UCL
Université catholique de Louvain

École polytechnique de Louvain (EPL)

Segmentation of fingerprint images on a glass plate

A multi-class fern approach

Dissertation presented by
Noé ANTOINE

for obtaining the Master's degree in
Mathematical Engineering

Supervisor(s)
Christophe DE VLEESCHOUWER

Reader(s)
John LEE, Laurent JACQUES

Academic year 2016-2017

Abstract

In the scope of a larger work studying the signature electrical signals of particular strains in the fingers, this work addresses the issue of segmenting the contact area between the finger tip and a glass plate, i.e. the fingerprint in the image. The automatized segmentation problem is defined as (Bazen and Gerez, 2001) deciding which parts of the image belong to the foreground, here the contact area, and which part to the background, the noisy area at the borders of the image that can contain various confusing elements for the algorithm like spurious fingerprints, artifacts or light spots. In addition to that, this works adds two more classes, one of each side of the delimitation line between the foreground and the background.

This report tries to solve the problem on the provided data set using random ferns, a tool that evaluates the probabilities of each pixel to belong to each one of the classes and allocates the latter to the one that is the most likely, based on a semi-naive Bayesian approach. Each fern consist of several binary tests in the neighbourhood of a pixel based on features of the fingerprint texture. Two of them yield good results: the coherence and the absolute deltas of pixels intensities.

After training the ferns probabilities on a small part of the manually segmented images, the results of segmentation are excellent for high-contrasted fingerprints, with more than 90% of pixels correctly classified, a percentage decreasing with the contrast level inside the fingerprint. Lastly, morphological operations are performed on the resulting region to obtain a compact segmented area and reduce the classification errors. Eventually, a few leads for improvement of the method and other potential useful features are discussed in the conclusion.

Contents

Chapter 1

Introduction

After an accident, some victims are unfortunately amputated of one of their hands which can be replaced with a prosthesis. However, this does not grant the victims as much autonomy as they had before since the latter acts like a numb limb. An important improvement would be to give the patient nervous information about the strain experienced by the prosthesis, which he could use to perform much precise tasks and gain self-government.[14] To this extent, some ongoing research is trying to relate human tactile afferent responses to specific skin strain patterns caused by fingerprint shearing.[1]

A crucial step to study the signature of specific strains is the data acquisition. In this case, it is achieved by filming the finger tip of different volunteers undergoing strains under controlled forces conditions. At the same time, a system of thin electrodes records the electrical impulses going through the median nerve of which the study is not addressed in this work.

This experience, described in [1], is at the heart of this thesis and supplies all the images to be worked with. It is briefly described and illustrated below. During the experience, a specially designed equipment firmly holds the finger (index or middle finger) still, the fingerprint facing a camera standing below as the patient is sat and his palm is facing downwards. Once the device is set, a transparent glass plate between the camera and the finger performs its routine in a few steps.

While the rest of the finger remains immobile all along, the glass plate goes to contact with the finger tip before loading a normal force on the latter that will act during the trial. Next, it slides horizontally $8mm$ one way then $12mm$ the other, before unloading, and finally going to cleaning. The duration of each of these steps is fixed, traveled distances are constant and the movement amplitude allows the finger tip to fully slide on the surface. Moreover, the loading step accounts for the inertia of the finger by allocating a full second for it to stabilize before beginning motion. Theses steps are illustrated in Figure 1.1.

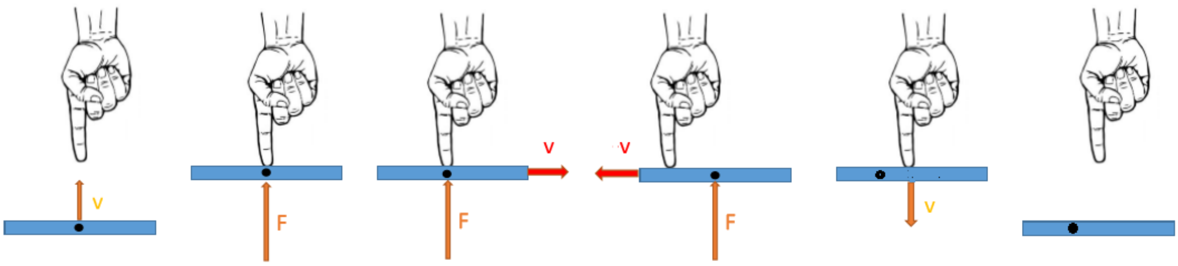


Figure 1.1: The different steps of the glass plate routine, ordered from left to right.

To measure the strains using a recently developed method (Delhaye et al. 2016) on the finger tip along the movement of the glass plate, the project team must know what part of the finger tip is experiencing the normal force, i.e. the contact area between the fingerprint and the surface. The main objective of this thesis is to develop a method which precisely determines the latter from a recorded fingerprint image, no matter how the subject's fingerprint is shaped and the state of the glass plate. Moreover, the solution must be automatized since there is a great number of images to segment.

To ease and normalize the work, the fingerprint segmentation will only be done on reference frames. The reference image of a recorded sequence is chosen as the one just before starting motion since the force is already loaded and the finger has had one second to stabilize. This implies a second part to the algorithm which performs tracking of the contact area along the frames of each sequence starting from the reference image, based among others on the difference of successive images.

This task is not as trivial as a rigid registration since the contact area changes in shape and must be regarded as a non rigid registration problem, where the area of interest can have variations in size and even disappear in the end of the sequence. This would need a complete dedicated study and is not addressed in this work.

To fulfill the segmentation, one can use different elements. First, the contact area, referred to as CA, has remarkable properties since it consists of a fingerprint, which is a very particular pattern. However, the quality of the images and the nature of the experience creates many challenges detailed later (e.g. artifacts). Second, the large number of images can be regarded as a considerable source of instances of fingerprints in experience conditions, differing from one trial to another. This provides an encompassing variety of situations including particular cases and should be taken advantage of.

The fingerprint segmentation problem is defined in this case as the task to decide which part of the image belongs to the contact area (i.e. the fingerprint) called the foreground and which pixels are members of the background, which is the potentially noisy area at the borders of the image. [3] A good algorithm would discard all the background without losing any pixels of the contact zone.

This issue is already addressed in the literature. Nevertheless, the images of this set have unique properties and require a customized solution to handle them since most of the papers tackle the issue on a specific data set which can arguably be quite different from the one considered here. Most of the methods divide the image in blocks of pixels (e.g. 16×16) and extract discriminating characteristics on the sub-blocks, training a linear classifier between foreground and background from images manually segmented.

Some authors try to perform first pre-processing like filtering to enhance the quality of the image: a local histogram equalization in [7] for instance. Features extracted on the blocks are, among others, the mean, the coherence of the gradients, the variance of pixel intensities [3], the spatial 2D Fourier frequency spectrum [5], the output of Gabor Filters [9] or even the histogram bin counts of some local feature in [15]. Finally, most papers use mathematical morphology as post-processing to smooth out the results and reduce the errors in the segmentation. [3] [5].

At the opposite, the approach is here pixel-wise. Instead of classifying sub-blocks of the reference image as a part of the foreground or of the background, the method classifies each pixel based on its neighbourhood. The random ferns provide a statistical method that allocates each pixel a class based on the distribution of some feature in a small area of the image around it. The idea is to use a set of binary tests on several pairs of pixels, comparing their values according to some useful feature with each other or according to some threshold. Afterwards, one decides from the obtained values and the trained distributions weather a pixel belongs to one class or another.

As it will be discussed when presenting the results, this method turn out to be able to perform a very accurate segmentation with a very small error margin at the border of the foreground. However, the outcome depend a lot on the design of the tests, the training set and mostly on the quality of the studied image: it happens to yield very bad segmentation on poorly contrasted images for instance, unfortunately. The objective is therefore redefined since one should also find features that efficiently differentiates the foreground from the background even in difficult cases.

This thesis is organized as follows. First, Section 2 defines the concept of random ferns. Then, Section 3 presents the methodology, starting from finding the useful features of the fingerprint texture to the design of the binary tests used in the random ferns. Finally, Section 4 gives a more practical insight of the data set before introducing the validation methodology and presenting some experimental results as well as some pre-processing and post-processing improving the classification.

Chapter 2

Random Ferns

The random ferns are the main tool used in this work. They allow to go from the ground truth, referred to as GT, to a quantification of the likelihood of each pixel to be a member of the different classes, using only its neighbourhood and a few binary tests based on useful features. This is done by assigning and computing the different class probabilities to each pixel.

In practice, one statistically train probability distributions of fern values for pixels that are part of different textures (e.g. the fingerprint or the background) as described below. This distribution is generated from the ground truth and evaluates the chances of belonging to the different classes using some simple properties of Probability. The concept presented as follows is mainly defined in [4] and [12] and is explained using similar terminology.

A fern is a group of N binary tests t_i, $i \in [1, N]$ in the neighbourhood of a pixel called a *patch* that is a square of size $(2R_{Patch} + 1)^2$ centered at this pixel. Each test usually compares a pair of pixels in the light of some metric like their intensities. It is chosen at random in the patch but can also be selected only in a more particular shape like a disk or a smaller part of these surrounding pixels. Once the N pairs of pixels are generated, they are fixed once and for all and are kept in memory as well as the corresponding tests.

A test t_i takes a value of 1 if succeeded and 0 otherwise. Therefore, there are 2^N possible realizations for a single fern. A simple way of enumerate these possibilities is to define a unique value F for each realization of a fern ranging in $[0, 2^N - 1]$ [4]:

$$F = \sum_{i=1}^{N} t_i 2^{i-1} \tag{2.1}$$

For instance, if $N = 4$ and tests take values $(t_4, t_3, t_2, t_1) = (1, 0, 1, 0)$, then $F = 1.8 + 0.4 + 1.2 + 0.1 = 10$. In practice, one will use a set of M ferns $\mathcal{F} = \{F_k = \{0, 1\}^N : 1 \leq k \leq M\}$. For a given class set $\mathcal{C} = \{c_i, 1 \leq i \leq H\}$, the ground truth allows to train them by evaluating them for pixels of which the class is known, and therefore computing the conditional probabilities $P(F_k|C = c_i)$ by accumulation of the training samples. The goal is to determine the posterior probabilities $P(C = c_i|F_1, F_2, \ldots, F_M)$. The Bayes formula writes:

$$P(C = c_i|F_1, F_2, \ldots, F_M) = \frac{P(F_1, F_2, \ldots, F_M|C = c_i)P(C = c_i)}{P(F_1, F_2, \ldots, F_M)} \tag{2.2}$$

If one admits the hypothesis that the prior is uniform, i.e. $P(C = c_i) = \frac{1}{H}$ then the probabilities are directly proportional since the denominator is independent from the class:

$$P(C = c_i|F_1, F_2, \ldots, F_M) \propto P(F_1, F_2, \ldots, F_M|C = c_i) \tag{2.3}$$

However, to capture a significant amount of information about the texture of the patch, one needs a great amount of tests $N \times M$, increasing with the size of the patch. As this would imply 2^{NM} different conditional probabilities to train and to handle per class per fern, which is not feasible as NM values is too large to keep in memory, one should use the solution presented in [12].

