided into two main categories: hardware based and software based:

1. hardware based: liveness detection can be incorporated into a system through the addition of hardware components to the capture device that can search for traits in the fingerprint using blood pressure, electrocardiogram, temperature or other methods.

2. Software based: liveness detection can also be implemented through the use of algorithms that are added to the system. It is a less expensive approach because no additional hardware is required. These methods extract features from the fingerprint images in order to determine liveness and, according to the number of images examined, they are called dynamic (the features extracted derive from the analysis of multiple frames of the same fingerprint) or static (the features are extracted from a single fingerprint impression or from the comparison of different impressions of the same fingerprint). Another subdivision of the software based methods can be grounded on the physical principle they exploit: the perspiration, the elastic distortion and the morphology of the fingerprint.



(a) latent fingerprint enhancement

(b) fingerprint photograph

(c) negative image

(d) latex poured over the printed image.

Figure 3.2: Non-cooperative method.

3.4 Software-based methods state of the art

As previously stated, the software-based approaches to liveness detection does not require additional hardware. We analyze the state of the art of dynamic and static methods.

3.4.1 Dynamic methods

During the contact between the finger and the scanner surface the skin gets wetter due to the increasing amount of sweat. Since the pores on the fingertip surface are the source of the perspiration process, Derakhshani et al. in [12] examined how they change in sequential frames during a fixed interval of few seconds in both live and fake fingertips. The variation of the wetness of the fingertip skin reflects on a variation of the gray-level profile of the acquired images. In order to evaluate this feature, the fingerprint skeleton of the image at 0 and 5 seconds is converted into a couple of mono-dimensional signals. Several statistical measures are proposed on the basis of the obtained signals such as the total swing ratio, the Min/Max growth ratio, the last-first fingerprint signal difference mean and the percentage change of standard deviation.

Parthasaradhi et al. [46] introduced some modifications. Since excessive amount of wetness can produce a saturated signal, they added two new features: dry saturation percentage change and wet saturation percentage change.

In [9] Coli et al. proposed two dynamic features: the time variation of the grey level mean value of the whole image and the L1-distance of its grey levels histogram.

In [28], Jia et al. introduced some new features extraction methods based on the analysis of human skin elasticity. Once the fingertip is on the scanner surface, a sequence of fingerprint images which describes the finger deformation process is captured. Then two features which represent the skin elasticity are extracted from the image sequence: 1) the correlation coefficient between the fingerprint area and the signal intensity; 2) the standard deviation of the fingerprint area extension in x and y axes. Finally the Fisher Linear Discriminant is used to discriminate the real skin from artificial materials.

The work of Antonelli et Al. [3] adopted a dynamic procedure in order to perform a liveness detection based on elastic deformation. Once the finger is on the scanner surface, the user is invited to rotate the fingertip. An elastic tension and, consequently, an elastic deformation is caused by the movement of the fingertip on the surface. The deformation degree depends on the level of elasticity of the skin that is supposed to be different for live or artificial fingers. A sequence of images is acquired at high frame rate and the elastic distortion from each frame to the next one, is encoded in a features vector ("distortion code").

Abhyankar e Schuckers in [1, 51], used the perspiration phenomenon, observed only in live people, as a measure to classify "live" fingers from "not live" (fake or belonging to a cadaver) fingers. Multiresolution analysis and wavelet packet analysis are used to extract information from low frequency and high frequency content of the images respectively. The images were acquired as soon as the finger was leaned on the sensor and then after 5 seconds.

A method based on fingerprint deformation analysis was proposed by Zhang et al. [61]. The subject is required to put a finger on the sensor surface, and then apply some pressure in four different directions. A thin-plate spline (TPS) is used to model the distortion of live and spoof fingerprints. Since spoof materials are typically much more rigid, compared to the human skin, their deformation is lower. This deformation is represented by the displacement of paired minutiae obtained before and after distortion.

3.4.2 Static methods

Tan e Schuckers in [53] developed a new method that quantify the perspiration phenomenon in a single image. The underlying process is to extract the ridge signal which represents the gray level values along the ridge mask and then use wavelet transform to decompose this signal into multi-scales. On each scale, static features are extracted to quantify the perspiration pattern to distinguish between live and non-live fingerprints.

In another work [54], the above mentioned authors proposed a liveness detection method based on noise analysis along the valleys in the ridge-valley structure of fingerprint images. Statistical features are extracted in multiresolution scales using the wavelet decomposition technique.

Several methods based on the analysis of a single image have been proposed by Nikam and Agarwal. In [43, 43] they proposed a new method based on the ridgelet transform. As a matter of fact, wavelets are very effective in representing objects with isolated point singularities, but ridgelet transform allows representing singularities along lines in a more efficient way than the wavelets. They use ridgelet energy and co-occurrence signatures to characterize the fingerprint texture.

These two authors also propose [44, 42] the use of curvelet transform for liveness detection. Curvelet transform allows representing singularities along curves in a more efficient way than the wavelets. Ridges oriented in different directions in a fingerprint image are curved, hence curvelets are very significant to characterize fingerprint texture. As for ridgelet, the fingerprint texture is characterized by energy and co-occurrence signatures [4].

In [40], they also proposed the integrated use of Local Binary Pattern (LBP) and wavelet transform. LBP histograms are used to capture the textural details while wavelet energy features characterize ridge frequency and orientation information.

In [41] they used again textural measures based on wavelet energy signatures and Gray Level Co-occurrence Matrix (GLCM) features. In this work, in order to extract textural characteristics, they introduced some statistical measures defined by Haralick [23].

In all of these methods, proposed by Nikam and Agarwal, an algorithm is employed to reduce the dimensionalities of the feature sets: PCA (Principal Component Analysis) or SFFS (Sequential Forward Feature Selection). The results have been tested with three classifiers and then fused with an hybrid classifier.

Since in a live finger it can be noticed the regular periodicity due to the pores on the ridges while this regularity is not evident for spoof fingerprint signals, Derakhshani et al. [13] used one static feature based on the Fast Fourier Transform (FFT) of the fingerprint skeleton converted into a mono-dimensional signal.

In [55] Tan and Schuckers fused together their previous works [13, 53, 54] with a measure of the image quality.

Moon et al. [39], looking at the finger surface with an high resolution camera, observed that the surface of a fake finger is much coarser than that of a live finger. They employed a 1000 dpi sensor, whilst current sensors exhibit 500 dpi on average, and they did not work with the entire image (too large because of its resolution) but with subsamples of a fixed size. For extracting this feature, the residual noise returned from a denoising process applied to the original sub-images is considered. The standard deviation of this noise is then computed to highlight the difference between live and fake coarseness.

Chen et al. [5] presented a static method that use multiple impressions and is based on elastic-deformation features. Given a genuine query-template pair of fingerprints, the entity of elastic distortion between the two sets of extracted minutiae is measured. The idea is that live and spoof fingerprints show different elasticity response repeating the acquisitions.

Coli et al. [10] analyzed the frequency domain with a two-dimensional Fourier transform. Some fingerprint micro-characteristics are less defined in an artificial fingerprint image due to the roughness of the skin or to the ridge line discontinuity. Thereby, high frequency details can be removed or strongly reduced. In order to measure that reduction they computed the modulus of the Fourier transform of the image, called “Power Spectrum”. They defined the HFE (High Frequency Energy) as the square module integral of the image Fourier transform calculated outside a circular region with R as radius and centered on the null frequencies along both axis. HFE measure can characterize the information which distinguish fake and live fingerprints.

H.Choi et al. [7] introduced a liveness detection method based on multiple static features, derived from a single fingerprint image. These features are individual pore spacing, residual noise and several first order statistics. A correlation filter is adopted to address individual pore spacing. The selected features are useful to reflect the physiological and statistical characteristics of live and fake fingerprint.

In [2] Abhyankar et al. proposed the use of a multiresolution texture analysis technique to minimize the energy associated with phase and orientation maps. Cross ridge frequency analysis of fingerprint images was performed by means of statistical measures and weighted mean phase was calculated. These different 36 features along with ridge reliability or ridge center frequency were given as inputs to a fuzzy c-means classifier.

Tidu et al. [36] proposed two measures: the use of the number of pores and the mean distance between them to discriminate live and fake. In addition, a third measure has been proposed to introduce the image quality under the assumption that the pores number will be higher in good quality images and lower in bad ones. It is also been made a dynamic analysis comparing the number of pores and the mean distance between them in two images acquired at 0 and 5 second.

In [33] Marasco and Sansone propose a novel solution based on static features derived from visual textures of the image. These measures are obtained using signal processing methods (individual pore spacing, residual noise of the fingerprint image), first order statistics (energy, entropy, median, variance, skewness, kurtosis, coefficient of variation), intensity-based features (gray level 1 ratio, gray level 2 ratio).

Another set of features is extracted by Galbally et al. [15]. They used several sources of information: angle information provided by the direction field, Gabor filters, which represent another implementation of the direction angle, pixel intensity of the gray-scale image and power spectrum. They also assess the fingerprint quality analyzing the image in a holistic manner, or combining the quality from local non-overlapped blocks of the image.

Gottschlich et al. [21] propose a new invariant descriptor of fingerprint ridge texture called Histograms of Invariant Gradients (HIG). They were inspired by invariant feature descriptors such as the Histograms of Oriented Gradients (HOG) and the Scale Invariant Feature Transform (SIFT). This descriptor is designed to preserve robustness to variations in gradient positions. Spoofed fingerprints are detected using multiple histograms of invariant gradients computed from spatial neighborhoods within the fingerprint.

3.4.3 Textural algorithms in Fingerprint Liveness Detection

Since the publication of Nikam and Agarwa work [40] it was evident the ability of the LBP algorithm to capture the different characteristics of live and fake fingerprints. After that the use of many other textural algorithms was introduced in the FLD field.

Ghiani et al. [19] proposed the use of a rotation invariant extension of the Local Phase Quantization (LPQ). The LPQ is a blur insensitive texture classification method that works at low frequency values in the frequency domain. Its blur invariance is probably the reason of its ability to distinguish between live and fake samples. The obtained results, competitive with the state-of-the-art, where further improved by a feature level fusion with the LBP.

In [22], Gragnaniello et al. used the Weber Local Descriptor (WLD), it consists of two components (the differential excitation and the orientation) and depends not only on the change of a stimulus such as lighting, but also on the original intensity of that stimulus. The better performances were obtained with the combination of the WLD with the LPQ.

In 2013, Ghiani et al. [17] introduced the use of the Binarized Statistical Image Features (BSIF). It is a local image descriptor constructed by binarizing the responses to linear filters but, in contrast to other binary descriptors, the filters are learnt from natural images using independent component analysis (ICA). A quantization-based representation encodes the local fingerprint texture on a feature vector in a highly effective way.

Jia et al. [30] used a novel fingerprint vitality detection method based on Multi-Scale Block Local Ternary Patterns (MBLTP). Instead of a single pixel, The computation is not based on a single pixel value but on the average value of blocks. The ternary pattern is adopted to reflect the differences between the pixels and a selected threshold.

Again Jia et al. [29] proposed a spoof fingerprint detection method based on two different types of Multi-Scale Local Binary Pattern (MSLBP). In the first type they simply increased the radius of the LBP operator, in the second one they applied a set of filters to the image and then they applied the LBP operator in the fixed radius. Both types of MSLBP, with proper scales selection, outperformed many of the spoof fingerprint detection methods.

3.5 Fingerprint Liveness Detection Competition

The Department of Electrical and Electronic Engineering of the University of Cagliari and the Department of Electrical and Computer Engineering of the Clarkson University organize the Fingerprint Liveness Detection Competition (LivDet). The first edition was held in 2009 [35] and other two took place in 2011 [59] and 2013 [20]. The fourth edition will be hosted in September by BTAS 2015. The competition purpose is to assess the achievements of the state of the art in FLD and it is open to academic and industrial institutions. Once a signed consent form is obtained, each user has access to a link for downloading the training set of several datasets (three in 2009 and four in the following editions) is given. Each participant gives their preference on whether to enter as anonymous.

3.5.1 Experimental protocol and evaluation

Due to the wide variety of current liveness detection algorithms, the competition defines some constraints for the submitted algorithms:

1. Methods must output, for each image, a “liveness degree” ranging from 0 to 100 (e.g. posterior probability of “true” class).

2. A training set of fake and live fingerprint images will be made available to each participant, freely downloadable from the LivDet site after the participant registration. These images are a subset (25% in 2009 and a half in the following editions) of the entire data set.

3. Each submitted algorithm, as a Win32 console application, must follow the input and output sequence required.

4. Each submitted algorithm is tested using a withheld dataset that is the remaining part (75% in 2009 and a half in the following editions) of the entire data set.

Each submitted algorithm should have the following list of parameters: LIVENESS_XYZ.exe [ndataset] [inputfile] [outputfile] Where:

- LIVENESS_XYZ.exe is the executable name, where XYZ is the identification number of the participant.

- [ndataset] is the identification number of the data set to analyse (it changes for each edition depending on the dataset).

- [inputfile] is a text file with the List of images to analyse.

- [outputfile] is a text file with the output of each processed image, in the same order of [inputfile]. Given the image, the output is a posterior probability of the live class or a degree of “liveness” normalized in the range 0 and 100 (100 is the maximum degree of liveness, 0 means that the image is fake). In the case that the algorithm has not been able to process the image, the correspondent output must be -1000 (failure to enroll).

Each parameter, related to the data set configuration, must be set before submission. Each user can configure his algorithm by the training set available after registration. Only Win32 console applications with the above characteristics will be accepted for the competition. Participants may also publish the source code of their algorithm, but this is not mandatory.

The following parameters are adopted for the performance evaluation:

- Frej: Rate of failure to enroll.
- Fcorrlive: Rate of correctly classified live fingerprints.
- Fcorrfake: Rate of correctly classified fake fingerprints.
- Ferrlive: Rate of misclassified live fingerprints.
- Ferrfake: Rate of misclassified fake fingerprints.

The threshold value for determining liveness was set at 50. This threshold is used for the values given for Ferrfake and FerrLive.

3.5.2 Datasets

The 2009 dataset for the final evaluation is constituted of three sub-sets, which contain live and fake fingerprint images from three different optical sensors. Table 3.1 lists the scanners we used for data collection and table 3.2 the databases characteristics.

Images have been collected by a consensual approach using different materials for the artificial reproduction of the fingerprint (gelatin, silicone, play-doh). The training sets contain the 25% of the data, the remaining 75% is in the testing sets.

Data set	Sensor	Model No.	Resolution(dpi)	Image size
#1	Crossmatch	Verifier 300 LC	500	480x640
#2	Identix	DFR2100	686	720x720
#3	Biometrika	FX2000	569	312x372

Table 3.1: Device characteristics of the LivDet 2009 datasets.

Data set	Sensor	Live samples	Fake samples
#1	Crossmatch	2000	2000
#2	Identix	1500	1500
#3	Biometrika	2000	2000

Table 3.2: Characteristics of the LivDet 2009 datasets (live/fake number of fingers).

The 2011 dataset for the final evaluation is constituted of four sub-sets, which contain live and fake fingerprint images from four different optical sensors. Table 3.3 lists the scanners used for data collection and table 3.4 the databases characteristics.

The dataset consists of images from four different devices; Biometrika, Digital Persona, Italdata and Sagem. There are 4000 images for each of these devices, 2000 live images and 2000 spoof images (400 of each of 5 spoof materials). The spoof materials were gelatine, latex, PlayDoh, silicone and wood glue for Digital Persona and Sagem (400 each) and gelatine, latex, ecoflex (platinum-catalysed silicone), silicone and wood glue for Biometrika and ItalData (400,each). The dataset of 4000 images per scanner were divided into two equal datasets, training and testing. Live images came from 400 fingers from 50 people for Biometrika and ItalData datasets, 200 fingers representing 100 people for, Digital Persona dataset, and 112 fingers from 56 people for Sagem dataset. Spoof images come from approximately 100 fingers representing 50 people for the Digital Persona and Sagem Datasets and 81 fingers representing 22 subjects for the Biometrika and ItalData datasets. The spoof images were collected using the consensual method that was described earlier.

Data set	Sensor	Model No.	Resolution(dpi)	Image size
#1	Biometrika	FX2000	500	315x372
#2	Digital Persona	4000B	500	355x391
#3	ItalData	ET10	500	640x480
#4	Sagem	MSO300	500	352x384

Table 3.3: Device characteristics of the LivDet 2011 datasets.

The 2013 datasets consists of images from four different devices; Biometrika, Crossmatch, Italdata and Swipe. There are 4000 or more images for each of these devices as detailed in tables 3.5 and 3.6. The spoof materials used for this experiment were Body Double, latex,

Data set	Sensor	Live samples	Fake samples
#1	Biometrika	2000	2000
#2	Digital Persona	2000	2000
#3	ItalData	2000	2000
#4	Sagem	2000	2000

Table 3.4: Characteristics of the LivDet 2011 datasets (live/fake number of fingers).

PlayDoh and wood glue for Crossmatch and Swipe and gelatine, latex, ecoflex (platinum-catalysed silicone), modasil and wood glue for Biometrika and Italdata. The images were divided into two equal datasets, training and testing. Live images came from 300 fingers from 50 subjects for the Biometrika and Italdata datasets, 940 fingers representing 94 subjects for the Crossmatch dataset, and 1000 fingers from 100 subjects for the Swipe dataset.

Spoof images come from approximately 225 fingers representing 45 people for the Crossmatch and Swipe Datasets and 100 fingers representing 15 subjects for the Biometrika and Italdata datasets. The spoof images of two of the LivDet 2013 datasets (Crossmatch and Swipe) were collected using the cooperative method that was described earlier. The other two datasets (Biometrika and Italdata) were created using the non-cooperative method (latent fingerprints). This is the reason of the great difference between the error rates: a fake created from a latent fingerprint will be, in many cases, less similar to the original one then the one created with cooperation.

Data set	Sensor	Model No.	Resolution(dpi)	Image size
#1	Biometrika	FX2000	569	315x372
#2	Italdata	ET10	500	640x480
#3	Crossmatch	L SCAN GUARDIAN	500	800x750
#4	Swipe		96	208x1500

Table 3.5: Device characteristics of the LivDet 2013 datasets.

In each one of three edition, at the end of the competition, the entire dataset has been made available to the scientific community after the signing of an appropriate license agreement.

Data set	Sensor	Live samples	Fake samples
#1	Biometrika	2000	2000
#2	Italdata	2000	2000
#3	Crossmatch	2250	2250
#4	Swipe	2250	2250

Table 3.6: Characteristics of the LivDet 2013 datasets (live/fake number of fingers).

Chapter 4

Textural features for fingerprint liveness detection

Many popular local descriptors can be seen as statistics of labels computed in the local pixel neighborhoods through filtering and quantization. These methods describe each pixels neighborhood by a binary code which is obtained by initially convolving the image with a set of linear filters and then binarizing the filter responses. The bits in the code string correspond to the binarized responses of the different filters. In different computer vision problems these methods showed very good results. In contrast to global descriptors which compute features directly from the entire image, local descriptors represent the features in small local image patches. As already stated in the previous chapter, local descriptors and, especially, textural algorithms provided excellent performance in the FLD field. In this thesis we present the results obtained with the LBP, LPQ and BSIF algorithms.

4.1 Investigated methods

4.1.1 Multiresolution LBP

The LBP operator was first employed for two-dimensional textures analysis and excellent results were obtained by the version invariant with respect to grey level, orientation and rotation [45]. It extracts certain uniform patterns corresponding to microfeatures in the image. The histogram of these uniform patterns occurrence is capable of characterize the image as it combines structural (it identifies structures like lines and borders) and statistical (microstructures distribution) approaches.

In a circular neighborhood of each pixel of a grayscale image we define the texture T as:

$$T = t(g_c, g_0, ..., g_{P-1}) \tag{4.1}$$

This represents the distribution of the P surrounding pixels. g_c is the grayscale value of the selected pixel and g_p are the pixels in the circular neighborhood of a given radius R with $R > 0$. If g_c is in position $(0;0)$, then the P points g_p, with $p = 0, ..., P-1$, are in position $(-Rsin(2\pi p/P); Rcos(2\pi p/P))$.

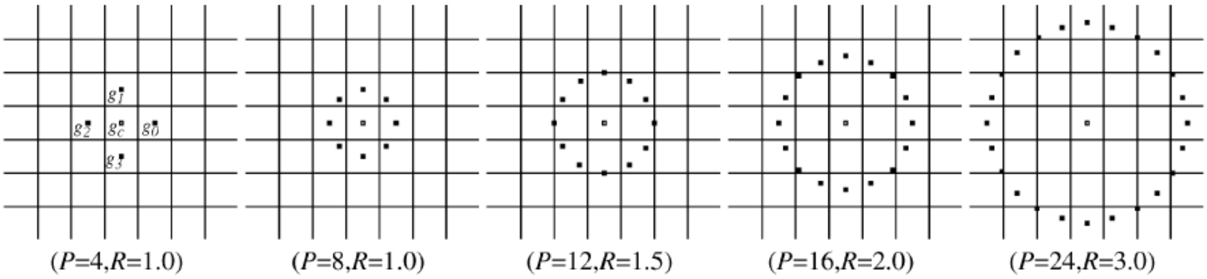


Figure 4.1: Examples of the points selected at different values of the position P and the radius R [45].

The interpolation is used in order to estimate the values of those points that are not precisely in the center of a pixel. Figure 4.1 shows examples of the points selected at different values of the position P and the radius R.

If we subtract the central value from the circular neighborhood values we obtain:

$$T = t(g_c, g_0 - g_c, g_1 - g_c, ..., g_{P-1} - g_c) \tag{4.2}$$

Assuming that $g_p - g_c$ values are independent from g_c:

$$T \approx t(g_c)t(g_0 - g_c, g_1 - g_c, ..., g_{P-1} - g_c) \tag{4.3}$$

Since $t(g_c)$ in 4.3 describes the overall luminance of the image, unrelated to local image texture, much of the information is contained in:

$$T \approx t(g_0 - g_c, g_1 - g_c, ..., g_{P-1} - g_c) \tag{4.4}$$

If we consider the signs of the differences instead of their exact values then invariance with respect to the gray levels scaling is achieved:

$$T \approx t(s(g_0 - g_c), s(g_1 - g_c), ..., s(g_{P-1} - g_c)) \tag{4.5}$$

with:

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{4.6}$$

A unique $LBP_{P,R}$ value can be obtained by assigning the factor 2^p for each sign $s(g_p - g_c)$ resulting in 2^p different binary patterns:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \tag{4.7}$$

In order to achieve rotation invariance namely to assign a unique identifier to each rotation invariant local binary pattern we define:

$$LBP_{P,R}^{ri} = min\{ROR(LBP_{P,R,i}) \quad | \quad i = 0, 1, ..., P-1\} \tag{4.8}$$

$ROR(x, i)$ rotates the neighbor set clockwise so many times that a maximal number of the most significant bits, starting from g_{P-1}, is 0.

The number U of spatial transitions (bitwise 0/1 changes) in the neighborhood pixels sequence is a measure of uniformity. Patterns that have a U value of at most 2 are defined "uniform" and the following operator is used:

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & if \quad U(LBP_{P,R}) \leq 2 \\ P+1 & otherwise \end{cases} \tag{4.9}$$

were:

$$U(LBP_{P,R}) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)| \tag{4.10}$$

The $LBP_{P,R}^{riu2}$ output values are $P+2$ instead of the 2^p obtained with the $LBP_{P,R}$. These values, extracted for each pixel of the image, are inserted in an histogram that is used as a feature vector.

In our experiment we will use the rotation invariant version with three different combination of the (P,R) values: $(8,1)$, $(16,2)$ and $(24,3)$. The three histograms are then united in a single feature vector of $10+18+26 = 54$ values.

4.1.2 Rotation Invariant LPQ

The Local Phase Quantization (LPQ) is a blur insensitive texture classification method [24]. It is able to represent all spectrum characteristics of images in a very compact feature representation, thus avoiding redundant or blurred information. The main reason of proposing this approach is to point out the spectrum differences between a "live" fingerprint and a "fake" one. Since different fingerprint orientations may arise on a sensor surface, we adopt the rotation invariant extension of LPQ.

Image blurring $g(\mathbf{x})$ can be expressed by a 2-D convolution between the original image $f(\mathbf{x})$ and the point spread function (PSF) of the blur $h(\mathbf{x})$, where the vector $\mathbf{x}$ represents the coordinates (x, y). In the frequency domain, the convolution become the product: $G(\mathbf{u}) = F(\mathbf{u}) \cdot H(\mathbf{u})$, where $\mathbf{u}$ is the frequency and $G(\mathbf{u})$, $F(\mathbf{u})$, and $H(\mathbf{u})$ are discrete Fourier transforms (DFT). If we just consider the phase of the spectrum, we obtain the sum: $\angle G = \angle F + \angle H$.

If the PSF is centrally symmetric, $\angle H \in \{0, \pi\}$ as the Fourier transform H is always real and, usually, its shape is close to a Gaussian or a sinc function, hence H is positive at low frequency values. In that frequency interval, $\angle H = 0$ and $\angle G = \angle F$ proving that the phase is blur invariant.

For every pixel $\boldsymbol{x}$, we compute the local spectra using a short term Fourier transform (STFT) in the local neighborhood N_x (defined by a rectangular window function ω_R), obtaining:

$$F(\mathbf{u}, \mathbf{x}) = \sum_{\mathbf{y}} f(\mathbf{y}) \omega_R(\mathbf{y} - \mathbf{x}) e^{-j2\pi \mathbf{u}^T \mathbf{y}} \tag{4.11}$$

That is a blur-insensitive representation, with four low frequency components: $\mathbf{u}_1 = [a, 0]^T$, $\mathbf{u}_2 = [0, a]^T$, $\mathbf{u}_3 = [a, a]^T$, $\mathbf{u}_4 = [a, -a]^T$, only if a is small enough to satisfy $H(u_i) > 0$. For each point x we can write:

$$\mathbf{F}(\mathbf{x}) = [F(u_1, \mathbf{x}), F(u_2, \mathbf{x}), F(u_3, \mathbf{x}), F(u_4, \mathbf{x})] \tag{4.12}$$

Given the vector $\mathbf{G}(\mathbf{x}) = [Re\{\mathbf{F}(\mathbf{x})\}, Im\{\mathbf{F}(\mathbf{x})\}]$, from his j-th component g_j:

$$q_j = \begin{cases} 1, & \text{if } g_j \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{4.13}$$

We can write these eight binary coefficients in the form of an integer value included between 0 and 255 through:

$$F_{LPQ}(\mathbf{x}) = \sum_{j=1}^{8} q_j 2^{j-1}$$

From all of these values (one for every pixel of the image), we obtain an histogram that can be represented as a 256 features vector.

In the rotation invariant version of LPQ we take advantage of the fact that, given a rotation matrix $\boldsymbol{R_\theta}$, the Fourier transform of a rotated function is the Fourier transform of the original function rotated by $\boldsymbol{R_\theta}$. Instead of a rectangular window, we use a circular Gaussian one and the coefficients of the local spectra 4.11 on a radius r around the point $\boldsymbol{x}' = \boldsymbol{R_\theta x}$ are calculated at frequencies $\boldsymbol{v}_i = r[cos(\phi_i)sin(\phi_i)]^T$, with $\phi_i = 2\pi i/M$ and $i = 0, ..., M-1$.

From the vector $\boldsymbol{V}(\boldsymbol{x}) = [F(v_0, \boldsymbol{x}), ..., F(v_{M-1}, \boldsymbol{x})]$, we calculate $\boldsymbol{C}(\boldsymbol{x}) = Im\{\boldsymbol{V}(\boldsymbol{x})\}$ and then we extract the characteristic orientation $\xi(x) = \angle b(x)$ from the complex moment:

$$b(x) = \sum_{i=0}^{M-1} c_i e^{j\varphi_i} \tag{4.14}$$

Instead of 4.12, we use the oriented frequency coefficients:

$$F_\xi(\boldsymbol{u}, \boldsymbol{x}) = \sum_{y} f(\boldsymbol{y}) \omega_R (R_{\xi(x)}^{-1}(\boldsymbol{y} - \boldsymbol{x})) e^{-j2\pi \boldsymbol{u}^T R_{\xi(x)}^{-1} \boldsymbol{y}}$$

If we apply a rotation, the position of the coefficients changes, but the 256-value histogram is the same (rotation invariant LPQ).

4.1.3 BSIF

Local image descriptors make the backbone of the current approaches for visual object recognition. The function of descriptors is to convert the pixel-level information into a useful form, which captures the most important image and video contents but is insensitive to irrelevant aspects caused by varying environment. In contrast to global descriptors which compute features directly from the entire image, local descriptors, which have proved to be more effective in real world conditions, represent the features in small local image patches.