The semi-naive Bayesian compromise in [16] prevents suppressing the dependencies between binary tests within the same fern but considerably reduces the amount of conditional probabilities to compute by assuming that all ferns are independent, and therefore makes their number much less important, i.e. 2^N per class per fern, a number of values that can be easily stored in memory for a reasonable N value. This can be summered in the following approximation of the joint conditional probability:

$$P(F_1, F_2, \ldots, F_M | C = c_i) \approx \prod_{k=1}^{M} P(F_k | C = c_i) \tag{2.4}$$

To avoid a computational overflow that will quickly occur when M is large, one will finally use the monotone logarithm function and defines the score s_c of the class c_i for a pixel:

$$s_{c_i} = \log(\prod_{j=1}^{M} P(F_k | C = c_i)) = \sum_{k=1}^{M} \log(P(F_k | C = c_i)) \tag{2.5}$$

A natural classification will then be the MAP estimator $\hat{c}$ for each pixel:

$$\hat{c} = \underset{c_i \in \mathcal{C}}{\arg\max}\, P(C = c_i | F_1, F_2, \ldots, F_M) \approx \underset{c_i \in \mathcal{C}}{\arg\max}\, s_{c_i} \tag{2.6}$$

The next task is therefore to find useful features that efficiently capture the texture of the different classes and chiefly distinguish them, avoiding as much as possible to increase the computational cost at the same time, if possible. Therefore, the simplicity of the tests is something that must be looked for as well as the deep insight they should offer about the different classes. This is addressed in Section 3.

Finally, one should already notice that the method already has different parameters: the shape of the fern and the size of the patches, the number of tests N, the number of ferns M and the distribution of the tests within a single fern withdrawn from a library T of distinct binary tests.

The parameterization will be discussed both in Sections 3 and 4. Before that, let us introduce useful notations to be employed in the rest of this work, especially for the training of the ferns and the conditional class probabilities.

The training simply consists of computing the conditional class probabilities $P(F_k | C = c_i)$ for each fern F_k and class c_i by a counting procedure of the fern values for each class on the trained subset of the ground truth. Independently for each fern F_m, $m \in [1, M]$ one can trivially estimate the p_{k,c_i} parameters:

$$p_{k,c_i} = P(F_m = k | C = c_i) = \frac{N_{k,c_i}}{N_{c_i}} \tag{2.7}$$

Where N_{k,c_i} is the count of occurrences of $F_k = m$ of pixels belonging to c_i and N_{c_i} is the total number of pixels of class c_i that has been trained on. However, when performing this phase, one might not encounter all the 2^N possible fern realizations for a given fern F_j despite the fact that it could still occur when evaluating the ferns to another image. In that case, $P(F_{new} | C = c_i) = 0$ and it cancels out the product $\prod_{j=1}^{M} P(F_j | C = c_i)$ (and by the same times creates an error when computing log0), leading this scheme to yield poor results.[12]

To cope with this issue, one should instead use adjusted probability values as in [12] by using a non-zero minimal $P(F_j|C = c_i)$ value under the constraint that the sum of conditional probabilities must still be equal to one: $\sum_{k=0}^{2^N-1} p_{k,c_i} = 1$ for each fern F_m. One will take p_{k,c_i} to be:

$$p_{k,c_i} = \frac{N_{k,c_i} + N_r}{N_{c_i} + 2^N N_r} \tag{2.8}$$

where N_r is a regularization term, set to $N_r = 1$. The hypothesis is made that this value does offer proper classification results like in [12]. A comparative study of the results depending on this regularization value is left to an eventual further work.

Chapter 3

Methodology

3.1 Classes

This thesis tries to depict the forces and weaknesses of a multi-class fern approach to meet the objectives described before. The principle is to evaluate each pixel of an image based on its surrounding and allocating it to the most likely class depending on some tests capturing therefore the nature of the pixel. Let us define the set of classes $\mathcal{C} = \{c_i : 1 \leq i \leq H\}$ to be used in this work.

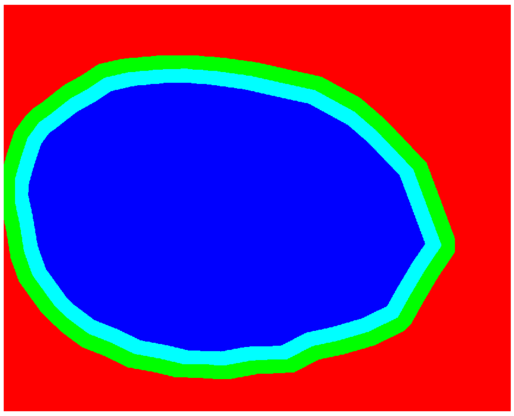
Here, $H = 4$ and the classes are defined as follows: the foreground which is the contact zone itself, the inner and outer borders which are a few pixels thick and spread on both side of the border of the contact area and finally the background which consists of all the remaining pixels. Both borders have a thickness of *rBorder* pixels, each spreading one side of the segmentation delineation. To ease the numerous calls to these classes, one can find in Table 3.1 the abbreviations of each of the four classes used later on and Figure 3.1 shows an example of the different classes for an image of the ground truth as well as the corresponding color code used from now on.

While the problem only consists of a two class problem, the decision to define two more entities resides in several points. First, since the ferns operate pixel-wise, a good solution is desired to be very accurate. In addition to that, the areas particularly sensitive to the separation stand a priori near the fingerprint borderline since the neighbourhood of the classified pixels can occur to be nearly of the same nature. Consequently, the method should separate BG and FG at first but also differentiate the four classes among themselves to eventually achieve a precise classification at the border. Second, this option allows a more complete study of the design of the random ferns and a deeper analysis of their experimental results when applied to this particular context.

Class	Abbreviation
Background	**BG**
Background border	**BBG**
Foreground border	**BFG**
Foreground	**FG**

Table 3.1

(a) Original Image



(b) Corresponding classes for $rBorder = 35$

Figure 3.1: Illustration of the four classes and color code definition. **BG** in red, **BBG** in green, **BFG** in cyan and **FG** in blue.

3.2 Image Sets

The camera records 8-bit gray-scale images at 50 frames per second for a total of 310 recorded frames per sequence. First, 25 frames correspond to the loading upon the finger, then during the next 50 the normal force is enabled to finally get to frame 76 which is the reference one of the sequence.

(a) Set 1

(b) Set 2

Figure 3.2: Illustration of the two images sets, displayed using the whole grey scale on the same image.

When displaying the reference frames, one can observe that the light intensity is not the same on the observation frame like in Figure 3.2a. Even though this could be probably used in practice in some cases, one should not consider this as a feature but rather as an imperfection. Moreover, the intensity disparity often considerably reduce the contrast of the fingerprint, especially when the contact area is small. To reduce the impact of this phenomenon, a second image set is introduced. Both sets are illustrated in Figure 3.2.

For each sequence, the corresponding image of Set 2 is the reference frame minus the first one of the sequence. As in the latter, the surface has not reached contact with the fingerprint yet, it does not contain relevant data yet. This process normalizes the contrast and removes unnecessary elements due to natural irregularities from the image. One can notice in Figure 3.2b that the fingerprint is well defined as far as the contrast goes. Also, on the borders of the finger there is a small lighter band which is due to the flatting of the finger undergoing the normal force. The ridges are lighter and the dark part is due to the shade created by the angle of the capture device.

3.3 Separation metric

Since the classification problem is now defined, the following essential step is to find a manner to quantify the quality of the separation between classes. To do so, one should determine the quantity representing them and after that choose a metric based on the latter that depicts the distances among the different classes. This section describes the classification measure to be used in this work.

For one fern, the scores of the different classes provides a useful tool that reflects the consequence of the test design on the classification. Indeed, they determine the most likely class of each pixel in Equation 2.6. The space between their distributions can be seen as an evaluation of the class disconnection. However, as the score is the logarithmic sum of conditional class probabilities of the fern realization value in Equation 2.5, the criterion will rather be the mean distance between the class distributions of values $P(F_k|C = c_i)$ on all the ferns, i.e. $1 \leq k \leq M$. This allows to do the measure independently from the uniform prior hypothesis that is not expected to be true. In practice though, the most important aspect is that a small score for a class correctly means that this class is the one of the considered pixel.

The metric used on two distributions p and q of respective random variables X and Y ranging over the same event set E is chosen as the Bhattacharyya distance $D_B(p, q)$. This distance is defined (in the case of continuous distributions) in [8] and writes:

$$D_B(p, q) = -log(BC(p, q)) \tag{3.1}$$

where $BC(p, q)$ is the Bhattacharyya coefficient defined as:

$$BC(p, q) = \sum_{e \in E} \sqrt{p(X = e)q(Y = e)} \tag{3.2}$$

Since p and q are probabilities, one can observe that $0 \leq BC(p, q) \leq 1$, a large value corresponding to similar distributions. Therefore, $0 < D_B(p, q) < +\infty$ and the larger the distance the best the separation between p and q. A priori, a good classification scheme will consequently yield a large value of this metric.

However, as discussed earlier, this only compare two classes and therefore different studies are of particular interest. First, the pair FG-BG must be given prime regard. Second, the pair BFG-BBG that discriminates the pixels on both sides of the fingerprint borderline.

3.4 Patch Rotation

In order to improve the results,a trick is to orientate the patch around the studied pixel according to its principal gradient before evaluating or training the ferns. Since each fern consists of tests of which the position remain fixed and that the patches may be highly oriented due to the fingerprint nature, this action allows the binary comparisons to capture more similar information and subsequently to be more meaningful. This is particularly true since the value of the side of the patch used in practice will be of the same order as the length of a ridge plus a valley. Therefore, for one idealized situation of a patch located near a ridge, the Figure 3.3 depicts the interest of turning the patches in the same principal direction : the test in blue will in all cases start in the lighter part and end up in the darker part. Of course, this is highly simplified but offers an insight of the potential influence of this step.