Many popular local descriptors such as LBP [45] and LPQ [24] can be seen as statistics of labels computed in the local pixel neighborhoods through filtering and quantization. These methods describe each pixel's neighborhood by a binary code which is obtained by first convolving the image with a manually predefined set of linear filters and then binarizing the filter responses. The bits in the code string correspond to binarized responses of different filters. These methods showed very good results in different computer vision problems [47].

For efficiently representing fingerprint images for liveness detection, we adopt a new local descriptor called BSIF (binarized Statisitcal Image features) which was recently proposed

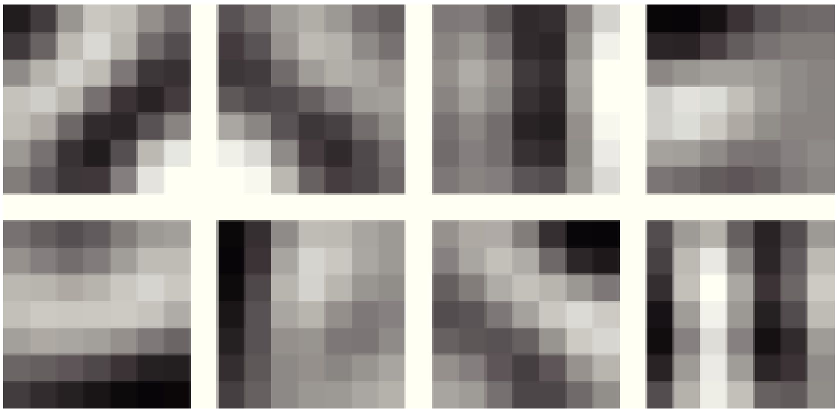


Figure 4.2: Learnt filters of size 7 × 7.

by Kannala and Rahtu for face recognition and texture classification [31]. Inspired by LBP and LPQ, the idea behind BSIF is to automatically learn a fixed set of filters from a small set of natural images, instead of using hand-crafted filters such as in LBP and LPQ. Our proposed approach for fingerprint representation consists of apply learning, instead of manual tuning, to obtain statistically meaningful representation of the fingerprint data, which enables efficient information encoding using simple element-wise quantization. Learning provides also an easy and flexible way to adjust the descriptor length and to adapt to applications with unusual image characteristics such as fingerprints.

To characterize the texture properties within each fingerprint sub-region, the histograms of pixel's BSIF code values are then used. The value of each element (i.e. bit) in the BSIF binary code string is computed by binarizing the response of a linear filter with a threshold at zero. Each bit is associated with a different filter and the desired length of the bit string determines the number of filters used. The set of filters is learnt from a training set of natural image patches by maximizing the statistical independence of the filter responses [25]. The details for computing the filters can be found in [31]. The set of the natural images that are used in the calculation of the filters are shown in Figure 4.3.

Figure 4.3: The set of the natural images that are used in the calculation of the BSIF filters.

Given an image patch X of size $l \times l$ pixels and a linear filter W_i of the same size, the filter response s_i is obtained by:

$$s_i = \sum_{u,v} W_i(u,v)X(u,v) = w_i^T x,$$

where vectors w and x contain the pixels of W_i and X.

The binarized feature b_i is obtained by setting $b_i = 1$ if $s_i > 0$ and $b_i = 0$ otherwise.

The filters W_i are learnt using independent component analysis (ICA) by maximizing the statistical independence of s_i.

To sum up:

- a certain number of "templates" in the spatial frequency space is set, let's say n. These templates correspond to filters previously computed. In the case of LPQ, these filters are computed over four pre-defined frequency values; in the case of BSIF, they are learnt by ICA from a set of natural images. Size and number of filters may vary depending on how many textural details must be pointed out in the image to be coded by BSIF. In other words, input to ICA is the set of natural images, and the related output is the set of filters (corresponding to the independent components generating those images).

- A fingerprint image X is subdivided in patches, whose size corresponds to that of adopted filters. For each patch, the set of filters is applied and each response is binarized. At the end of this step, each patch is coded into a string of bits whose size is equal to the bit length n. This string embeds the main textural characteristics of the related patch (for example, in the case of three bits/filters, the code 010 means that the patch is positively correlated with the second filter).

- At this point, a unique bit string is associated to each patch. Different patches may share the same string. The final step is to compute the frequency of all possible configurations of n bits over all patches, thus obtaining the BSIF feature vector whose length is 2^n. The value of j-th component of such feature vector is the number of occurrences of the related binary string corresponding to the decimal value $j-1$. This is also the reason for which we referred to those feature vectors as "histograms".

In our experiments, we used the set of filters provided by the authors of [31] and learnt from the set of 13 natural images previously shown. There are two parameters in the BSIF descriptor: the filter size l and the length n of the bit string. The filters W_i were learnt using different choices of parameter values, each set of filters was learnt using 50000 image patches. The filters obtained with $l = 7$, $n = 8$ are illustrated in Figure 4.2.

4.1.4 Multi-Scale BSIF

One of the unsolved problems in the use of BSIF is the setting of the parameters, namely the filters number (bit number) and their size (window size). Regarding the number of filters, a rule of thumb is the more is the better, but there is the problem of the feature vector size. In our experiments we have never used more than 12 filters of a given window size resulting in a 4096 elements feature vector. Although previous work [17] indicates that, in the FLD case, a higher accuracy has usually produced through filters of a relatively small size, this is not always true. Furthermore, since the performance produced by filters of different sizes do not vary much, our hypothesis is that filters of different sizes actually capture different live and fake characteristics. Therefore the features generated by filters of different size would be, at least partially, complementary. In addition to the accuracy that we will show in the experimental results, one of the key points of the BSIF is the filters speed in the features extraction. Parallel computation based on the use of multiple windows could further improve the performances.

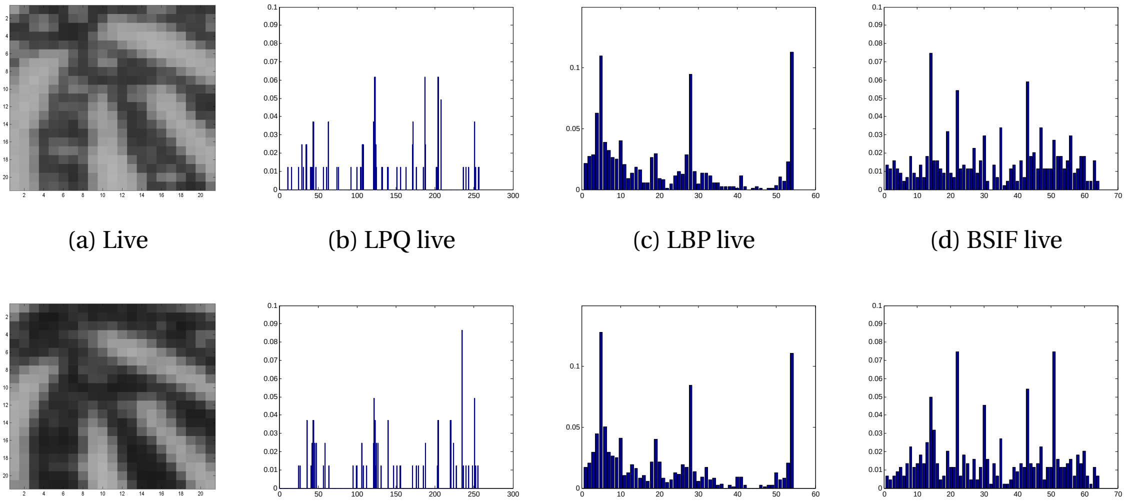
For these reasons it is possible to apply filters of different dimensions and use different fusion techniques both at a features and at a score level. In this thesis we simply fuse together the feature vectors extracted with the BSIF using 12 bits and all the odd windows size from $5x5$ to $17x17$. Before the fusion, in order to reduce the high dimensionality of the feature vector, we performed a feature reduction with the PCA (Principal Component Analysis) [14].

The PCA is performed on a matrix containing all the extracted feature vectors. The columns of the new matrix are sorted in order of decreasing eigenvalues. The cumulative energy content of the first c columns is the sum of the first c eigenvalues:

$$Energy(c) = \sum_{i=1}^{c} eig[i] \tag{4.15}$$

In our experiments we reduced the features using the 90%, 95% and 99% of the cumulative energy content. In the next chapter we will show the obtained experimental results.

4.2 Textural characteristics of a fingerprint



(a) Live

(b) LPQ live

(c) LBP live

(d) BSIF live

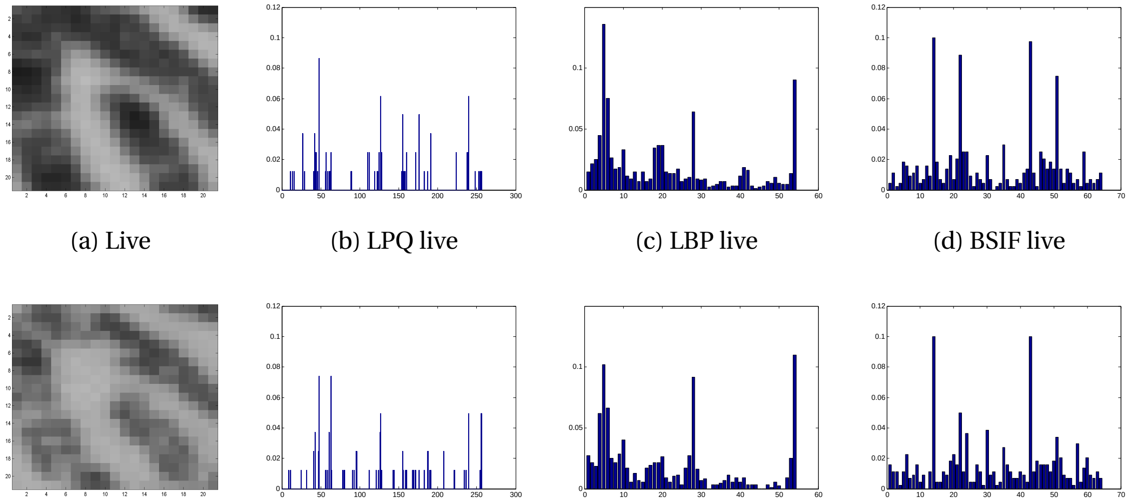
(e) Fake

(f) LPQ fake

(g) LBP fake

(h) BSIF fake

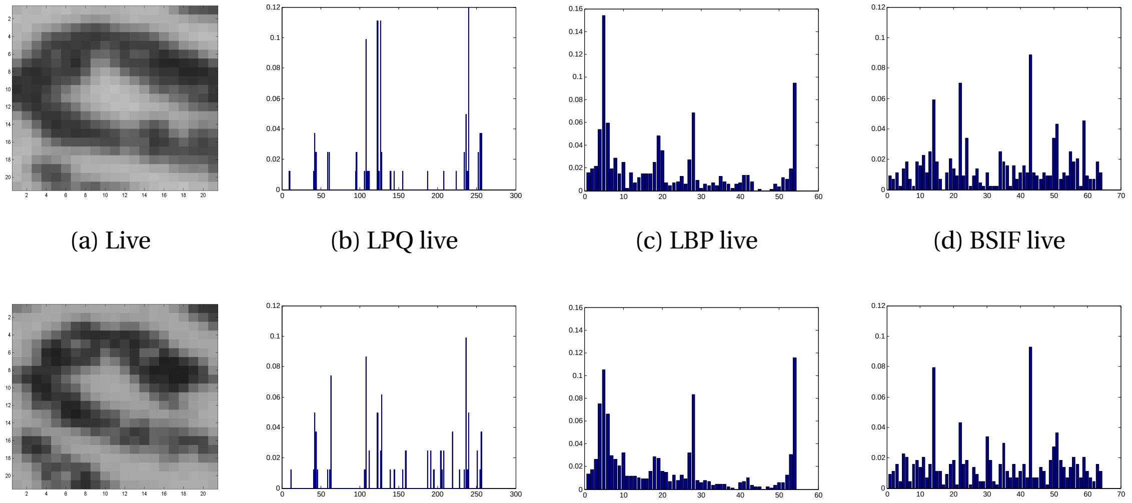
Figure 4.4: ROI of a live fingerprint (a) and corresponding fake (e). Histograms of the features extracted with LPQ, LBP and BSIF from the ROI of the live fingerprint (b,c,d) and from the ROI of the corresponding latex fake (f,g,h).



(a) Live (b) LPQ live (c) LBP live (d) BSIF live

(e) Fake (f) LPQ fake (g) LBP fake (h) BSIF fake

Figure 4.5: ROI of a live fingerprint (a) and corresponding fake (e). Histograms of the features extracted with LPQ, LBP and BSIF from the ROI of the live fingerprint (b,c,d) and from the ROI of the corresponding wood glue fake (f,g,h).



(a) Live (b) LPQ live (c) LBP live (d) BSIF live

(e) Fake (f) LPQ fake (g) LBP fake (h) BSIF fake

Figure 4.6: ROI of a live fingerprint (a) and corresponding fake (e). Histograms of the features extracted with LPQ, LBP and BSIF from the ROI of the live fingerprint (b,c,d) and from the ROI of the corresponding gelatine fake (f,g,h).

The comparison between the histograms obtained with different algorithms from images of real and fake samples of the same fingerprint can help to better understand their functioning. As shown in figures 4.4-4.6 (a) and (e), we only considered as a ROI (Region Of Interest) a square constructed around the core of the fingerprints. The histograms in figures 4.4-4.6

(b,f), (c,g) and (d,h) are extracted respectively with the LPQ (256 values), LBP (54 values) and BSIF with 6 bits and 5x5 filters (64 values).

It can be observed, from the feature vector values, that all the algorithms are able to capture some invariant characteristics of the fingerprint shape. An example is the three peaks of all the LBP histograms. In particular, the peak on the fifth bin indicates a large number of the textural templates corresponding to an edge, that is, the textural template corresponding to the fifth value of the LBP feature vector. The distribution of other peaks clearly shows common characteristics among fingerprints, which are captured by the LBP descriptor. Unfortunately it is not easy to point out features which establish the difference between live and fake images (figures 4.4-4.6(c) and 4.4-4.6(g)), but what reported testifies that LBP is a good fingerprint descriptor. Similar results are obtained by the BSIF (figures 4.4-4.6(d) and 4.4-4.6(h)). On the other hand, LPQ-based live and fake feature vectors are strongly different, but we can also see a notable intra-class difference among feature vector values (figures 4.4-4.6(b) and 4.4-4.6(f)). It is almost impossible to identify specific characteristics captured by LPQ, since the distributions of the frequency of textural templates appear as completely random. However, its high performance shows that the LPQ feature space separates well live and fake patterns by the phase changes at low frequencies.