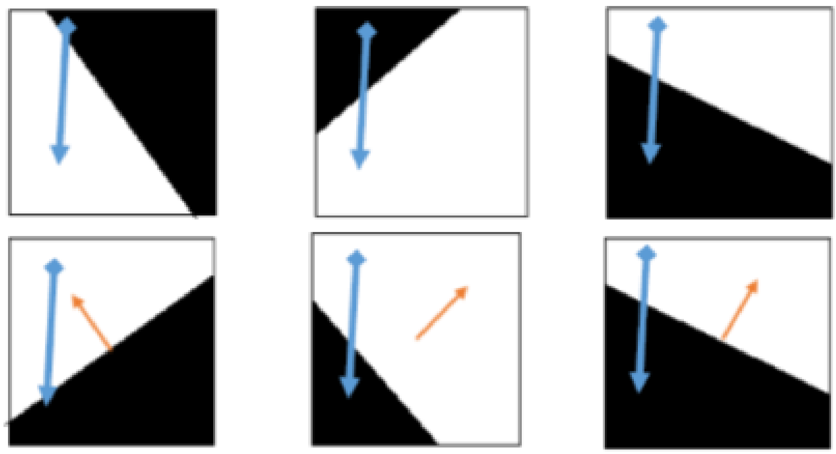
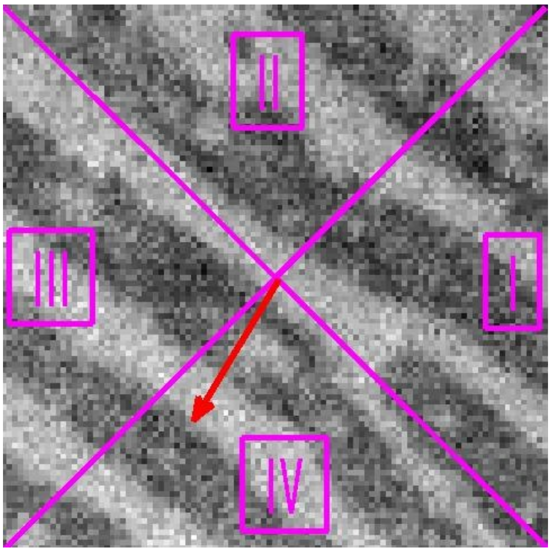


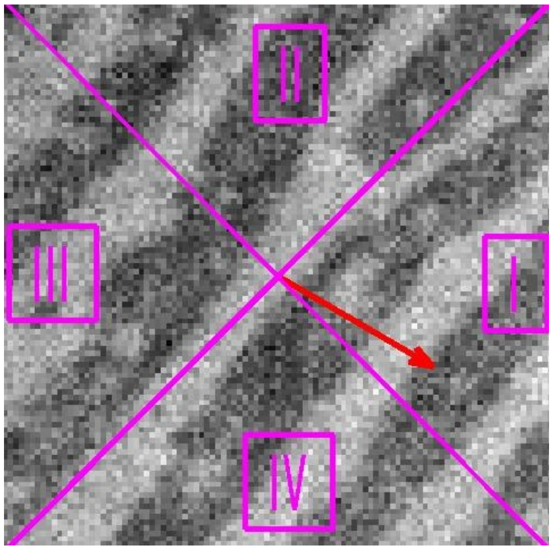
Figure 3.3: Idealized patch rotation. Binary test in blue. Principal gradient in orange. Upper row: initial patches. Lower row: patches after rotation with principal gradient in the upper quadrant.

In practice, the principal gradient is here determined by the following procedure. In a first time, the space is divided into four quadrants ninety degrees wide each: I, II, III and IV. Each of these quadrants points towards one of the main directions of the image, respectively $\{+y, -x, -y, +x\}$ (As implemented in the code: the x value is vertical, the y value is horizontal and the origin is top left of the image) as shown in Figure 3.4. Then, the gradient of intensities at each pixel of the patch is allocated to the quadrant it points to and its norm is added to a quadrant counter initialized to zero. Once all the gradients have been evaluated, the quadrant containing the principal gradient is chosen as the one with the greatest norm counter value.

Finally, the rotation is simply performed by turning the patch by 90 degrees the required number of times to move the principal gradient towards the quadrant I, as displayed in Figure 3.4. In practice though, the patch is not rotated since it requires a certain amount of memory and time. Instead, simple indices operations will suffice for this purpose.



(a) Patch before rotation



(b) Patch after rotation

Figure 3.4: Illustration of the rotation of a 100×100 patch and of the four quadrants in pink. In red is the principal gradient.

3.5 Features of the fingerprint and test design

This section describes the different tests that are studied in two points. On the one hand, the definition of the test as well as its potential interest and on the other hand a study of the actual fern value separation on small subsets of the ground truth. These subsets of image Set 2 consist of 4 easy images and of 4 difficult images. This allows at the same time to briefly evaluate the interest of the test with respect to the image quality.

The three image groups in this section are therefore referred to as EASY, HARD and ANY (defined as the set containing both EASY and HARD). An easy image is chosen such that the full contact area is very contrasted whereas an hard image is chosen to be contrasted but not everywhere: its segmentation is still quite clear for the human eye but there is a presence of some badly contrasted zones. These group definitions do not account for the quality of the BG and therefore both sets are chosen to contain diverse BG situations. This approach is only used as an helping tool to select the most promising features and to understand how and when to use them properly. The detailed results and potential adjustments to the full set are discussed in the next Chapter.

In practice, the ferns will be trained in this section with $R_{Patch} = 15$, since it make the patch side approximately corresponds to the width of a valley and a ridge on the images, $M = 200$, $N = 10$ and $nSamples = 15000$ which is defined as the number of pixels trained per class in the whole set and finally $rBorder = 20$.

3.5.1 Absolute Deltas of Intensities

The first inherent feature of the fingerprint is that it is highly contrasted, since it consists of ridges of high intensities and of valleys of lower intensities. It is the basis of the first binary test.

This test compares a pair of pixels of intensities p_1 and p_2 and is defined as follows:

$$T_{S_\Delta}(p_1, p_2) = \begin{cases} 1 & \text{if } |p_1 - p_2| \geq \Delta \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

where $\Delta \geq 0$ is uniformly drawn at random for each binary test from the small set of integer values S_Δ differentiating the classes, and is generated at the same time as the ferns, being also kept in memory. As the FG is more contrasted, the T_Δ success rate should be representative of the likelihood of a pixel to belong to the contact area.

To find discriminant values for S_Δ, one train the ferns on the ANY subset and search the histogram of $|p_1 - p_2|$ values per class for pivotal Δ_i, $i = 1, \ldots, |S_\Delta|$. When looking at Figure 3.5, one could be first tempted to only use only $\Delta_1 = 3$ since all classes are almost at the same level at that point only. However, to improve the robustness of the test to various contrast situations, one should instead use a few more values. Simultaneously, this allows to use pivots that separate different classes. Therefore, different pertinent sets such as $S_\Delta = 3, 6$ are tested, containing pivotal values where classes switch ordering in the histogram.

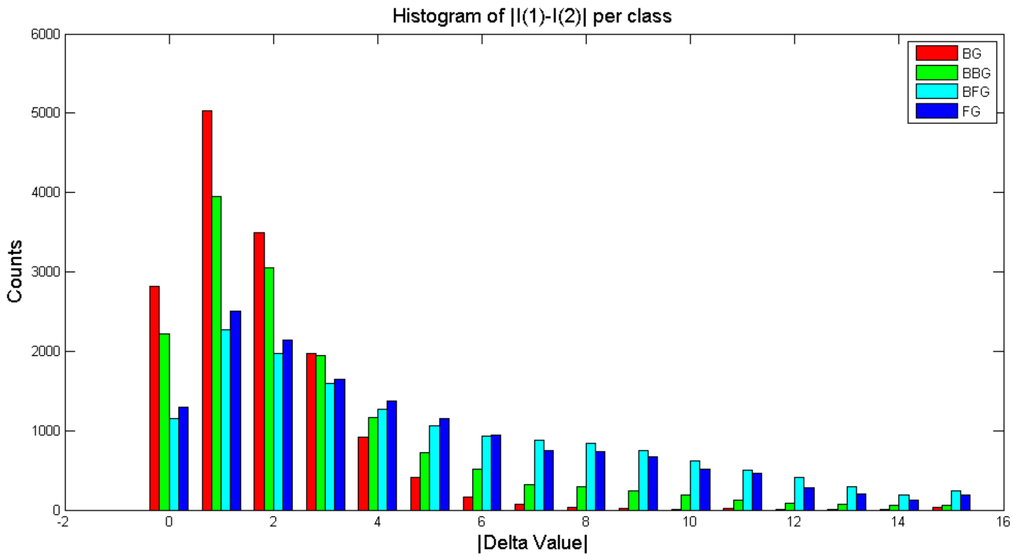


Figure 3.5: Histograms of $|I(1)$-$I(2)|$ values where $I(1)$ and $I(2)$ are intensities of a pair of pixel compared in a binary test within a fern.

The obtained Bhattacharyya distances are found in Table 3.2. They are obtained by considering only this type of test in T and computed for the training sets defined earlier. One can see that the tests offer a more important separation between FG and BG than BFG and BBG. Also, the HARD set is much less distinguished than the EASY one. Therefore, this test does not offer good results if the image is not well contrasted. Other features must be looked for to classify the HARD set better. The main advantage of this test is that it is requires a very short computation time.

Test	BC(FG,BG)	BC(BFG,BBG)
T_{S_Δ}, $S_\Delta = \{3\}$	0.48/0.20	0.18/0.07
T_{S_Δ}, $S_\Delta = \{2,4,6\}$	0.56/0.20	0.23/0.07
T_{S_Δ}, $S_\Delta = \{3,6,9\}$	0.55/0.18	0.19/0.06
T_{S_Δ}, $S_\Delta = \{3,6\}$	0.57/0.21	0.24/0.08
T_{S_Δ}, $S_\Delta = \{3,6\}$, Aligned pairs	0.67/0.25	0.20/0.06

Table 3.2: Mean over all ferns of the Bhattacharrya distances for the different tests on both image subsets. First entry of each cell is the distance for EASY and the second one the distance for HARD.

Finally, to improve the results, one does not take positions at random within the neighbourhood but instead considers pairs of pixels that are on the same horizontal line after the patch rotation since it is then aligned with the quadrant of the principal gradient. The first point is chosen at random within the patch and stands left of the second one, between 1 and R_{Patch} pixels away. This helps to capture contrast since the principal gradient is ideally normal to the crossed ridges. The results visible in Table 3.2 indicate that the distances are increased for both sets between FG and BG but not specially improved for the second distance.

The second test is directly inspired of the first one and uses similar pair of points. It is not strictly speaking one single trial on a couple of pixels but is instead based on a span of comparisons. Therefore, it is much slower than the previous one. The goal is to capture a broader information in a single test instead of considering a single pair of positions. Instead of considering one T_{S_Δ}, the test is this time based on $N_T = 10$ of them. The test is succeeded if the number of times that $T_{S_\Delta} = 1$ is at least equal to $N_T \times Ratio_{thresh}$ where $0 \leq Ratio_{thresh} \leq 1$.

Let us define T_{N_T,S_Δ} as follows:

$$T_{N_T,S_\Delta}(Ratio_{thresh}) = \begin{cases} 1 & \text{if } \sum_{i=1}^{N_T} T_{S_\Delta} \geq Ratio_{thresh} N_T \\ 0 & \text{otherwise} \end{cases} \tag{3.4}$$

where $Ratio_{tresh} \in S_{Ratio}$ with S_{Ratio} a set of a few values between 0 and 1. When creating the ferns, each test T_{N_T,S_Δ} is allocated with a $Ratio_{tresh}$ value taken uniformly at random in S_{Ratio} ans is also kept in memory.

Hopefully, this test is designed to increase the distance between BFG and BBG since the outside of the borderline can be described by the fact that when going from the fingerprint towards the BG, the frequency of T_{S_Δ} that are still equal to 1 decreases as the high-contrasted ridges quickly vanish of the considered neighbourhood. This time, one therefore considers only $S_\Delta = \{3\}$ but uses different values of $Ratio_{thresh}$ to account for this quick change at the border. To find an appropriate set S_{Ratio}, one look first at the histogram in Figure 3.6. It depicts the number of tests per class whose number of success among the $N_T = 10$ corresponding T_Δ is equal to the $N_{success}$ in the abscissa axis.

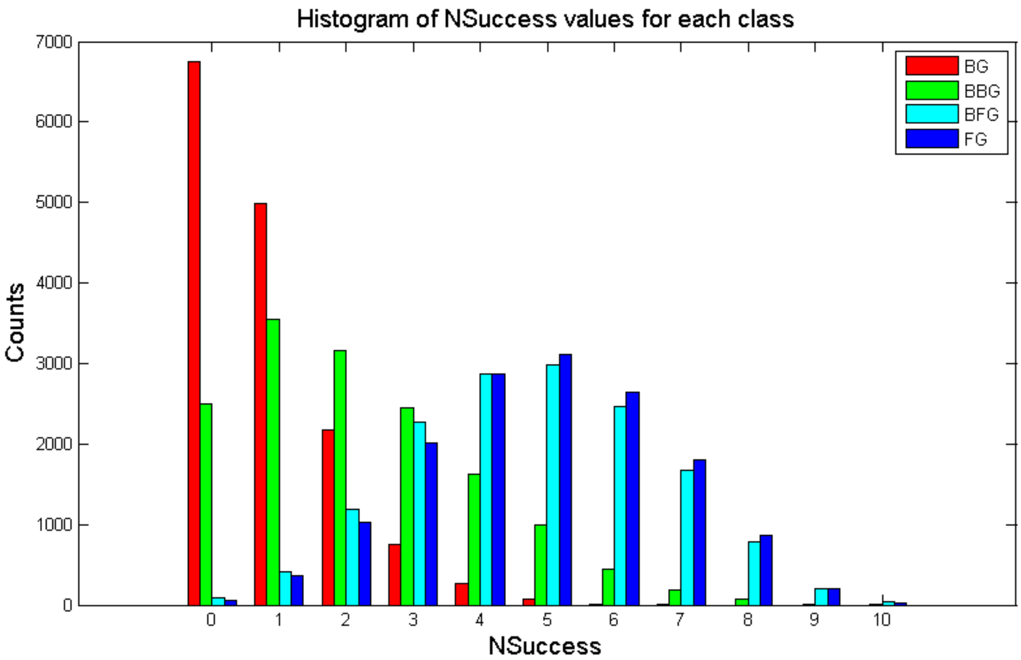


Figure 3.6: Histograms of $N_{success} = \sum_{i=1}^{N_T} T_{S_\Delta=\{3\}}$ for $T_{10,S_\Delta=\{3\}}$ for each class on the ANY set.

As one can see on the histogram, the classes distributions are very different, and as expected, FG and BFG take larger values of $N_{Success}$. For $0 \leq N_{Success} \leq 1$ and $4 \leq N_{Success} \leq 8$, one can clearly distinguish between the BFG and BBG counts, when the difference is large for any $N_{Success}$ between FG and BG except for the pivotal value $NSuccess = 3$. This feature seem to offer an appreciable separation between the four classes. Results for two S_{Ratio} sets are displayed in the Table 3.3. As hoped for, the distances are increased in all cases with respect to the previous test. The S_{Ratio} set is picked as the values where a class bin becomes greater than another one previously dominant. For instance, the pivotal value between BG and BBG is $N_{Success} = 2$, between FG and BG $N_{Success} = 3$, between FG and BFG is $N_{Success} = 4$, etc. One can notice that the distinction between FG and BFG is very small because of the similar nature of these pixels.

Test	BC(FG,BG)	BC(BFG,BBG)
$T_{10,S_\Delta=\{3\}}$, $S_{Ratio} = \{0.1, 0.3, 0.5\}$	1.07/0.47	0.28/0.11
$T_{10,S_\Delta=\{3\}}$, $S_{Ratio} = \{0.2, 0.3, 0.4, 0.5\}$	1.10/0.45	0.32/0.10

Table 3.3: Mean over all ferns of the Bhattacharrya distances for T_{N_T,S_Δ} tested for two S_{Ratio}. First entry of each cell is the distance for EASY and the second one the distance for HARD.

Even though it considerably improves the results, this test takes much more computational time and increases the number of binary tests by a factor N_T if the test set T does not contain any other test. Therefore, as the number of tests largely increase, it should naturally improve the results, especially in such a large patch. Nevertheless, the separation in the histogram of Figure 3.6 should and will be exploited. This feature capture information over the whole patch and even a little further, whereas the one introduced in the next section is a local feature : the $ULBP$.

3.5.2 Uniform Local Binary Pattern

A third test is defined in [15], for the fingerprint segmentation context. This time, the test only captures local information, e.g. only the eight pixels around the studied one. The feature is called *Uniform Local Binary Pattern* or ULBP and is explained in the following paragraphs.

To begin with, the local binary pattern (LBP) operator was introduced by Ojala [11] for the description of the image texture feature. It is notably invariant to monotonic gray-scale changes and has a low computational complexity. Despite this measure is local by construction, it can capture a texture in broader areas by using a different resolution scale. For each pixel the calculation of LBP is based on the intensities of its direct neighbours. Each of them is compared with the center pixel and the ones whose intensities equal or exceed the center pixel's are marked as 1, otherwise as 0. Just like the fern realization value, the $LBP_{P,R}$ number characterizing the image texture over a neighborhood is defined from a sequence of bits as follows:

$$LBP_{P,R} = \sum_{i=0}^{P-1} a_i 2^i \tag{3.5}$$

where R is the radius of the pixel disk, P is the number of neighbouring pixels (x_i, y_i) considered and $A(x,y) =< a_0, a_1, a_2, \ldots, a_{p-1} >$ denotes the ring feature for an image pixel (x, y), where:

$$a_i = \begin{cases} 1, & \text{if } Intensity(x_i, y_i) - Intensity(x, y) \geq 0 \\ 0, & \text{if } Intensity(x_i, y_i) - Intensity(x, y) < 0 \end{cases} \tag{3.6}$$

An example of $LBP_{8,1}$ from [15] is illustrated in Figure 3.7. The ULBP is used in particular for the fingerprint segmentation. A LBP descriptor is called uniform if and only if it contains at most two bitwise transitions between 0 and 1. The binary string is considered circular, just like the structure. For instance, $(1,1,1,0,0,1,1,1)$ or $(0,1,0,0,0,0,0,0)$ are uniforms when $(1,1,0,0,1,0,1,0)$ is not and contains 6 transitions.

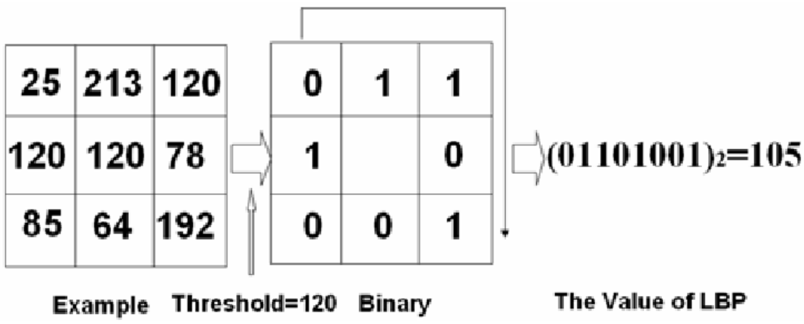


Figure 3.7: Illustration of $LBP_{8,1}$ calculation.

The ULBP allows to distinguish between pixels with the immediate proximity of a ridge or not since ideally that situation corresponds to one or two transition at most in the LBP sequence. A clean flat area will correspond to zero up to two transitions if there is a slight intensity gradation or more in the case of noise on the image. If there is more than two transitions, i.e. if a pattern is not uniform, it is allocated a specific value $P + 1$:

$$\text{ULBP} = \begin{cases} \sum_{i=0}^{P-1} a_i, & \text{if } Transitions(LBP) \leq 2 \\ P+1, & \text{otherwise} \end{cases} \tag{3.7}$$

,where:

$$Transitions(LBP) = |a_{P-1} - a_0| + \sum_{i=0}^{P-1} |a_i - a_{i+1}| \tag{3.8}$$

In [15], they divide the image in sub-blocks, compute the ULBP values and draw their histogram. According to certain thresholds on the number of ULBP values within certain ranges, they decide weather a block is segmented or not. For instance, if there is too many noise, the number of values equal to $P + 1$ will be large, therefore if there is too much of them, the block is not considered as a part of the FG. The pixels nearby the ridges should have ULBP values between approximately 3 and 5 so the number of points in that area indicate if the block is likely to belong to the fingerprint. To perform the same reasoning, the first step is to train the ferns and study the histograms of ULBP values for the different classes before finding a test distinguishing between them. The corresponding histogram of the ANY set is visible on Figure 3.8.

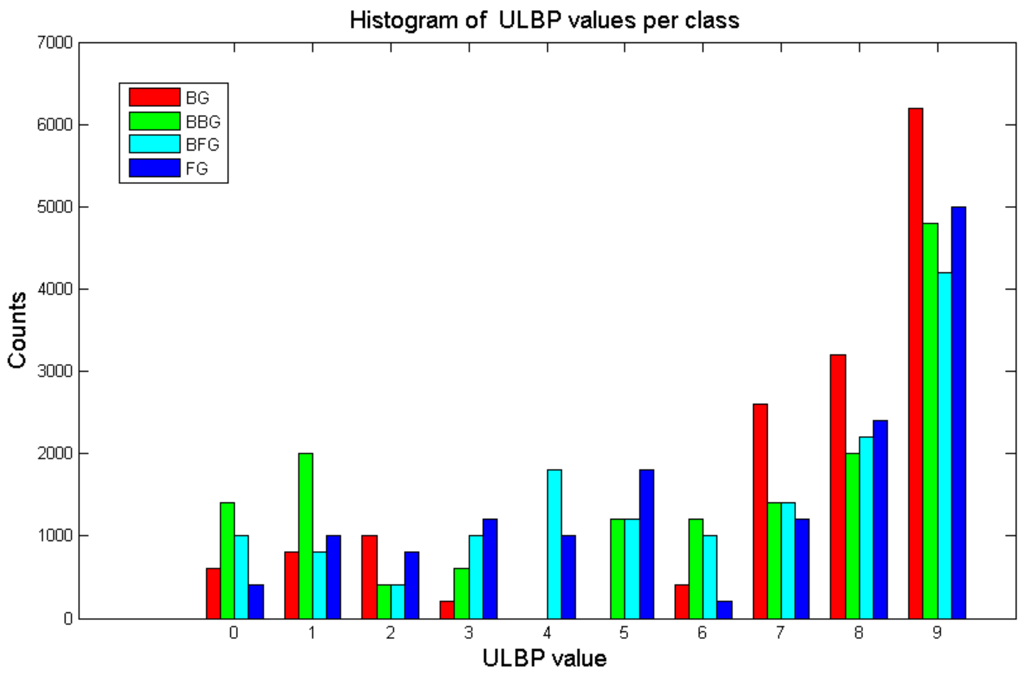


Figure 3.8: Histogram of ULBP value per class.