These examples also suggest the presence of common characteristic between different samples of the same finger, no matter if real or fake. The presence of this information, not related to the fact that the fingerprint is true or fake, could be considered a sort of noise in FLD. In the next section we will further analyze what we will call the "user-specific" problem.

4.3 The "User-Specific" problem

The feature set aimed at liveness detection is supposed to be able to measure the differences among "live" and "fake" fingerprint independently on the set of users at hand and, obviously, on the one who will be eventually targeted by potential attackers. Therefore, intrinsic characteristics of the users' fingerprint should be avoided by these measurements, in order to avoid poor sampling representation and "polarized" classification. On the basis of this hypothesis, the problem is expressed by a bayesian network where the state of nature is two-classes ($live/fake$ or $live/\overline{live}$), and the measurement, usually a feature vector, is generated from this state. The module is therefore managed individually. This aspect impacts on the design phase: the fingerprint liveness detector is designed to obtain a certain performance independent on the so-called user population targeted by the fingerprint verification system. It should guarantee, for instance, the same "ROC curve" on very different user populations. This "generic-user" approach can be reasonable when the system must be used alone (for example in forensic applications or border checks), where it is supposed to reveal a fraud trial, if any. On the other hand, when integrated in a fingerprint verification system, the liveness detector should protect the system from very peculiar attacks, against the targeted clients stored in the system's memory. Therefore, in this case, the system should be equipped with a "user-specific" fingerprint liveness detector. This view has been already pointed out in two recent works [6, 60] dealing with face liveness detection. In particular, the authors proposed several declinations of such a system: a generative approach in which live and fake samples are represented as user-driven clusters; a discriminative approach when live and fake samples of the targeted client are used to train one or more classifiers, and a domain adaption approach when some fake samples of a certain user are missed.

Starting from above previous works and observations, we will show that the user-specific information may help improving the performance of a fingerprint liveness detector. For this purpose, we will adopt three state-of-the-art feature sets (LBP [45], LPQ [56] and BSIF [31]) and two classifiers [14] (KNN with $K = 1$ and linear SVM) in order to show and explain this phenomenon. We don't have the ambition to state that user-specific is an intrinsic "characteristic" of all possible features sets but only to show that in the observed cases this phenomenon exists and can be exploited.

In the next chapter we will show evidences of this "user-specific effect" presence and we will also take advantage of this effect in order to improve the accuracy of a FLD system.

Chapter 5

Experimental Results

5.1 Data sets

All the presented algorithms were tested using the data sets collected for the three editions of the International Fingerprint Liveness Detection Competition (LivDet 2009, 2011 and 2013). In 2009 the data sets were collected by three optical sensors (Biometrika, Crossmatch and Identix) [35], in 2011 by four optical sensors (Biometrika, Italdata, Digital Persona, Sagem) [59] and in 2013 by two optical sensors (Biometrika and Italdata) and by a capacitive sensor (Swipe) [20]. Each dataset is divided into two parts, one used to train a classifier and the other to test the classifier performances (see Fig. 5.1). Device and datasets characteristics were previously shown in Chapter 3. The overall number of analyzed live and fake fingerprint images is 40,000.

Almost all of the fakes were created using a consensual method: a volunteer put his finger on a mould of plasticine like material and afterwards another material such as gelatine or liquid silicon is poured over the mould. The result, after a certain time interval, is an artificial replica of the fingertip. The only exception was the Biometrika and Italdata images of the LivDet 2013 data sets that were obtained acquiring fakes created with a non consensual method: a latent fingerprint left on a surface is enhanced, digitized through the use of a camera, and, finally, the negative image is printed on a transparency sheet, over which the fake material is poured.

5.2 Experimental protocol

As stated in the previous chapter, we identified a “user-specific effect” presence. In order to improve the accuracy taking advantage of this effect we had to adopt a different experimental protocol with respect to the one usually adopted in FLD. For this reason, hereinafter, will distinguish the latter case, that we will call “General-Purpose”, from the former, namely the “user-specific” case, which will be discussed in a separate section.

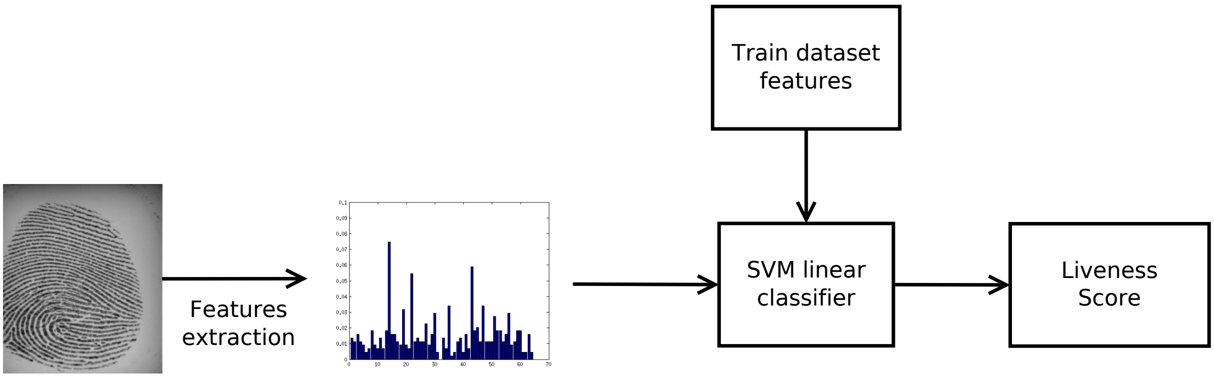


Figure 5.1: Classification of a fingerprint in a liveness detection system.

5.2.1 The “General-Purpose” case

For the BSIF algorithm, we used 60 different sets of learnt filters for all possible combinations of window size (odd pixel squares from 3x3 to 17x17) and number of bits (from 5 to 12). The only exception is the 3x3 window for which the number of bits ranges from 5 to 8. As a matter of fact, the space of 3x3 patches is 9 dimensional, but discarding the constant ("DC") component (discovered by ICA) leaves 8 dimensions, and hence there are eight filters representing the independent components. The spaces of larger patches are higher dimensional and more independent components can be found. Due to space limitations and to the fact that a greater number of filters prove to produce better results, in tables we will only show results obtained with a bit number equal to 12. Therefore all the feature vectors will contain $2^{1}2 = 4096$ values. Both LBP and LPQ have been optimized in performance with respect to data sets adopted, thus their performance is fairly comparable with that of BSIF. The LBP feature vector contain 54 values while the LPQ one contains 256 values.

In order to compare our results with the state-of-the-art, we also reported those presented in [29], [22], [15] and [34].

The classifier adopted in all our tests (with BSIF, LBP and LPQ) was a linear Support Vector Machine[1]. The output of each SVM, a real value included in the $[0,1]$ interval, is interpreted as *a posteriori* probability of the “Live" class given the input pattern. By thresholding this value we are able to obtain the performance at different operational points. In particular, we computed:

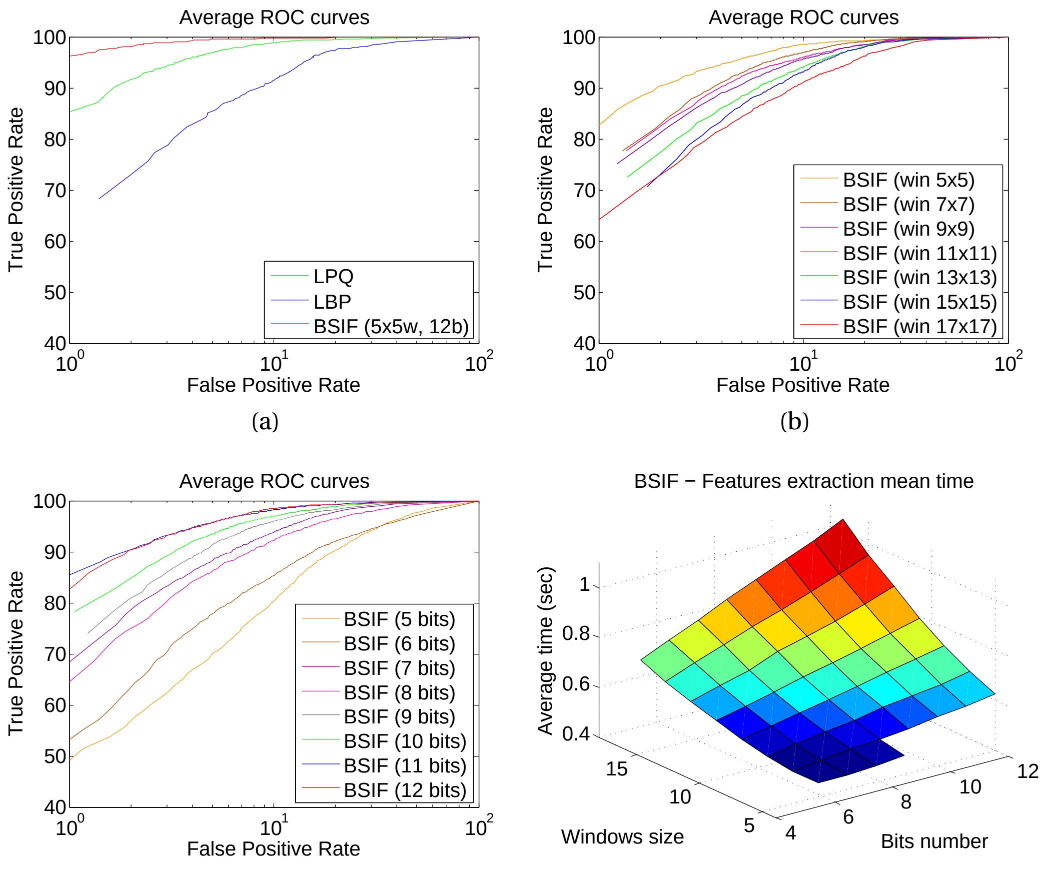
- Ferrlive: the rate of misclassified live fingerprints.
- Ferrfake: the rate of misclassified fake fingerprints.
- The Average Classification Error (ACE) which is defined as:

  $ACE = (Ferrlive + Ferrfake)/2.$

The threshold value for determining liveness was set at 0.5.

[1] This classifier gave the best results for all investigated feature vectors, including LBP and LPQ and other textural-based algorithms, as also reported in [18, 19, 16]. Moreover, by tuning the classifier, the performance difference among methods are substantially unaltered, thus we reported results only for the linear SVM classifier.

5.3 “General-Purpose” experimental results



(a)

(b)

(c)

(d)

Figure 5.2: Average ROC Curves for the BSIF with a (5x5) window size and 12 bits compared with LPQ and LBP (a), with different window sizes and best bit length (12 bits) (b), with different bit length and best window sizes (5x5 pixels) (c) and BSIF features extraction times (sec) in a 3D graph (d).

A full overview of the performance of the classifier is given by the ROC curves, which plot the Ferrfake-Ferrlive pair for all possible threshold values in the range $[0,1]$. For sake of space, since the curves are very similar between them, we only present the average results obtained from all ten datasets (using the average values of Ferrfake and Ferrlive). In particular we plotted:

- Figs. 5.2 (a): average LPQ, LBP and best BSIF ROC in order to compare our proposal against state-of-the-art algorithms. “Best" BSIF means that we selected, for each data set, the feature vector that was extracted with a 5x5 windows size and a 12 bit length.
- Figs. 5.2 (b): average BSIF ROCs by setting the best value of bits number (12 bits) and varying the window size, for each data set, in order to evaluate the dependency of BSIF performance on the window size parameter.

- Figs. 5.2 (c): average BSIF ROCs by setting the best value of window size (5x5 window) and varying the bits number, for each data set, in order to evaluate the dependency of BSIF performance on the bits number parameter.

	Mean	Standard Deviation
LBP	1.79	0.83
LPQ	4.43	2.66
BSIF	0.58	0.34

Table 5.1: LBP, LPQ and BSIF (5x5 win, 12 bits) mean and standard deviation of the computational time (sec).

Finally, Fig. 5.2(d) shows the computational time as function of the BSIF parameters. An average value is also given in Tab. 5.1, and compared with that of LPQ and LBP. It can be noticed that it increases with both window size and bits number, but this second parameter is the most relevant. However, it is noticeable that computational time is much better than that of LPQ, comparable with that of LBP in the worst case, and considerably less when considering the best BSIF configuration (Tab. 5.1).

All the results have been computed on a platform based on Matlab 7.9.0 r2009b, Windows 7 PRO 64 bit, Pentium(R) Dual-Core CPU ES200 @ 2.50GHz 8 GB[2].

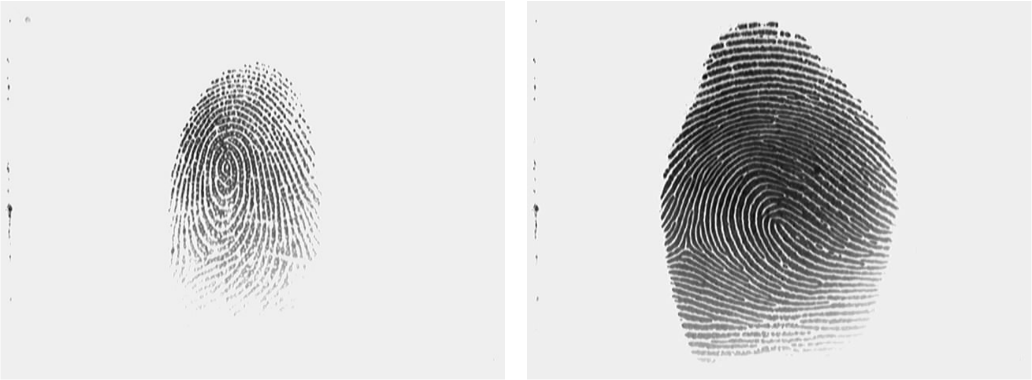


Figure 5.3: Some fingerprint images collected with the Italdata sensor.