This histograms teaches two things. First, the background is indeed almost absent of the center values $[3, 5]$. Also, all the classes register a large number of ULBP values at the $P + 1 = 9$ bin, which means that the ANY set is quite noisy, even in the FG. To differentiate the classes, a possible binary test would subsequently be T_{ULBP} defined as:

$$T_{ULBP}(x, y) = \begin{cases} 1 & \text{if } 3 \leq ULBP(x, y) \leq 5 \\ 0 & \text{otherwise} \end{cases} \tag{3.9}$$

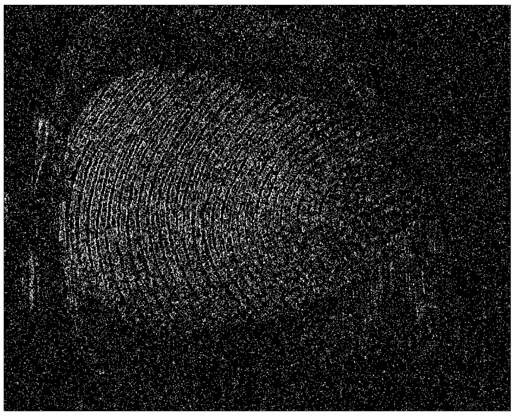
This test provides the distances in Table 3.4 when evaluated on the first pixel of each pair in the ferns. One can see that it does not yield larger distances unfortunately, probably because of the quality of the different images and their level of noise, especially for the HARD set where values are very low. If a second test $T_{ULBP=9}$ is added to the test set to account for the frequency of non uniform LBP, the distance is even less important since the images are all noisy and the bin $ULBP = 9$ does not offer a particularly large separation in Figure 3.8.

Test	BC(FG,BG)	BC(BFG,BBG)
T_{ULBP}	0.23/0.08	0.07/0.03
$T_{ULBP} + T_{ULBP=9}$:	0.15/0.06	0.05/0.03

Table 3.4: Mean over all ferns of Bhattacharrya distance for T_{ULBP} and and $T_{ULBP} + T_{ULBP=9}$ on both images subsets. First entry of each cell is the distance for EASY and the second one the distance for HARD.

Despite the short distances obtained, this feature is indeed of interest. One can observe a binarized illustration of all pixels such that $T_{ULBP} = 1$ in Figure 3.9. One can see that a low contrasted image appear as much noisy when binarized and this makes the test less appropriate. In particular, the test takes these values on the edges of the ridges, and therefore appear on the fingerprint, but with a very low density. That is why the ferns capture with difficulty this feature, especially at low-contrasted areas like in the right part of Figure 3.9b. It should be given an extra attention in case of a further work, especially in case of a multi-scale approach since it would possibly make the LBP radius match with the width of ridges and valleys, making therefore the test succeed in a larger part of the fingerprint and not only on the edges of ridges.

(a) Clear, high-contrasted frame



(b) Noisy, low-contrasted frame

Figure 3.9: Binarization of images for pixels such that $T_{ULBP} = 1$.

3.5.3 Coherence

Since the last test did not improve the separations of the distributions, the next feature will once more be defined over a larger neighbourhood, i.e. a square centered at the pixel whose the feature is evaluated. Instead of considering the intensities of the pixels directly, this feature is gradient-based and describes the alignment of squared gradients within some neighbourhood. It is called the *Coherence* and is first proposed in [10]. It is described and applied to the context of fingerprints in [2]: it measures how well the norm-squared gradients share the same orientation.

A fingerprint has ridges and valleys. The main gradient direction should be normal to the ridges, locally. However, when evaluating how aligned are the gradients, a simple mean would not be useful. Indeed, gradients on both sides of a ridge or a valley would be opposite and therefore a mean would cancel out and not be a good descriptor for the feature since they indicate the same orientation and their mean is equal to zero. This is due to the fact that locally the ridge-valley structures remain unchanged when rotated over 180 degrees [2] [13].

In consequence, an appropriate method is used to deal with this issue, i.e. the one described in [2]. Let us give each gradient its corresponding angle in the orientation span ranging from $-\pi/2$ to $\pi/2$. If one doubles these angles, the gradients in opposite direction will be identically oriented. Indeed, the distance between them will be doubled and go from 180 degrees to 360 degrees. At the opposite, perpendicular gradients will face opposite directions as the angle between them will be increased from 90 to 180 degree. In addition to that, the gradients will be squared in norm in order to give more weight to large changes in intensities. Moreover, this approach results in clean expressions. A brief description of how the coherence is derived follows, but more details can be found in [2].

The gradients vectors are first estimated in the Cartesian system as $[G_x, G_y]^T$. For the purposes described in the last paragraph, it is converted to the polar system to $[\rho, \phi]$ with $-\pi/2 \leq \phi \leq \pi/2$, as follows:

$$\begin{bmatrix} \rho \\ \phi \end{bmatrix} = \begin{bmatrix} \sqrt{G_x^2 + G_y^2} \\ tan^{-1} G_y / G_x \end{bmatrix} \tag{3.10}$$

And the reverse transformation is represented by:

$$\begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \rho cos\phi \\ \rho sin\phi \end{bmatrix} \tag{3.11}$$

The squared gradients vectors after doubling the angle are defined as $[G_{s,x} G_{s,y}]^T$ and can be expressed without any term in ϕ nor ρ using trigonometric identities and substitutions:

$$\begin{bmatrix} G_{s,x} \\ G_{s,y} \end{bmatrix} = \begin{bmatrix} \rho^2 cos2\phi \\ \rho^2 sin2\phi \end{bmatrix} = \begin{bmatrix} \rho^2 (cos^2\phi - sin^2\phi) \\ \rho^2 (2sin\phi cos\phi) \end{bmatrix} = \begin{bmatrix} \rho^2 [(\frac{G_x}{\rho})^2 - (\frac{G_y}{\rho})^2] \\ \rho^2 [2(\frac{G_x}{\rho})(\frac{G_y}{\rho})] \end{bmatrix} = \begin{bmatrix} G_x^2 - G_y^2 \\ 2G_x G_y \end{bmatrix} \tag{3.12}$$

Finally, the Coherence over a window W is given by [10]:

$$Coh = \frac{|\sum_W (G_{s,x}, G_{s,y})|}{\sum_W |(G_{s,x}, G_{s,y})|} \tag{3.13}$$

Therefore, $0 \leq Coh \leq 1$. If the squared gradients all point towards the same direction, their impact on the denominator and the numerator is the same and it results a coherence value of 1. At the opposite, it they are distributed in all directions, the resulting coherence value is equal to 0. Thus, the greater the coherence the stronger oriented is the window W. The fingerprint should then correspond to high coherence values.

The coherence equation can be re-expressed in terms of Cartesian coordinates G_x and G_y by substituting squared gradients using Equations 3.12 and defining $G_{xx} = \sum_W G_x^2$, $G_{yy} = \sum_W G_y^2$ and $G_{xy} = \sum_W G_x G_y$:

$$
\begin{aligned}
Coh &= \frac{|\sum_W (G_{s,x}, G_{s,y})|}{\sum_W |(G_{s,x}, G_{s,y})|} && (3.14) \\
&= \frac{|(\sum_W G_{s,x}, \sum_W G_{s,y})|}{\sum_W \sqrt{G_{s,x}^2 + G_{s,y}^2}} && (3.15) \\
&= \frac{|(\sum_W G_x^2 - G_y^2, \sum_W 2G_x G_y)|}{\sum_W \sqrt{(G_x^2 - G_y^2)^2 + 4G_x^2 G_y^2}} && (3.16) \\
&= \frac{|(G_{xx} - G_{yy}, 2G_{xy})|}{\sum_W \sqrt{(G_x^2 - G_y^2)^2 + 4G_x^2 G_y^2}} && (3.17) \\
&= \frac{\sqrt{(G_{xx} - G_{yy})^2 + 4G_{xy}^2}}{\sum_W \sqrt{G_x^4 + G_y^4 + 2G_x^2 G_y^2}} && (3.18) \\
&= \frac{\sqrt{(G_{xx} - G_{yy})^2 + 4G_{xy}^2}}{\sum_W \sqrt{(G_x^2 + G_y^2)^2}} && (3.19) \\
&= \frac{\sqrt{(G_{xx} - G_{yy})^2 + 4G_{xy}^2}}{G_{xx} + G_{yy}} && (3.20)
\end{aligned}
$$

The implementation requires to select a W window. First, the shape is selected as a square of $R_{Coh} \times R_{Coh}$ pixels centered in the evaluated pixel. The R_{Coh} is set to different values to evaluate the corresponding separations of the feature histograms. However, as the *coherence* defined previously is specially designed to account for gradients pointing towards opposite directions, one can take a large value for R_{Coh}. Indeed, the inside of the fingerprint is made of ridge-valleys structures and therefore it prevents the coherence to dramatically fall down when W is small and potentially entirely contained within a ridge or a valley. Instead, it ensures to give credit to the transitions ridge-valley and valley-ridge.

However, to avoid the loss of precision at the border and to avoid a large computational cost, one should not use very large values. A compromise would be to set its value to the precision scale of the method, i.e. $R_{Coh} \approx 2R_{patch}$ Histograms displayed in Figures 3.10a, 3.10b and 3.10c give the distribution of Coh values for $R_{Coh} = 7$, $R_{Coh} = 15$ and $R_{Coh} = 2R_{Patch} = 30$ respectively.

In these Figures, one can see that when increasing R_{Coh}, the different Coh class distributions grow narrower. At the same time, their maximal value is decreasing, since the ridges are not straight lines and a larger window induces more noise. Most importantly, the different classes are already distinguishable for a small value of R_{Coh} as seen in Figure 3.10a. However, BG and BFG are almost the same, which is due to the same nature of both classes at local scale. One can see there that the BG takes lower values than the other classes.

In particular, the BG and BBG have a pivotal value with FG and BFG at $Coh = 0.35$ in Figure3.10a. However, the BG and BBG distributions are much better separated for a larger value, i.e. $R_{Coh} = 30$ in Figure 3.10c, as the BG does almost not take values larger than $Coh = 0.15$, which is not the case for BBG. At the same time and less perceptibly, BFG is shifted left of the FG, which is expected since on a larger scale, the nature of BFG and FG behave differently.

In order to select binary tests that efficiently capture the differences between the classes, the first step is to define a test $T_{Coh,R_{Coh}}$ before selecting a few representative pivotal values as for the previous tests. It is defined as follows:

$$T_{S_{Coh},R_{Coh}}(x,y) = \begin{cases} 1 & \text{if } Coh_W(x,y) \geq Coh_{tresh} \\ 0 & \text{otherwise} \end{cases} \tag{3.21}$$

where S_{Coh} is a set of a few values in $[0,1]$, Coh_{tresh} is uniformly drawn at random in the latter, W is a square of $R_{Coh} \times R_{Coh}$ pixels centered in (x,y).

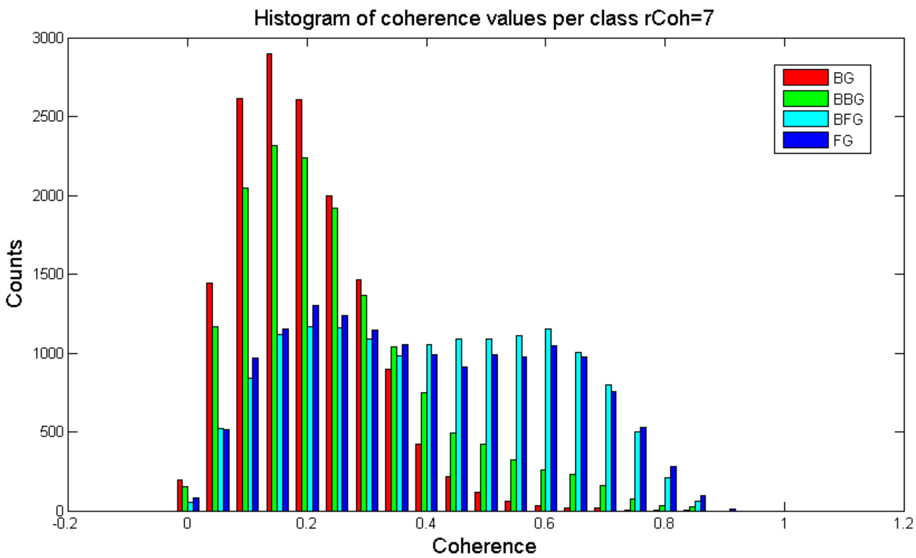
A comparison between the images of Coh values with R_{Coh} values of respectively 10 and 30 is visible in Figure 3.11 on an arbitrary reference frame in EASY. Despite a very clear capture of the ridges in the first image, the other parts of the picture take coherence values that are closer to the ones inside the fingerprint, which is not wished for. To avoid a noisy feature distribution such as the one for T_{ULBP} displayed in Figure 3.9b (because the test captures information around a too small area) and in sake of a better separation, one will accordingly choose a large value of $R_{Coh} = 30$. Therefore, in accordance with the values of the ANY set depicted in Figure 3.10c, two sets will be tested: $S_{Coh} = \{0.15, 0.25, 0.45\}$ and $S_{Coh} = \{0.1, 0.2, 0.45\}$. A third situation is added to compare with the case of a lower R_{Coh}.

Distances in Table 3.5 are obtained by evaluating $T_{S_{Coh},R_{Coh}}$ on the first pixel of each pair in the ferns. Moreover, the test set does not contain any other test, as for the previous results. One can observe the the separation is the best for BC(FG,BG) with $R_{Coh} = 30$ and $S_{Coh} = \{0.1, 0.2, 0.45\}$ but is not very sensitive to the parameters. However, $S_{Coh} = \{0.25\}$ provides better results for BC(BFG,BBG) because that value is pivotal for these two classes. Therefore, adding other values to the set deteriorates the separation between them but increases the one between FG and BG. The distances are of the same order as the ones obtained for T_{N_T,S_Δ} in table 3.3.

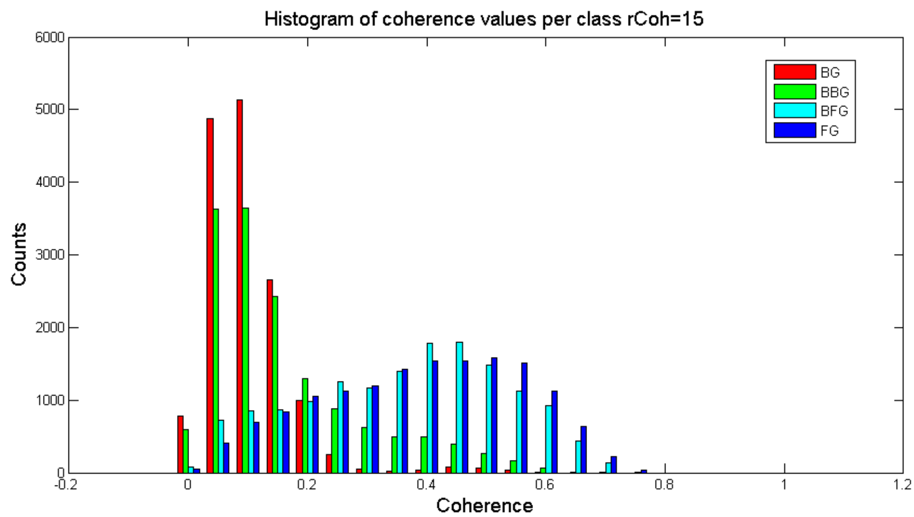
Test	BC(FG,BG)	BC(BFG,BBG)
$R_{Coh} = 10$, $S_{Coh} = \{0.25\}$:	0.93/0.40	0.33/0.08
$R_{Coh} = 30$, $S_{Coh} = \{0.15, 0.25, 0.45\}$	1.03/0.38	0.25/0.05
$R_{Coh} = 30$, $S_{Coh} = \{0.1, 0.2, 0.45\}$	1.03/0.042	0.25/0.06

Table 3.5: Mean over all ferns of Bhattacharrya distance for $T_{S_{Coh},R_{Coh}}$ on both images subsets. First entry of each cell is the distance for EASY and the second one the distance for HARD.

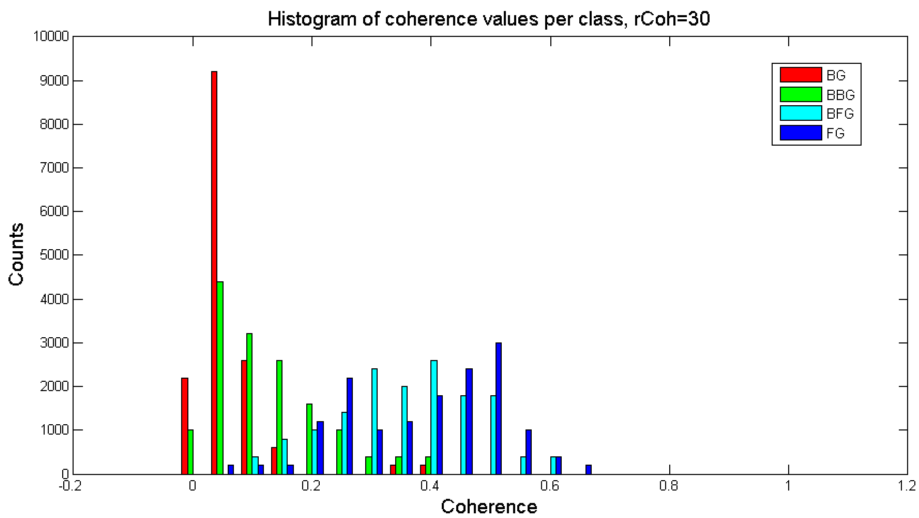
To conclude with, the tests on features on a window of about the size of the patch provides better results in term of Bhattacharyya distances than the ones on local features shown in the Tables 3.2 and 3.4. The set of tests will therefore contain two types of sets: $T_{S_{Coh},R_{Coh}}$ and T_{N_T,S_Δ}. The validation of this method and the results on the data set follow in the next chapter.



(a) $R_{Coh} = 7$

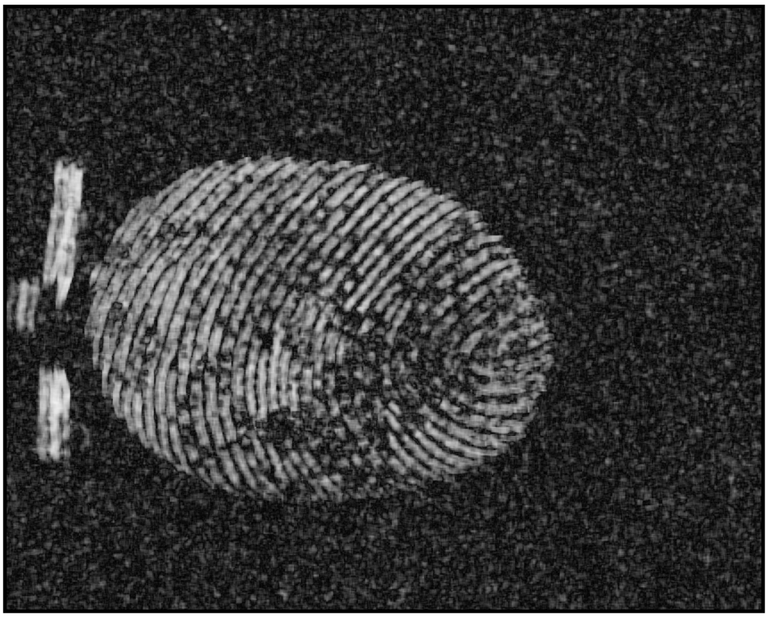


(b) $R_{Coh} = 15$

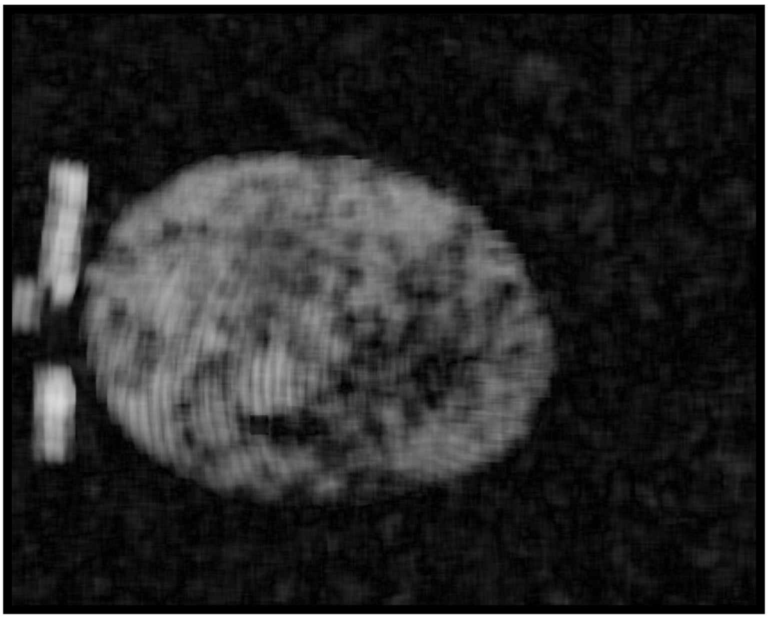


(c) $R_{Coh} = 30$

Figure 3.10: Histograms of Coherence values per class for different R_{Coh} values on ANY.



(a) $R_{Coh} = 10$



(b) $R_{Coh} = 30$

Figure 3.11: Coherence values for a frame in EASY with different R_{Coh} values displayed on the whole gray scale range.

Chapter 4

Results and Validation

4.1 Dataset

The first step of the segmentation task has been to select representative images and manually draw the contact areas on those, the latter being visible for the human eye, by selecting a number of points on the borderline. Successive points in the segmentation line are connected with linear interpolation. This is referred to as the *Ground Truth* or GT and will be essential for the validation of the method. The data set consists of 80 reference frames from distinct sequences of 1024×1280 pixels each. The resolution of the camera is 1200 dpi. [1]

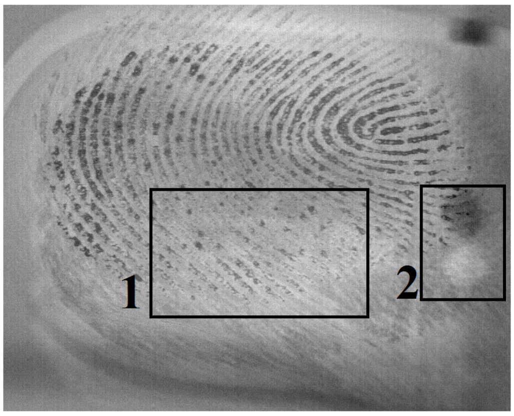
4.1.1 Challenges

Despite the exploitable statistical properties of the fingerprints, the segmentation task is quite challenging. The fact that the finger slides on a glass plate creates artifacts, spurious fingerprints from previous positions of the finger misunderstood as part of the contact area, light reflects or a sometimes visible device. In addition to that, the fingerprints have sometimes been altered with a marker pen on research purposes. Most of the challenges come from the fact that an altered contrast brings the fingerprint away from its idealized nature from which tests have been designed or, at the opposite, that a good contrast in the BG is mistaken as a part of the FG. The Figure 4.1 shows some of these challenges.