From Figs. 5.2 (a), it is evident that the BSIF performance is superior than that of the LPQ and the LBP algorithms. The only exception is given by the performance on the Italdata data

[2] Matlab, Windows, Pentium are property of Mathworks, Microsoft, Intel, respectively. Source codes of LBP, LPQ and BSIF is available at the following website: http://www.cse.oulu.fi/CMV/Research.

sets (Tab. 5.3), which appears to be only slightly better than of the LPQ. This can be explained by the fact that Italdata images are automatically pre-processed during the acquisition step which is impossible to be avoided. Therefore, we can suppose a certain loss of details making Italdata images apparently "cleaner" than those acquired with other sensors, as shown in Fig. 5.3, with a sort of visible "blurring effect" that helps explaining why a blurring invariant algorithm like LPQ equates the BSIF in this case [56]. In general, the very good performance of BSIF confirmed the results reported in [17].

Algorithm	Biometrika	Italdata	Swipe
MSBSIF (90% energy)	1.00	0.50	3.05
MSBSIF (95% energy)	1.05	0.50	3.20
MSBSIF (99% energy)	1.05	0.50	3.65
BSIF	0.55	0.60	4.76
LBP [45]	1.30	3.35	18.96
LPQ [56]	2.15	1.45	9.65

Table 5.2: Average Classification Error (ACE) values comparison for the LivDet 2013 datasets. The BSIF features were extracted with a 5x5 windows size and a 12 bit length.

Tables 5.2-5.4 point out that the BSIF outperform all of the others state-of-the-art algorithms with the only partial exception of the two Multi-Scale LBP algorithms on the 2011 datasets (Table 5.3). As a matter of fact the MSLBP1 performances are much better than those of the BSIF for the Digital dataset and slightly better for the Sagem dataset, the MSLBP2 performances are similar to those of the BSIF for the Italdata dataset and slightly better for the Sagem dataset. For this reason we extended our work on the BSIF focusing on the fusion of feature vectors extracted at different window sizes as already done with the original LBP algorithm. In the first three rows of Tables 5.2-5.4 we inserted the results obtained using the fusion of features after a PCA features reduction using as stopping criterion the cumulative energy content at 90%, 95% and 99%.

Results show a general error reduction with respect to the BSIF algoritm with the only exception of the Biometrika sensor in 2013 (Table 5.2) and in 2009 (Table 5.4). In both cases this is probably due to the fact that the results for the different window sizes are unbalanced with those obtained for a 5x5 window size much better than the others. Therefore, instead of an improvement, we notice a deterioration of the results.

Algorithm	Biometrika	Italdata	Digital	Sagem
MSBSIF (90% energy)	4.65	9.70	1.70	3.90
MSBSIF (95% energy)	4.45	9.85	1.70	3.80
MSBSIF (99% energy)	4.40	9.60	1.80	3.80
BSIF	6.15	12.60	4.00	5.91
LBP [45]	11.20	18.45	10.60	8.51
LPQ [56]	14.70	13.60	11.45	8.01
MSLBP1 [29]	7.30	14.80	2.50	5.30
MSLBP2 [29]	10.60	12.60	6.70	5.60
WLD [22]	13.25	27.67	13.75	6.66

Table 5.3: Average Classification Error (ACE) values comparison for the LivDet 2011 datasets. The BSIF features were extracted with a 5x5 windows size and a 12 bit length.

Tabs. 5.2 and 5.3 also point out a strong performance difference on Biometrika and Italdata data sets between edition 2011 and 2013 of LivDet which can be explained by the fact that these data sets for LivDet 2013 were made up of fake fingerprint images provided by latent prints. As a matter of fact, the third edition of LivDet [20] was also aimed to compare liveness detectors when facing more realistic spoofing trials. However also in this case it can be noted that BSIF performance is far better than that of LBP and LPQ.

Figs. 5.2(b-c) show the BSIF performance as functions of the windows size and bits number parameters. With regard to the window size (Fig. 5.2(b)), the performance trend appears as generally decreasing with it: the smaller the window size, the better the performance. However, it should be observed a performance decrease, beyond a certain size. This means that each data set/sensor requires a careful calibration of this parameter, due to the difference of the images captured.

With regard to the bits number (Figs 5.2(c)), the trend is strongly increasing: the higher

Algorithm	Biometrika	CrossMatch	Identix
MSBSIF (90% energy)	6.86	3.33	0.78
MSBSIF (95% energy)	6.86	3.40	0.51
MSBSIF (99% energy)	6.96	3.50	0.42
BSIF	3.45	4.58	1.02
LBP [45]	19.09	10.25	4.04
LPQ [56]	6.51	5.98	2.27
WLD[22]	1.16	5.70	2.00
QRF [15]	37.40	14.80	4.90
PMBF [34]	12.60	15.20	9.70

Table 5.4: Average Classification Error (ACE) values comparison for the LivDet 2009 datasets. The BSIF features were extracted with a 5x5 windows size and a 12 bit length. QRF stands for Quality Related Features [15] and PMBF for Perspiration- and Morphology-Based Features [34]

the number the better the performance. Among the analyzed datasets there are no exceptions for this trend. It seems that the optimal number of bits is 10-12. This is motivated by the fact that the number of bits defines the expressive power of the BSIF feature vector because it represents the number of filters as well as the number of possible filter configurations responses, that is, the textural configurations on each subimage.

As in the case of the window size, we have a similar behavior for the number of bits. As a matter of fact, it could be supposed that, by increasing this number, we should always obtain a correspondent performance improvement. But Fig. 5.4 clearly shows a sort of saturation effect beyond a certain bits number. Therefore, the number of bits points out an intrinsic

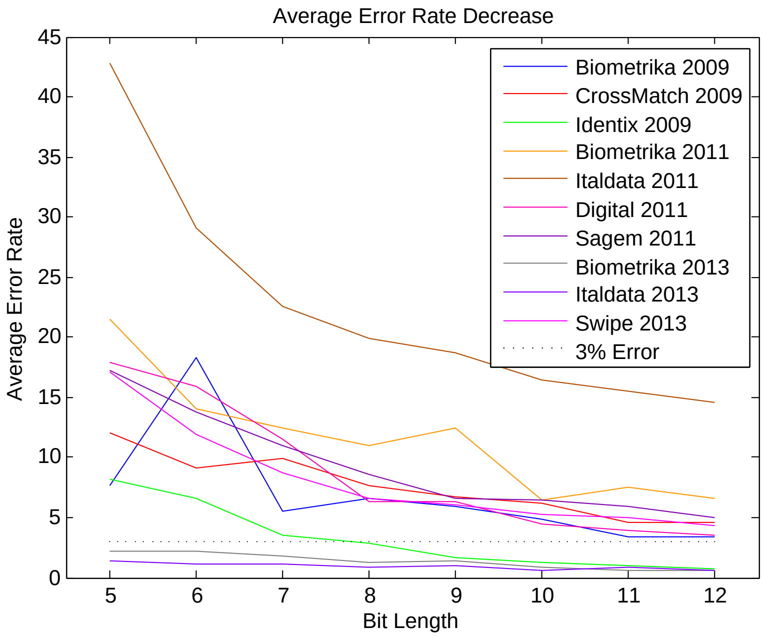


Figure 5.4: Decrease of the Average Error Rate as the bit length grows.

limitation of the BSIF method. We suppose that this “limit” is correlated with the resolution of each image, but we have not yet a clear evidence about that, since all LivDet sensors are characterized by the same resolution in terms of dpi. In other words, we believe that a higher resolution could lead to a higher "optimal" number of bits, due to the increase of details detectable in the image.

In our opinion, the very good performance of BSIFs, and the clear indication that this is obtained by using a very small window size (5x5 pixels) with a large number of bits are the most concrete evidence that differences between live and fake fingerprints are embedded in minute details that cannot be visible even by human experts. Filters learnt from natural images allow capturing many of these details, thus leading to a performance far better than that of other algorithms which use a less flexible approach to derive filters and textural templates.

5.4 “User-Specific” experimental results

Results presented in the previous section prove that FLD is still an open issue. We will show by experiments that the user-specific approach can be of great help in the improvement of a detector performances. In our analysis we first bring some evidence of this effect and then we propose a method that takes advantage of it.

5.4.1 Generic-user vs. user-specific Fingerprint Liveness Detection

In this Section, we show by experiments the user-specific effect using a simple discriminative approach to FLD. Given the feature vector, we trained a Nearest Neighbour (NN) classifier and a linear Support Vector Machine (SVM) according to the standard generic-user method. In parallel, we trained the same classifiers using the user-specific method, namely, the training set is made up of live and fake samples of the targeted user population. The adopted data sets are the four ones of the Second Edition of the International Fingerprint Liveness Detection (LivDet2011 [59]). Let's indicate the client set in a certain LivDet data set with Ω . Firstly, we randomly partitioned Ω into Ω_1 and Ω_2, with $\Omega_1 \cap \Omega_2 = \phi$, and $\Omega_1 \cup \Omega_2 = \Omega$. Let's extract all live and fake samples belonging to clients in Ω_1, thus generating the subset of images D_1. The same is done with Ω_2, thus generating D_2. Finally, let's partition D_1 in $D_1^{(1)}$ and $D_1^{(2)}$ such that different samples of clients in Ω_1 are present. We obtained three sets, which are intended as follows:

- $D_1^{(1)}$ is the user-specific training set, made up of samples from the targeted population, namely, the one in Ω_1. This set is used for training the user-specific classifiers;
- D_2 is the generic-user training set, made up of samples from another, generic population. This set is used for training the generic-user classifiers;
- $D_1^{(2)}$ is the user-specific test set, made up of samples from the targeted population Ω_1. This set is used to test the performance of classifiers above.

BIOMETRIKA	KNN (k=1)	SVM (lin.)
External train	19.99	16.06
Internal train	5.73	10.71

ITALDATA	KNN (k=1)	SVM (lin.)
External train	33.02	25.24
Internal train	9.43	13.91

DIGITAL	KNN (k=1)	SVM (lin.)
External train	24.35	17.21
Internal train	5.35	11.07

SAGEM	KNN (k=1)	SVM (lin.)
External train	12.31	7.09
Internal train	3.40	5.70

Table 5.5: Percentage classification rate using the LBP feature set.

According to this protocol, D_2 is averagely two times bigger than $D_1^{(1)}$, thus the generic-user population is more dense than the user-specific one in the feature space. We repeated this procedure for ten runs and averaged results in Tables 5.5 - 5.7. Reported results are impressive, because they show at which extent the user-specific contribution can improve the performance of the system under the targeted user population.

Potentially, if well-exploited, the user-specific contribution may lead to a very high classification rate, with values which have not yet been attained by any FLD algorithm proposed

BIOMETRIKA	KNN (k=1)	SVM (lin.)
External train	30.43	18.56
Internal train	8.00	7.58

ITALDATA	KNN (k=1)	SVM (lin.)
External train	29.63	24.35
Internal train	10.37	6.79

DIGITAL	KNN (k=1)	SVM (lin.)
External train	22.40	14.10
Internal train	3.57	8.87

SAGEM	KNN (k=1)	SVM (lin.)
External train	15.85	10.34
Internal train	2.99	6.51

Table 5.6: Percentage classification rate using the LPQ feature set.

so far in these LivDet datasets [11, 32, 52]. In order to explain such astonishing results, at least on the observed data, we must consider that a live/fake sample from a certain user can be closer to live/fake samples of the same user but also closer to a fake/live of another. This impacts on the inter-intra class variability. Such effect can be measured according to what reported in Table 5.8. We reported the output of an extensive analysis of clients distributions along the LivDet 2011 data sets and is currently under extension on the others LivDet datasets.

Let's see the live class first. The "NN live" column reports the percentage of live sample whose nearest neighbour is another live sample. The "NN live from same user" column reports the fraction of above live samples which comes from the same user. It is worth noting that many samples, on average, are closest to samples of the same user. But this is not the rule: in a certain sense, this tells us that all features adopted are good liveness measurements and, according to the generic-user method, they are not too much distinctive on the specific users. Where the descriptor is more effective, namely, in the BSIF case, the "user-specific" component is more present.

Different observations can be done from the fake class. The "NN fake" column of Table 5.8 reports the percentage of fake samples whose nearest neigbour is another fake sample. The "NN fake from same user" is the fraction over the above samples, of images whose fake is closest to another one of the same user. Finally, the "NN fake of same user from same material" is the fraction of those samples nearest to a sample of the same material. It is evident that the dependency on the same user is very noticeable for fake samples. This means that having fake samples of the same targeted user appears more important than collecting his/her live samples. This may be due to the fact that a replica removes some textural information of the user's alive fingerprint, but it is also able to add novel, user-specific textural information, although "spoofed". We believe that this spoofed information can be referred to the artifacts peculiar of each material, on the basis of what Table 5.8 reports in the last

BIOMETRIKA	KNN (k=1)	SVM (lin.)
External train	28.10	18.13
Internal train	5.01	2.14

ITALDATA	KNN (k=1)	SVM (lin.)
External train	33.39	20.05
Internal train	7.29	2.18