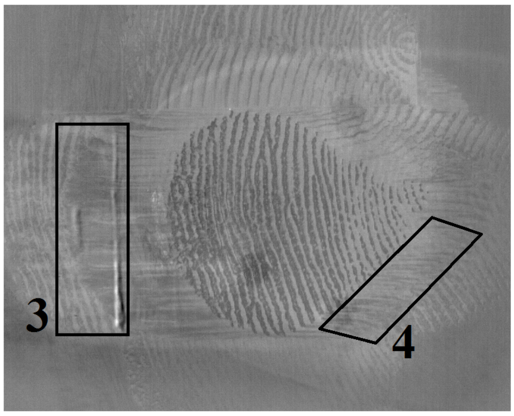
Also, the shape of the contact area is most of the time convex but can also be heavily non convex and this may be problematic as well since the convex hull of the remaining area after morphological operations does not correspond to a permissive approximation of the contact zone. Sometimes the images have a terribly low contrast on a large part of the fingerprint and this creates dramatic issue to fulfill the segmentation as they are supposed to have clear ridge-valley structures, at least in a large part of it.

4.1.2 Image Categories

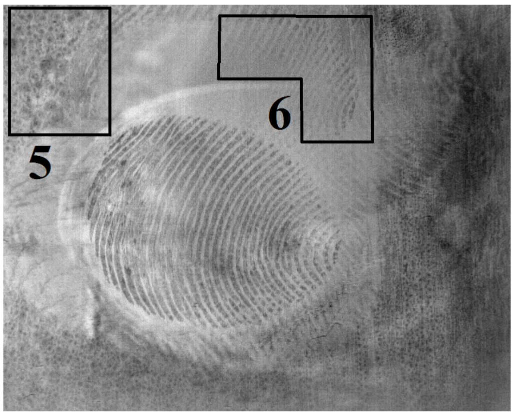
Inspired by [6], the GT is divided into several categories. These are not strictly objective but can help giving an insight of the impact of the contrast in the fingerprint on the results: according to the kind of contrast in the contact area in Set 2, the images are catalogued into four distinct sets. These sets all contain different challenges cited in the previous section and are not defined in accordance to them. These sets are defined in Table 4.1, alongside with the number of images of GT for each set. Examples of each sets are shown in Annex A.2.



(a) 1: Discontinuities in the contrast in FG. 2: Marking points for research purpose.



(b) 3: Light of the device. 4: Lubrification trail.



(c) 5: Artifacts due to lubricant. 6: Spurious Fingerprint.

Figure 4.1: Illustration of the challenges.

Category	Definition	Hits
A	Excellent contrast on the full fingerprint	32
B	Good contrast on the fingerprint except for small parts	38
C	The CA is distinguishable but many parts have bad contrast	6
D	The CA is guessed, not visible everywhere and have a terrible contrast	4

Table 4.1: Definition of the different image categories and their populations.

4.1.3 Limitations

Even though a fingerprint has a very recognizable structure, each fingerprint is unique and the images offer a certain span of various situations as far as both the background and the foreground are concerned. According to that, 80 images is not an extremely large amount of data. However, this algorithm is not supposed to be able to adapt to any situation but rather to perform segmentation on a set of frames of the same nature obtained in a similar context and on a limited set of fingerprints. Nevertheless, when considering for instance the very low number of images in category D, one can understand that these are not numerous enough in the GT to ensure an adequate training of the ferns adapted to that situation. The results will therefore not be very accurate for the latter category.

An important compromise is to use the Set 2. The images are indeed more contrasted but this can induce large areas of shade or light due to the differences of positions between the first image of the sequence and the reference frame as shown in Figure 4.2. This is a problem since there is a large contrast between that band and the neighbouring pixels: it will be mistaken as FG. As a rule of thumb though, the border of the finger gets a brighter band due to the flattening of the finger. As the gradients are aligned in that band, one should care for it when studying the results.

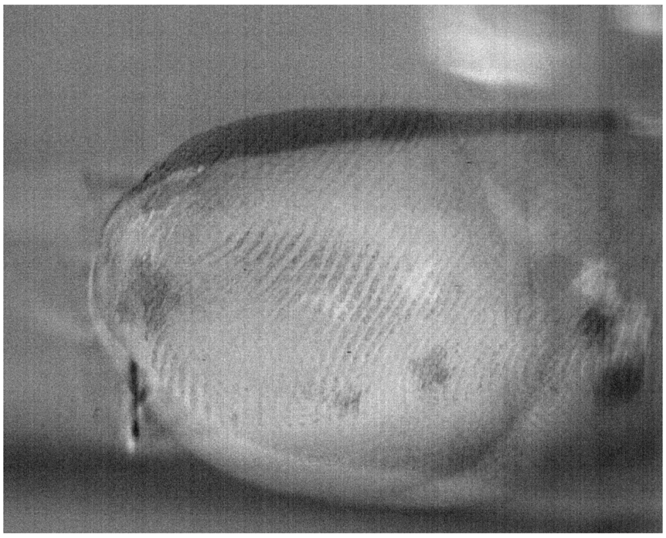


Figure 4.2: Large shade band in Set 2.

As described in the previous paragraph, the nature of the images is challenging. In particular, the set contains a large range of sizes for the FG : if r_{FG} is the ratio of pixels belonging to the FG, then it has a mean value of 0.23 and a standard deviation of 0.072 on the FG. This means that r_{FG} can be rather small, and this large gap between the size of the FG and of the BG can lead to errors when performing mathematical morphology to improve the results. Indeed, a step will be to save the largest connected component, which in the case of a difficult image like in Figure 4.2 can lead to dramatic errors.

Unfortunately, this set contains images with very bad quality in category D. The method introduced here is not able to guess the FG. This would require an additional work that considers the image in a larger scale and the fingerprint as a single object instead of this pixel-wise method.

4.2 Validation Methodology

According to the uniform prior hypothesis, the number of pixels trained is the same for all the classes. The maximal number of pixels that can be trained for a single class in a single image is equal to the minimal number of pixels belonging to BFG in the FG and is about 28000 pixels. In practice, the number of pixels trained per class will be of 6000 for each image of the training set. Of course, this means that the density of pixels trained is much higher for BFG and BBG than for BG and FG. In the sections of results, this hypothesis will be corrected by adjusting the prior class probability by multiplying the conditional probabilities by a certain factor which translates in shifting the scores of the ferns when using the logarithm in Equation 2.5.

The training set will not use any pixels that are classified as a part of light of the device, using the appropriate pre-processing method detailed in Section 4.2.1. In the implementation, the pixels at a distance of *safetyMargin* pixels or less from the border of the image are not trained nor evaluated with the ferns. They are automatically set as BG. An improvement would be to make tests fit with these situations where a part of the patch. However, for a value of the safety margin corresponding to the values of R_{Patch}, i.e. $safetyMargin = 30$, this choice does not lead to any classification mistake in the corresponding part of the GT.

The set of parameters used in both the training and the evaluation of ferns is available in Table 4.2. The tests will be drawn uniformly at random in the set of tests T, consisting of the tests providing the greater Bhattacharrya distances between FG and BG.

Parameter	Value
R_{Patch}	15
R_{Border}	20
$safetyMargin$	30
N	10
M	200
T	$\{T_{S_{Coh},R_{Coh}}, T_{N_T,S_\Delta}\}$
S_{Coh}	$\{0.1, 0.2, 0.45\}$
R_{Coh}	30
N_T	10
S_Δ	$\{3\}$
S_{Ratio}	$\{0.2, 0.3, 0.4, 0.5\}$

Table 4.2: Definition of the parameters of validation.

The training set will contain one half of the images in category A, B and C, selected at random. Therefore, 47.5% of the set will be trained. The images from set D are willingly discarded in order to avoid their training to drastically decrease the distance between the different classes, since a large part of their FG is almost of the same noisy nature as the BG. However, the results will evaluate the quality of the segmentation on all sets and on all untrained images.

Also, the images will not be evaluated on all pixels but in a grid of pixels separated from their neighbours by $Step_{Eval}$ pixels. The rest of the pixels take the same scores as their nearest neighbour on the grid. The value of $Step_{Eval}$ must remain small in order to be precise but can significantly reduce the complexity. In practice, a value of $Step_{Eval} = 3$ is chosen since it almost reduces the computational time of evaluation of one order of magnitude in comparison with a full evaluation, without loosing too much accuracy.

4.2.1 Pre-processing

First of all, before an image is trained or evaluated, the eventual light of the device present in the picture is captured and all the pixels that belong to it are classified as BG. This is made possible because the structure has a very recognizable shape and does not vary much since its relative position to the camera is almost unchanged. On the original image, it appears as a vertical bar of very high intensities. Therefore in the second set, as it is removed from the reference frame, it appears in very low intensities, in the negative actually, as shown in Figure 4.4.

The width of the bar has been measured directly on the images and is about 20 pixels. Therefore a specially designed filter will be used to capture the fact that a pixel belongs to it. An horizontal slice of this filter is shown in Figure 4.3. First, it takes high values for about 20 pixels to give positive weight to high intensities, then negative values for about 30 pixels. This large range is necessary because a lot of areas inside the fingerprint take high intensities for about the same width. Therefore, they must be discarded. This is the case when filtering over a long range to lower the filter value as new ridges of high intensities are encountered.

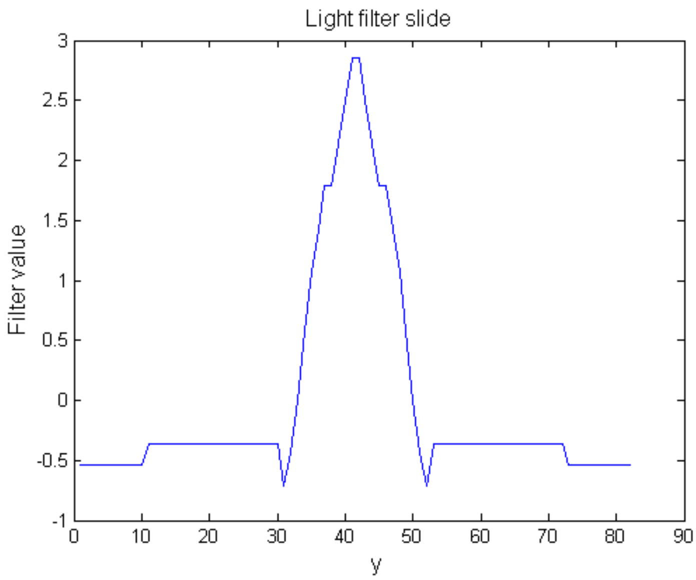


Figure 4.3: Slice of the light filter

Finally, the filtered image is binarized over a threshold $Tresh_{filter}$ before a morphological opening in the shape of the device but smaller, i.e. a rectangle of width W_{light} and height H_{light}. The parameters are chosen to minimize the number of light pixels actually belonging to the FG while maximizing the number of pixels of the light on the whole ground truth. The graphs of the parameter optimization presented in Annex A.1 yield optimal values of $W_{light} = 9$, $H_{light} = 70$ for $Tresh_{filter} = 200$ as well as an optimal number of slices, i.e. the height of the filter, of 100. A light segmentation is shown in Figure 4.4.