DIGITAL	KNN (k=1)	SVM (lin.)
External train	21.97	6.73
Internal train	1.86	1.06

SAGEM	KNN (k=1)	SVM (lin.)
External train	11.07	6.05
Internal train	2.48	1.76

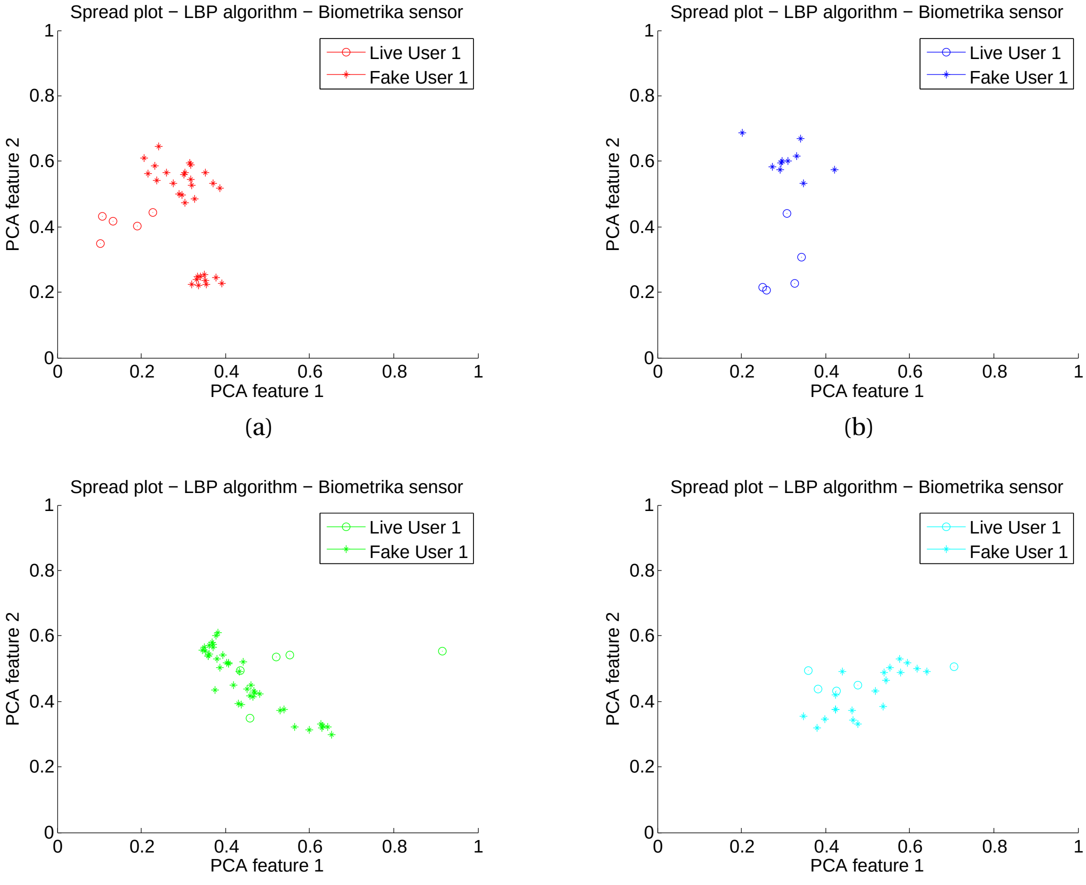
Table 5.7: Percentage classification rate using the BSIF (window 5x5, 12 bit) feature set.

column: the closeness to the same user’s fake fingerprint is coupled with the fact that that fingerprint is also close to one made up of the same material, but “far” from a corresponding live finger. Our hypothesis is that the “response” of the finger skin to a mold/cast combination is different from each person, thus leading to “user-specific” replica which is “more” than expected: it is not a simple replica, but a “novel” identity from the liveness perspective. This explains the performance difference when adopting the generic user approach, even if the materials used for the fake fingerprints are the same employed by the attacker.

5.4.2 Feature analysis

The experiments carried out so far have shown that some of the features used for liveness detection not only capture the intrinsic characteristics of “live” and “fake”, but also capture texture characteristics that distinguish a “client” from another. A reasonable hypothesis is to provide that the features are arranged in such a way as to characterize “live” and “fake”, but that their distribution is also related to the particular “client”. As a matter of fact, if you study and try to discern the distribution of live and fake, with no distinction between the client, you will find that these distributions are superimposed: both “live” and “fake” show a spreading throughout almost all of the allowable range. An analysis of these distributions, however, will highlight small groupings of features. These groupings determine overlaps between fake samples of some material other live samples. What, however, we still had not noticed is that some fake samples are superimposed, in the feature space, to live samples of different clients, whose fakes are much less overlapped.

In the following examples, we put some figures of four different clients. Figures 5.5 (a-d) show the spread plot of the first two PCA features of four client of the Biometrika LivDet 2011 dataset [59]. In the images we have differentiated the live (circles) from fake (asterisks). As you can see, fake and live of the same user are well separated for red and blue users but not for green and cyan users.



(a)

(b)

(c)

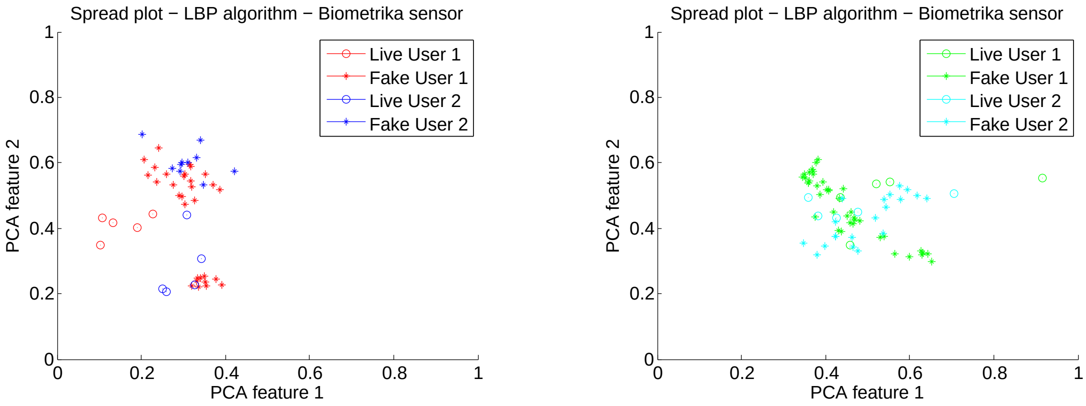
(d)

Figure 5.5: Separated spread plots of the fist two PCA features of four users of the LivDet 2011 dataset [59].

We can assume that in the first two cases 5.5 (a-b) we can easily train a classifier able to discriminate between live and fake, while in the other two we will get some error. If we put together the spread plot of the red and blue users, we can see that the live samples of the blue, one in particular, are very close to some fake samples of the red 5.6 (a). In the case of green and cyan the confusion is even higher and some live samples of both colors, which were originally isolated, are superimposed to fakes of the other color 5.6 (b).

But if we put together all these distributions, as is usually done, you will get a distribution much more difficult to characterize. Many live samples of one client can be very close to the fakes of many other clients (or some fake to other lives) 5.7.

This is because information "related" to each client quite often properly "isolate" fake and live, but the user-specific information become an obstacle to the correct classification



(a) (b)

Figure 5.6: Spread plots of the fist two PCA features of red and blue user (a) and green and cyan user (b) of the LivDet 2011 dataset [59].

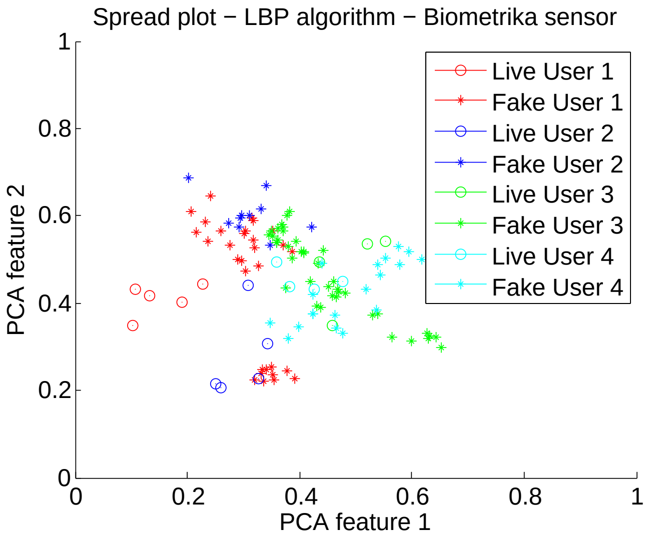
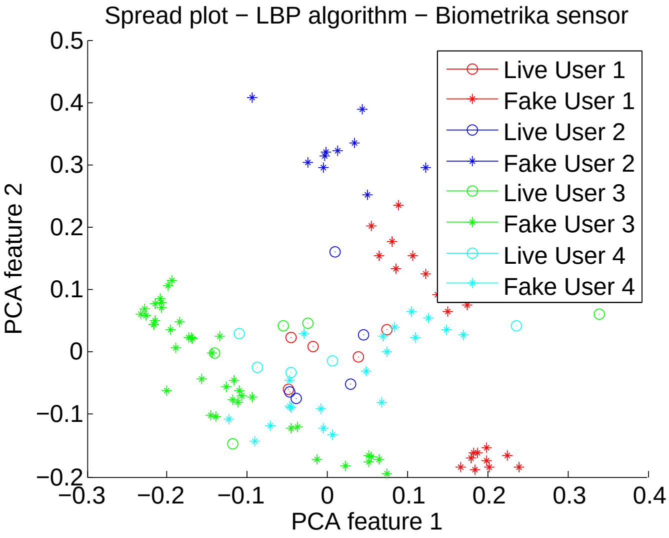


Figure 5.7: United spread plots of the fist two PCA features of four users of the LivDet 2011 dataset [59].

of the samples when they are treated as a single distribution. We consider this user-specific information as a bias that we would like to eliminate. The purpose of our work is thus to “de-personalize”, that is eliminate or reduce, the user-specific component that characterizes the chosen features. As previously noted, this characteristic invests all feature set considered so far, namely LBP, LPQ and BSIF, on which we will continue our analysis.

5.4.3 User shifting

The classification problem just analyzed it would simplify if we could turn it into a problem in which the live samples were “grouped” together, regardless of the client. Our hypothesis is the following: each single feature of the chosen feature set is characterized by a component inherently distinctive of the live or fake nature of the fingerprint to which is added a component characterizing the particular client. In order to remove or reduce this component, we propose to subtract from each feature (of live or fake samples) a value corresponding to the average of the corresponding values of that feature on all live samples. The effect of this shift will be that while the feature of the live class of a given client will now present a zero mean, the respective fake will be shifted by a constant amount. Considering the entire feature set, all the live samples will be centered around the origin, being all at zero mean, while fake samples will be moved on the margins of the various distributions. The result will be a general increase in the separability between live and fake samples, with a lower probability of overlapping between the fakes of a given client with the lives of another. In Figure5.8 it is evident that the live samples are centered around the origin while the fake are moved outside.



(a)

Figure 5.8: United spread plots of the fist two shifted PCA features of four users of the LivDet 2011 dataset [59].

5.4.4 Experimental results

We investigated the effectiveness of our approach on the Biometrika and Italdata data sets of LivDet 2011 [59]. Beside the analogue data sets of LivDet2013 [20], these are the only ones in which live samples of all the user population are always available. Each LivDet 2011 data set is made up of 4000 images per class (2000 live and 2000 fake), collected over 80 different people. For each finger there are five live and a varying number of fake acquisitions. We strictly followed the LivDet protocol which imposes using the first 2000 images as training set and the remainder as test set, where images in this set come from users not present in the training set. This protocol assures that "general-purpose" systems are trained and tested.

In order to analyze our proposal, we had to modify the LivDet protocol as follows. For each user in the training set, we computed by using the sample mean over all available live samples. Then, we removed such component from all live and fake fingerprint images of the related user. The obtained feature set, characterized now by a zero mean for the live samples of each user, is used for training a linear SVM classifier.

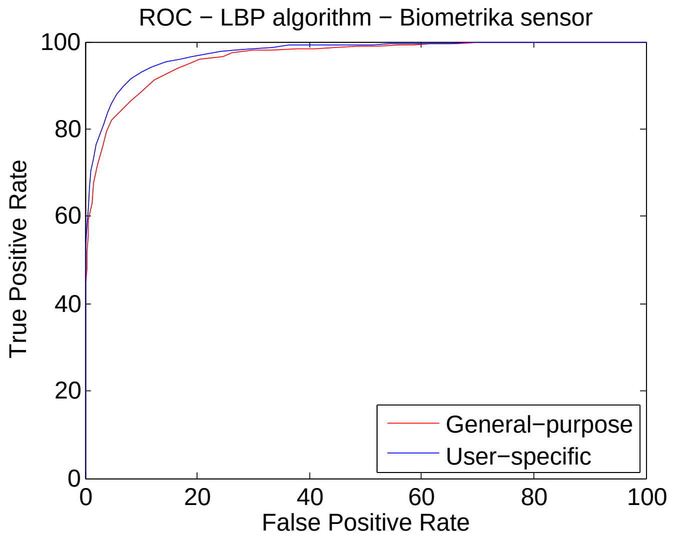
In the test phase, the mean for each user was computed using a similar procedure (live patterns of the user population are always available in real fingerprint verification systems, thus this procedure is completely fair) and subtracted to the related value of each feature vector. The obtained test set is used for evaluating the performance of the system. The only difference is that, since there are five live samples for each finger, in the test phase we randomly selected three feature vectors to compute the mean for each user and leaved the remaining two in the dataset. Basically, in a test dataset composed of 1000 lives and 1000 fakes, we use 600 lives to compute the mean vectors and we classify the remaining 1400 samples (400 lives and 1000 fakes). Results are compared with those obtained classifying the same 1400 samples without the mean subtraction, namely the original samples.

Performance is reported in terms of ROC curves, namely the plot of the True Positive Rate (percentage of live fingerprints correctly classified) and False Positive Rate (percentage of fake fingerprints wrongly classified). Results for both Biometrika and Italdata datasets are reported in Figures 5.9 (a,b) using the LBP, Figures 5.10 (a,b) using the LPQ and Figures 5.11 (a,b) using the BSIF with a 5x5 window size and 12 bits.

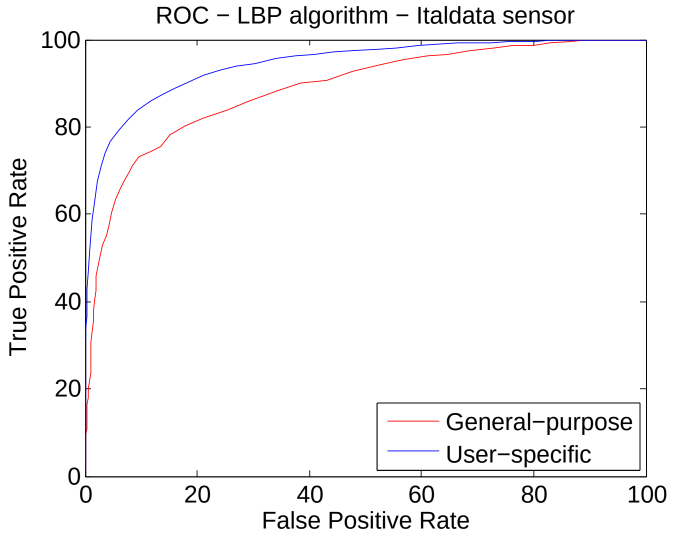
In Tables 5.9, 5.10 and 5.11 we also present the EER (Equal Error Rate) computed on the shifted and on the original feature vectors extracted with the LBP, LPQ and BSIF (with a 5x5 window size and 12 bits) algorithms.

In all cases, we can see a noticeable improvement of the performance with respect to the standard design protocol. This suggests that proposed method is effective in removing, partially at least, the user specific component from the related feature vector. It is worth noting that this method does not require collection of fake data from the user population taken into account, and training and design, once the user specific component is removed, can be done by following the usual procedure.

On the other hand, the same plots show that this improvement is not systematic, that is, it is not effective for all operational points. The reasons for this can be due to the strong hypothesis behind our model, in particular, the choice of the model of the user-specific component. If the basic assumption of our work is true, that is, a user-specific component exists into textural descriptors as LBP, LPQ and BSIF, modelling this relationship appropriately is the main issue.



(a)

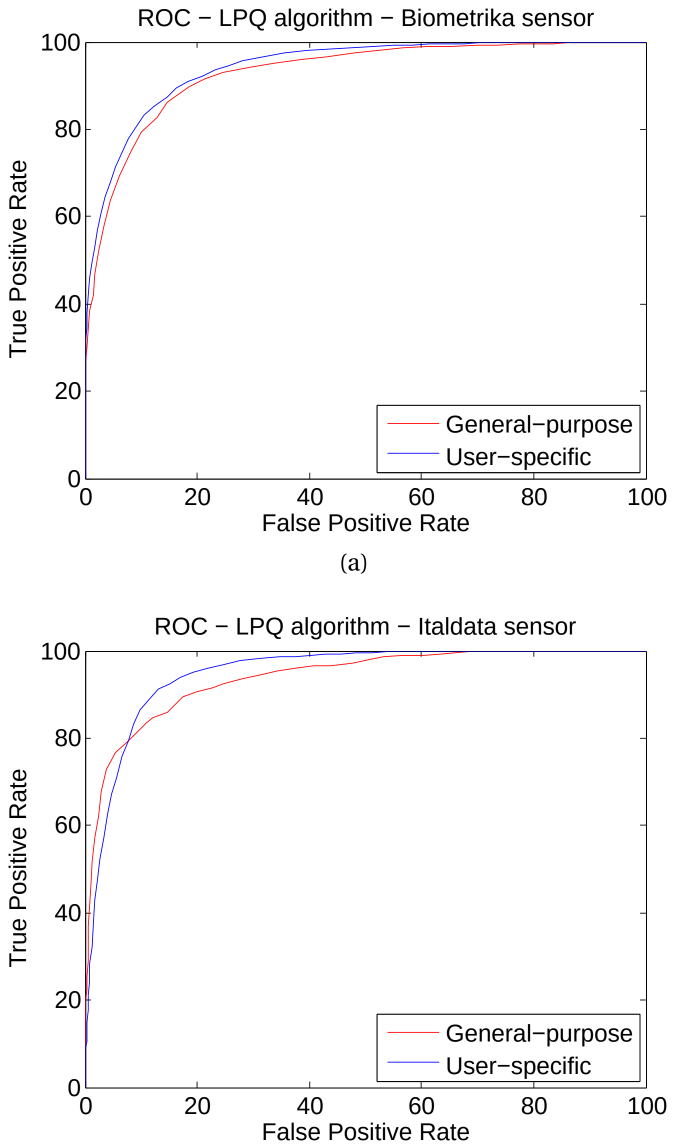


(b)

Figure 5.9: Roc curves computed for the general-purpose and the user-specific case using the LBP algorithm on the Biometrika and Italdata LivDet 2011 dataset [59].

		Live		Fake		
		NN live	NN live from same user	NN fake	NN fake from same user	NN fake of same user from same material
LBP	Biometrika	93,6%	18,5%	99,2%	96,4%	99,3%
	Italdata	88,6%	17,1%	99,0%	96,1%	98,9%
	Digital	96,7%	79,5%	98,5%	88,8%	100,0%
	Sagem	97,7%	59,6%	98,9%	89,6%	98,7%
LPQ	Biometrika	87,2%	21,1%	99,4%	93,6%	98,7%
	Italdata	87,0%	22,9%	98,9%	92,8%	98,0%
	Digital	96,3%	79,3%	98,5%	84,1%	99,9%
	Sagem	97,5%	70,5%	99,3%	91,7%	98,7%
BSIF	Biometrika	95,8%	46,5%	99,8%	99,6%	99,3%
	Italdata	91,2%	44,4%	99,5%	97,9%	98,9%
	Digital	98,9%	91,6%	99,9%	95,4%	99,9%
	Sagem	98,0%	82,0%	99,3%	93,9%	98,4%

Table 5.8: For each of three textural feature sets investigated, the "NN live" column reports the percentage of live sample whose nearest neighbour is another live sample; the "NN live from the same user" column reports the fraction of above live samples which comes from the same user; the "NN fake" column reports the percentage of fake samples whose nearest neighbour is another fake sample; the "NN fake from same user" is the fraction over the above samples, of images whose fake is closer to another one of the same user; the "NN fake of same user from same material" is the fraction of samples nearest to a sample of the same material.



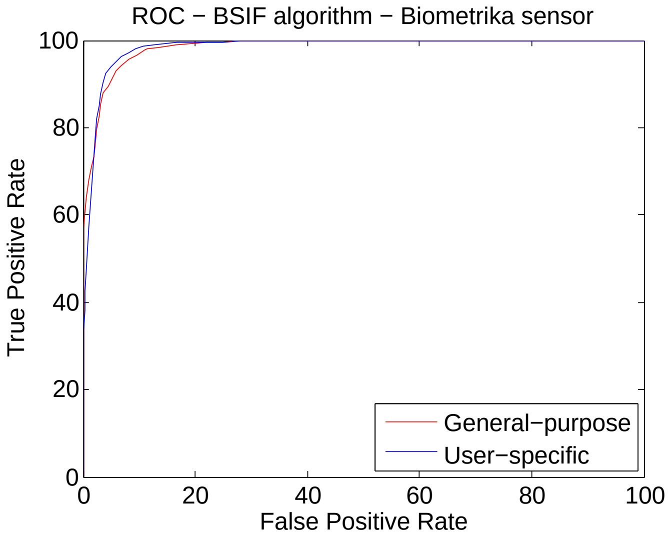
(a)

(b)

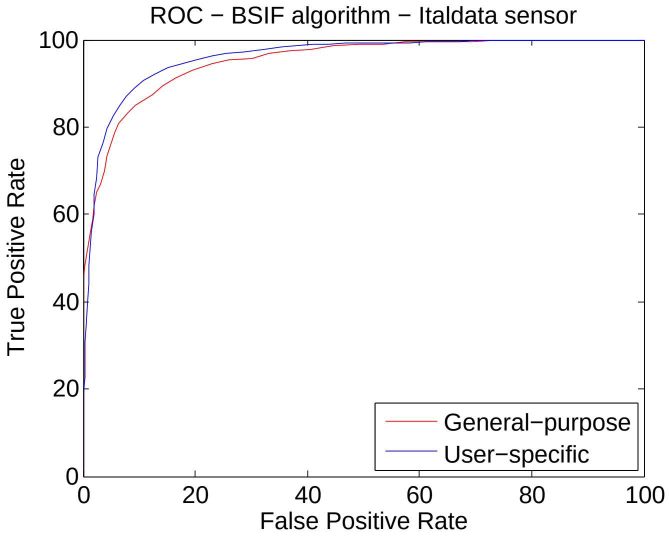
Figure 5.10: Roc curves computed for the general-purpose and the user-specific case using the LPQ algorithm on the Biometrika and Italdata LivDet 2011 dataset [59].

	Biometrika	Italdata
EER shifted feat.	8.36	13.07
EER original feat.	10.66	19.04

Table 5.9: EER computed on the shifted and on the original LBP feature vectors



(a)



(b)

Figure 5.11: Roc curves computed for the general-purpose and the user-specific case using the BSIF (with a 5x5 window size and 12 bits) algorithm on the Biometrika and Italdata LivDet 2011 dataset [59].

	Biometrika	Italdata
EER shifted feat.	13.45	11.11
EER original feat.	14.18	14.14

Table 5.10: EER computed on the shifted and on the original LPQ feature vectors

	Biometrika	Italdata
EER shifted feat.	5.17	10.05
EER original feat.	6.37	12.31

Table 5.11: EER computed on the shifted and on the original BSIF (with a $5x5$ window size and 12 bits) feature vectors

Chapter 6

Conclusions and Future Research

During the last three years I thoroughly explored the state-of-the-art of the FLD paying particular attention to the textural feature descriptors. I worked on the introduction of the LPQ and the BSIF in the FLD field carrying out a deep comparison with the other algorithms as well as an analysis of pros and cons depending on the algorithm's parameters. Results are very interesting because the proposed algorithms were comparable the state-of-the-art in terms of performance in the case of the LPQ, and outperformed the state-of-the-art in terms of performance and computational time in the case of the BSIF.

The main open discussion is related to set the optimal parameters for the BSIF extraction. Our analysis over ten data sets showed that the best values converge to be the same. In particular, performance does not vary appreciably over a certain window size and bits number, thus it could be stated that, for the FLD application, setting the windows size to $5x5$ pixels and the bits number to 11-12, leads to a more than reasonable trade-off in terms of computational time and performance. It should be noted that the characteristics of the considered fingerprints sensors were more or less the same for all data sets (around 500 dpi of resolution and gray-level images extracted), thus reported results are consistent. Nonetheless, we could expect that the parameters values could change when, for example, sensors significantly different in terms of dpi were available. It is true that, so far, the majority of fingerprint sensors on the market are similar to those investigated in this thesis, and the standard in this field is set to 500 dpi of resolution and 8-bit sized gray-level values. Anyhow, a specific investigation aimed to point out pros and cons of BSIF according to different kinds of

fingerprint sensors should be done in the future.

Another interesting focus of attention is on the filters derived from natural images. As it has been shown here by experiments it is true that they allow to compute a feature vector whose expressive power is more than that of, let's say, LBP. However, it might be possible that learning filters specialized on live and fake fingerprint images could lead to further, significant improvements. This is an on-going work that makes BSIF approach for FLD worth of further investigations. To sum up, thanks to its very good performance and computational time, the BSIF represents a significant step ahead to the real integration of a fingerprint liveness detector into a personal verification system.

An important step ahead in this direction was the introduction of the Multi-Scale BSIF. The first step to reduce the feature number with the PCA was necessary because of the large size of the feature vector and the subsequent feature fusion led, in almost all of the analyzed cases, to a significant accuracy improvement. The features reduction was performed using the 90%, 95% and 99% of the cumulative energy content and results were very similar but an deeper analysis of the feature selection criteria must be done carefully, evaluating the better size of each feature vector, the individual performance, and the real complementarity of the selected data.

I also investigated an interesting effect suggested by some empirical evidences which partially motivates the bad performance of current fingerprint liveness detectors using textural features: the presence of "user-specific" components. In particular, I analyzed LBP, LPQ and BSIF algorithms since they exhibited some of the best performance. We also proposed a model for trying to remove this component from the feature vectors. The proposed model is based on assumptions worthy to be analyzed in depth and this is what we are doing in these days. In particular, we are trying to obtain further evidences about the "user-specific" presence. Preliminary results seem to indicate that our approach is correct. With a simple translation of the feature based on the calculation of the average of the live feature vector of each user, it is possible to obtain an error reduction. If this would be confirmed, it implies that the design protocol, which we called "general-purpose" here, should be partially modified in order to "filter" the user-specific component from the feature vectors before the training phase. We believe that what we pointed out in this thesis marks the beginning of a new topic, which must be deeply investigated in future works for allowing an effective design of fingerprint liveness detector and leading to a performance acceptable for their integration in fingerprint verification systems.

Bibliography

[1] A. Abhyankar and S. Schuckers. A wavelet-based approach to detecting liveness in fingerprint scanners. In *Defense and Security*, pages 278–286. International Society for Optics and Photonics, 2004. [cited at p. 20]

[2] A. Abhyankar and S. Schuckers. Fingerprint liveness detection using local ridge frequencies and multiresolution texture analysis techniques. In *Image Processing, 2006 IEEE International Conference on*, pages 321–324. IEEE, 2006. [cited at p. 21]

[3] A. Antonelli, R. Cappelli, D. Maio, and D. Maltoni. Fake finger detection by skin distortion analysis. *Information Forensics and Security, IEEE Transactions on*, 1(3):360–373, Sept 2006. [cited at p. 20]

[4] S. Arivazhagan, L. Ganesan, and S. Kumar. Texture classification using curvelet statistical and co-occurrence features. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 2, pages 938–941. IEEE, 2006. [cited at p. 21]

[5] Yi Chen, AK Jain, and Sarat Dass. Fingerprint deformation for spoof detection. In *Biometric Symposium*, 2005. [cited at p. 21]

[6] I. Chingovska and A.R. dos Anjos. On the use of client identity information for face antispoofing. *Information Forensics and Security, IEEE Transactions on*, 10(4):787–796, April 2015. [cited at p. 37]

[7] H. Choi, R. Kang, K. Choi, and J. Kim. Aliveness detection of fingerprints using multiple static features. In *Proc. of World Academy of Science, Engineering and Technology*, volume 22, 2007. [cited at p. 21]

[8] R. Clarke. Human identification in information systems: Management challenges and public policy issues. *Information Technology & People*, 7(4):6–37, 1994. [cited at p. 6]

[9] P. Coli, G. L. Marcialis, and F. Roli. Analysis and selection of features for the fingerprint vitality detection. In *Structural, Syntactic, and Statistical Pattern Recognition*, pages 907–915. Springer, 2006. [cited at p. 19]

[10] P. Coli, G.L. Marcialis, and F. Roli. Power spectrum-based fingerprint vitality detection. In *Automatic Identification Advanced Technologies, 2007 IEEE Workshop on*, pages 169–173, June 2007. [cited at p. 21]

[11] P. Coli, G.L. Marcialis, and F. Roli. Vitality detection from fingerprint images: a critical survey. In *IEEE/IAPR 2nd International Conference on Biometrics (ICB 2007), Springer LNCS 4642, Seoul (Korea)*, pages 722–731, August 27-29, 2007. [cited at p. 18, 48]

[12] R. Derakhshani, S. Schuckers, L. Hornak, and L. O'Gorman. Determination of vitality from a non-invasive biomedical measurement for use in fingerprint scanners. *Pattern Recognition*, 36(2):383 – 396, 2003. Biometrics. [cited at p. 19]

[13] R. Derakhshani, S. Schuckers, L. Hornak, and L. O'Gorman. Determination of vitality from a non-invasive biomedical measurement for use in fingerprint scanners. *Pattern recognition*, 36(2):383–396, 2003. [cited at p. 21]

[14] R.O. Duda, P.E. Hart, and D.G. Stork. *Pattern classification*. John Wiley & Sons, 2012. [cited at p. 35, 38]

[15] L. Galbally, F. Alonso-Fernandez, J. Fierrez, and J. Ortega-Garcia. A high performance fingerprint liveness detection method based on quality related features. *Future Generation Computer Systems*, 28(1):311–321, 2012. [cited at p. vii, 22, 40, 45]

[16] L. Ghiani, P. Denti, and G.L. Marcialis. Experimental results on fingerprint liveness detection. In *AMDO'12 - 7th international conference on Articulated Motion and Deformable Objects, Mallorca, Spain*, pages 210–218, July 11-13, 2012. [cited at p. 40]

[17] L. Ghiani, A. Hadid, G.L. Marcialis, and F. Roli. Fingerprint liveness detection using binarized statistical image features. In *IEEE 6th International Conference on Biometrics: Theory, Applications, and Systems, (BTAS 2013), Washington DC (USA)*. [cited at p. 1, 22, 34, 43]

[18] L. Ghiani, G.L. Marcialis, and F. Rol. Experimental results on the feature-level fusion of multiple fingerprint liveness detection algorithms. In *Proceedings of the on Multimedia and Security*, MMSec 2012, pages 157–164, New York, NY, USA, 2012. ACM. [cited at p. 2, 40]

[19] L. Ghiani, G.L. Marcialis, and F. Roli. Fingerprint liveness detection by local phase quantization. In *21st International Conference on Pattern Recognition (ICPR 2012), Tsukuba, Japan*, pages 537–540, 11-15 Nov. 2012. [cited at p. 1, 22, 40]

[20] L. Ghiani, D. Yambay, V. Mura, S. Tocco, G.L. Marcialis, F. Roli, and S. Schuckers. Livdet 2013 fingerprint liveness detection competition 2013. In *Biometrics (ICB), 2013 International Conference on*, pages 1–6, 2013. [cited at p. 2, 3, 23, 39, 44, 52]

[21] C. Gottschlich, E. Marasco, A. Yang, and B. Cukic. Fingerprint liveness detection based on histograms of invariant gradients. In *Biometrics (IJCB), 2014 IEEE International Joint Conference on*, pages 1–7. IEEE, 2014. [cited at p. 22]

[22] D. Gragnaniello, G. Poggi, C. Sansone, and L. Verdoliva. Fingerprint liveness detection based on weber local image descriptor. In *Biometric Measurements and Systems for Security and Medical Applications (BIOMS), 2013 IEEE Workshop on*, pages 46–50. IEEE, 2013. [cited at p. 22, 40, 44, 45]

[23] R. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, (6):610–621, 1973. [cited at p. 21]

[24] J. Heikkilä and V. Ojansivu. Methods for local phase quantization in blur-insensitive image analysis. In *Local and Non-Local Approximation in Image Processing, 2009. LNLA 2009. International Workshop on*, pages 104–111. IEEE, 2009. [cited at p. 31, 32]

[25] A Hyvarinen and E Oja. Independent component analysis: algorithms and applications. *Neural Netw.*, 13(4-5):411–430, May 2000. [cited at p. 33]

[26] A. Jain, P. Flynn, and A. Ross. *Handbook of Biometrics*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007. [cited at p. 1, 5, 6, 8, 17]

[27] A. Jain. and D. Maltoni. *Handbook of Fingerprint Recognition.* Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2003. [cited at p. 1, 10, 15]

[28] J. Jia, L. Cai, K. Zhang, and D. Chen. A new approach to fake finger detection based on skin elasticity analysis. In Seong-Whan Lee and StanZ. Li, editors, *Advances in Biometrics*, volume 4642 of *Lecture Notes in Computer Science*, pages 309–318. Springer Berlin Heidelberg, 2007. [cited at p. 20]

[29] X. Jia, X. Yang, K. Cao, Y. Zang, N. Zhang, R. Dai, X. Zhu, and J. Tian. Multi-scale local binary pattern with filters for spoof fingerprint detection. *Information Sciences*, 268:91–102, 2014. [cited at p. 23, 40, 44]

[30] X. Jia, X. Yang, Y. Zang, N. Zhang, R. Dai, J. Tian, and J. Zhao. Multi-scale block local ternary patterns for fingerprints vitality detection. In *Biometrics (ICB), 2013 International Conference on*, pages 1–6, June 2013. [cited at p. 22]

[31] J. Kannala and E. Rahtu. Bsif: binarized statistical image features. In *Proc. 21st International Conference on Pattern Recognition (ICPR 2012), Tsukuba, Japan*, pages 1363–1366, 2012. [cited at p. 2, 33, 34, 38]

[32] E. Marasco and A. Ross. A survey on antispoofing schemes for fingerprint recognition systems. *ACM Comput. Surv.*, 47(2):28:1–28:36, November 2014. [cited at p. 2, 18, 48]

[33] E. Marasco and C. Sansone. An anti-spoofing technique using multiple textural features in fingerprint scanners. In *Biometric Measurements and Systems for Security and Medical Applications (BIOMS), 2010 IEEE Workshop on*, pages 8–14. IEEE, 2010. [cited at p. 22]

[34] E. Marasco and C. Sansone. Combining perspiration- and morphology-based static features for fingerprint liveness detection. *Pattern Recognition Letters*, 33:1148–1156, 2012. [cited at p. vii, 40, 45]

[35] G. L. Marcialis, A. Lewicke, B. Tan, P. Coli, D. Grimberg, A. Congiu, A. Tidu, F. Roli, and S. Schuckers. First international fingerprint liveness detection competition–livdet 2009. In *Proceedings of the 15th International Conference on Image Analysis and Processing*, ICIAP '09, pages 12–23, Berlin, Heidelberg, 2009. Springer-Verlag. [cited at p. 2, 3, 23, 39]

[36] G.L. Marcialis, F. Roli, and A. Tidu. Analysis of fingerprint pores for vitality detection. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 1289–1292. IEEE, 2010. [cited at p. 22]

[37] T. Matsumoto. Gummy and conductive silicone rubber fingers. In *Advances in Cryptology - ASIACRYPT 2002, 8th International Conference on the Theory and Application of Cryptology and Information Security, Queenstown, New Zealand, December 1-5, 2002, Proceedings*, pages 574–576, 2002. [cited at p. 1]

[38] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino. Impact of artificial gummy fingers on fingerprint systems. *Datenschutz und Datensicherheit*, 26(8), 2002. [cited at p. 1, 17]

[39] Yiu Sang Moon, JS Chen, KC Chan, K So, and KC Woo. Wavelet based fingerprint liveness detection. *Electronics Letters*, 41(20):1112–1113, 2005. [cited at p. 21]

[40] S. Nikam and S. Agarwa. Texture and wavelet-based spoof fingerprint detection for fingerprint biometric systems. In *Emerging Trends in Engineering and Technology, 2008. ICETET'08. First International Conference on*, pages 675–680. IEEE, 2008. [cited at p. 21, 22]

[41] S. Nikam and S. Agarwa. Wavelet energy signature and glcm features-based fingerprint anti-spoofing. In *Wavelet Analysis and Pattern Recognition, 2008. ICWAPR'08. International Conference on*, volume 2, pages 717–723. IEEE, 2008. [cited at p. 21]

[42] S. Nikam and S. Agarwa. Curvelet-based fingerprint anti-spoofing. *Signal, Image and Video Processing*, 4(1):75–87, 2010. [cited at p. 20]

[43] S. Nikam and S. Agarwal. Fingerprint anti-spoofing using ridgelet transform. In *Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on*, pages 1–6. IEEE, 2008. [cited at p. 20]

[44] S. Nikam and S. Agarwal. Fingerprint liveness detection using curvelet energy and co-occurrence signatures. In *Computer Graphics, Imaging and Visualisation, 2008. CGIV'08. Fifth International Conference on*, pages 217–222. IEEE, 2008. [cited at p. 20]

[45] T. Ojala, M. Pietikäinen, and T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(7):971–987, July 2002. [cited at p. v, 1, 29, 30, 32, 38, 43, 44, 45]

[46] S. Parthasaradhi, R. Derakhshani, L. Hornak, and S. Schuckers. Time-series detection of perspiration as a liveness test in fingerprint devices. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 35(3):335–343, 2005. [cited at p. 19]

[47] Matti Pietikäinen, Abdenour Hadid, Guoying Zhao, and Timo Ahonen. *Computer Vision Using Local Binary Patterns.* Springer, 2011. [cited at p. 32]

[48] N. K. Ratha, J. H. Connell, and R. M. Bolle. Enhancing security and privacy in biometrics-based authentication systems. *IBM Syst. J.*, 40(3):614–634, March 2001. [cited at p. 8]

[49] A.R. Roddy and J.D. Stosz. Fingerprint features-statistical analysis and system performance estimates. *Proceedings of the IEEE*, 85(9):1390–1421, Sep 1997. [cited at p. 13]

[50] S. Schuckers. Spoofing and anti-spoofing measures. *Inf. Sec. Techn. Report*, 7(4):56–62, 2002. [cited at p. 17]

[51] S. Schuckers and A. Abhyankar. Detecting liveness in fingerprint scanners using wavelets: Results of the test dataset. In *Biometric Authentication*, pages 100–110. Springer, 2004. [cited at p. 20]

[52] C. Sousedik and C. Busch. Presentation attack detection methods for fingerprint recognition systems: a survey. *IET Biometrics*, 3(4):219–233, 2014. [cited at p. 18, 48]

[53] B. Tan and S. Schuckers. Liveness detection for fingerprint scanners based on the statistics of wavelet signal processing. In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW'06. Conference on*, pages 26–26. IEEE, 2006. [cited at p. 20, 21]

[54] B. Tan and S. Schuckers. New approach for liveness detection in fingerprint scanners based on valley noise analysis. *Journal of Electronic Imaging*, 17(1):011009–011009, 2008. [cited at p. 20, 21]

[55] B. Tan and S. Schuckers. Spoofing protection for fingerprint scanner by fusing ridge signal and valley noise. *Pattern Recogn.*, 43(8):2845–2857, August 2010. [cited at p. 21]

[56] V. Ojansivu V and J. Heikkilä. Blur insensitive texture classification using local phase quantization. Proc. Image and Signal Processing (ICISP 2008), Cherbourg-Octeville, France, 5099:236-243, 2008. [cited at p. 2, 38, 43, 44, 45]

[57] T. van der Putte and J. Keuning. Biometrical fingerprint recognition: Don't get your fingers burned. In *Proceedings of the Fourth Working Conference on Smart Card Research and Advanced Applications on Smart Card Research and Advanced Applications*, pages 289–303, Norwell, MA, USA, 2001. Kluwer Academic Publishers. [cited at p. 17]

[58] D. Willis and M. Lee. Six biometric devices point the finger at security. *Netw. Comput.*, 9(10):84–96, June 1998. [cited at p. 17]

[59] D. Yambay, L. Ghiani, P. Denti, G.L. Marcialis, F. Roli, and S. Schuckers. Livdet 2011 - fingerprint liveness detection competition 2011. In *5th IAPR/IEEE Int. Conf. on Biometrics, New Delhi (India)*, pages 208–215, March, 29th, April, 1st, 2012. [cited at p. vi, 2, 3, 23, 39, 47, 49, 50, 51, 52, 54, 56, 57]

[60] J. Yang, Z. Lei, D. Yi, and S.Z. Li. Person-specific face antispoofing with subject domain adaptation. *Information Forensics and Security, IEEE Transactions on*, 10(4):797–809, April 2015. [cited at p. 37]

[61] Y. Zhang, J. Tian, X. Chen, X. Yang, and P. Shi. Fake finger detection based on thin-plate spline distortion model. In *Advances in Biometrics*, pages 742–749. Springer, 2007. [cited at p. 20]

List of Publications

Published papers related to the thesis

Conference papers

- D. Yambay, L. Ghiani, P. Denti, G.L. Marcialis, F. Roli, and S. Schuckers, LivDet 2011 - Fingerprint Liveness Detection Competition 2011, 5th IAPR/IEEE Int. Conf. on Biometrics, March, 29th, April, 1st, 2012, New Delhi (India), DOI: 10.1109/ICB.2012.6199810, pp. 208-215.

- L. Ghiani, G.L. Marcialis, and F. Roli, Experimental Results on the Feature-level Fusion of Multiple Fingerprint Liveness Detection Algorithms, 14th ACM Workshop on Multimedia and Security (MMSEC 2012), September, 6-7, 2012, Coventry (UK), ISBN: 978-1-4503-1417-6, DOI: 10.1145/2361407.2361434, pp. 157-164.

- L. Ghiani, P. Denti and G.L. Marcialis, Experimental Results on Fingerprint Liveness Detection, AMDO'12 - 7th international conference on Articulated Motion and Deformable Objects, July 11-13, 2012, Mallorca, Spain, ISBN: 978-3-642-31566-4, DOI: 10.1007 978-3-642-31567-1_ 21, pp. 210-218.

- L. Ghiani, G.L. Marcialis, and F. Roli, Fingerprint Liveness Detection by Local Phase Quantization, 21st International Conference on Pattern Recognition (ICPR 2012), 11-15 Nov. 2012, Tsukuba, Japan, ISBN: 978-1-4673-2216-4, pp. 537-540.

- G.L. Marcialis, L. Ghiani, K. Vetter, D. Morgeneier and F. Roli, Large Scale Experiments on Fingerprint Liveness Detection, Joint IAPR International Workshop, SSPR& SPR 2012, Hiroshima, Japan, November 7-9, 2012, ISBN: 978-3-642-34166-3, DOI: 10.1007/978-3-642-34166-3_ 55, pp. 501-509.

- L. Ghiani, D. Yambay, V. Mura, S. Tocco, G.L. Marcialis, F. Roli, S. Schuckers , LivDet 2013 - Fingerprint Liveness Detection Competition 2013, International Conference on Biometrics (ICB), pp.1,6, 4-7 June 2013 doi: 10.1109 ICB.2013.6613027.

- L.Ghiani, A.Hadid, G.L.Marcialis, and F. Roli. Fingerprint liveness detection using binarized statistical image features. In IEEE 6th International Conference on Biometrics: Theory, Applications, and Systems, (BTAS 2013),Washington DC (USA).

Published papers not related to the thesis

Conference papers

- G. Fadda, G. Marcialis, F. Roli, and L. Ghiani, Exploiting the golden ratio on human faces for head-pose estimation, In A. Petrosino, editor, Image Analysis and Processing ICIAP 2013, volume 8156 of Lecture Notes in Computer Science, pages 280-289, Springer Berlin Heidelberg, 2013.