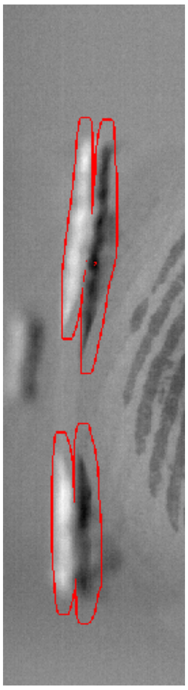


Figure 4.4: Light of the device in an image of Set 2. In red, the area classified as light.

4.2.2 Post-processing

Two steps are found in the post processing. First, the scores of the different classes are adjusted with dynamic shifts of the different class scores for the evaluated pixels. The shifts are based on means of the scores of the different classes in the image and are performed so that the means of the classes take the same value.

Second, a serial of morphological operations are performed on the resulting CA, in order to smooth the results out and to avoid including too much pixels in the FG, especially parts of the noisy image that are mistaken as a part of the fingerprint. These are successive opening and closing operations with increasing size of structuring element, as well as only saving the largest connected component. Graphical illustrations will be presented in the next section alongside the resulting images.

4.3 Results

This section presents the results obtained with the method on the whole Ground Truth, and separates the different values between the four categories previously defined. First, some values are given in terms of false positive and true positives rates for the classification obtained without altering the scores as well as graphical representations. Then one can find the comparative study with the situation of an alignment of the different class scores according to their mean value. Afterwards, a short overview of the morphological operations that can improve the segmentation is described. To conclude with, a review of the different results defines the limitations of the method and the corresponding challenging situations.

First, let us define the True Positive Rate (TPR) and the False Positive Rate (FPR):

$$TPR = \frac{\sum \text{True positive}}{\sum \text{Condition positive}}, \quad FPR = \frac{\sum \text{False positive}}{\sum \text{Condition negative}} \tag{4.1}$$

In this case, the condition to fill is to be classified as part of the FG mask. Since there are four classes, the latter consists of both the pixels classified as FG and BFG by default. Two classifications are considered in this section. On the one hand, the one that considers the different class independently and defines the FG mask as the union of FG and BFG. On the other hand, the one that sums the scores of FG and of BFG and compares it to the sum of the scores of BG and BBG to determine if one pixel belongs to one group or to the other. This is referred as the *SUM* classification.

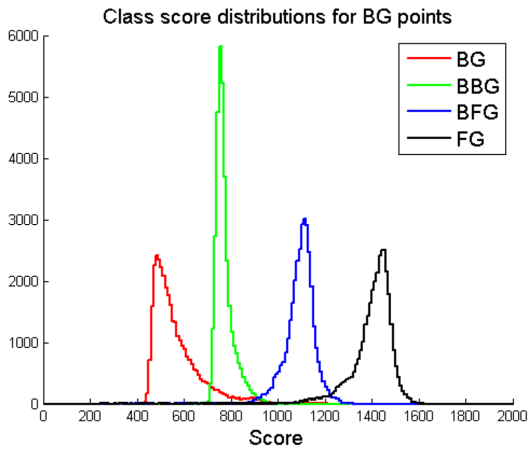
When not performing any shift on the scores before classifying a pixel, one can see the obtained mean values of TPR and FPR in Table 4.3 with respect to the distinct categories. This table indicates that the quality of the classification is strongly dependant of the category, i.e. of the contrast of the fingerprint. Notably, the results of categories C and D are very poor. In D, the contrast is very low, and the values indicate that almost no pixels are classified as FG or BFG, which is expected since the hypothesis that the prior is uniform is not exact. Therefore, it shows that the scores need to be readjusted in order to give more importance to the FG and BFG in low contrasted images. However, one should notice that SUM takes better values in any cases, i.e. lower FPR and higher TPR.

Category	Default Classification	Sum Classification
A	0.969/0.013	0.976/0.01
B	0.832/0.059	0.852/0.052
C	0.49/0.19	0.52/0.18
D	0.02/0.178	0.03/0.176

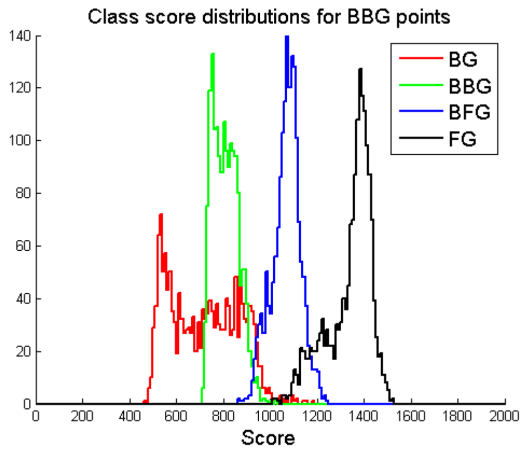
Table 4.3: Mean value of FPR and TPR of the different image categories. First entry is FPR and second entry TPR.

To understand the origin of this problem, Figure 4.5 displays the distributions of class scores for pixels that are known to belong to each class for an arbitrary image of category D called *sampleD*. The image used and its default classification are displayed in Figure 4.6. First of all, let us recall the images in D were not trained and therefore the chances of a good classification are obviously decreased. However, this allows to understand the importance of contrast in the quality of the method.

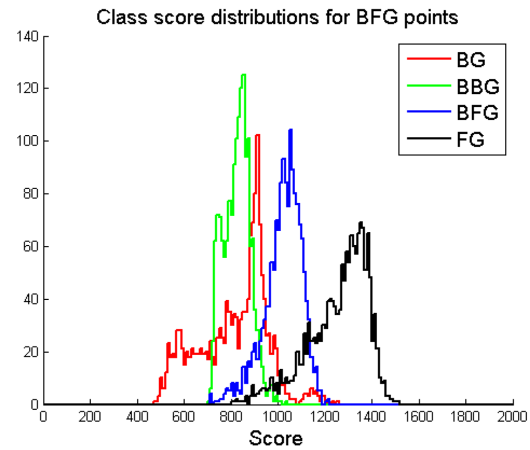
The Figure 4.5 shows that the scores of BG and BFG are almost always smaller than the other two classes. While the BG shows very distinguishable distributions with single peaks due to the low level of noise in the BG of *sampleD*, the FG has distributions that are over-crossing, with two preaks in BG and BBG. These peaks have a simple origin. The left peak corresponds to the values that behave like the BG in the training set, and the second one to the values that do not and that whose fern values rather correspond to the other classes. This separation in two peaks allows a proper classification when no class has two peaks left from the ones from another class as it is the case here. One can also notice that the parts of the foreground with a good contrast are creating a spread of the scores values of FG to the left between Figure 4.5a and Figure 4.5d.



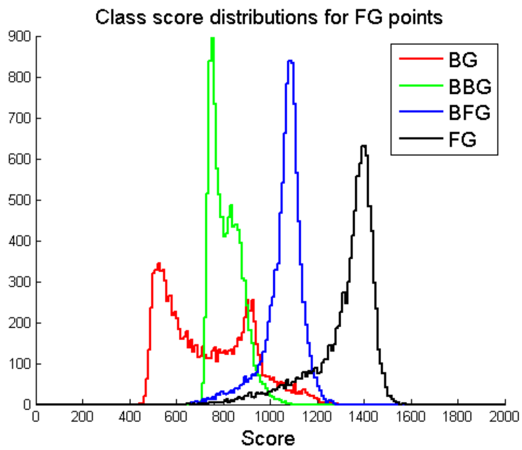
(a) Scores distributions for BG pixels



(b) Scores distributions for BBG pixels

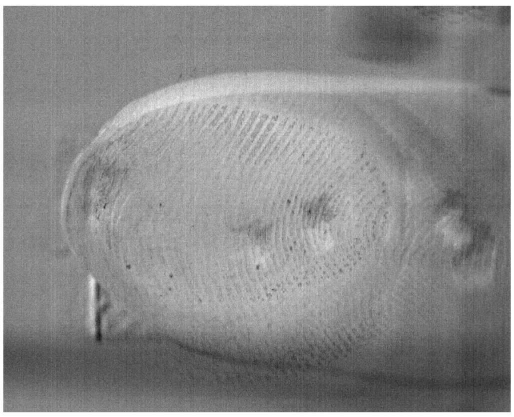


(c) Scores distributions for BFG pixels

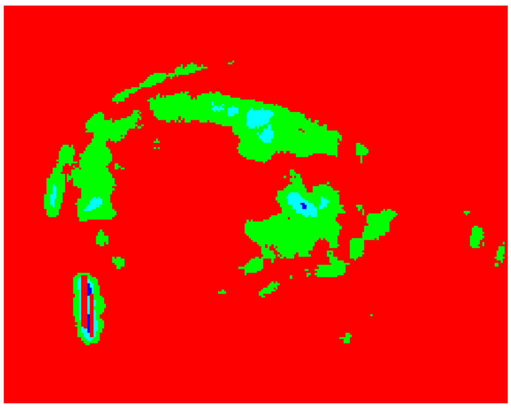


(d) Scores distributions for FG pixels

Figure 4.5: Distributions of scores for *sampleD*.



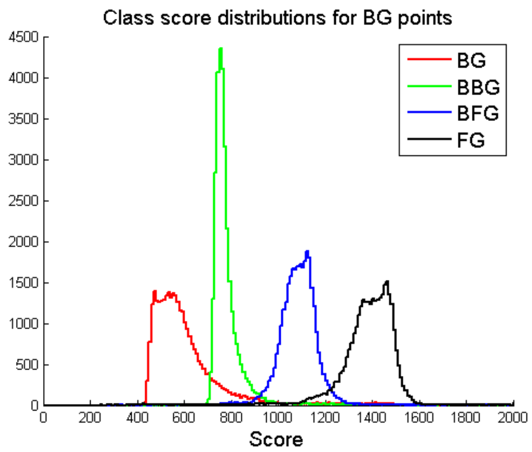
(a) *sampleD*



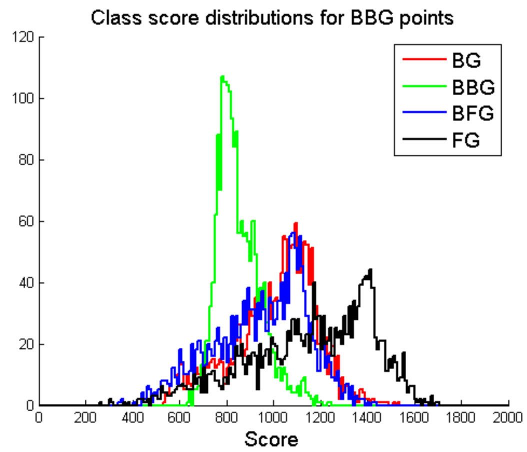
(b) Classification

Figure 4.6: *sampleD* and its default classification.

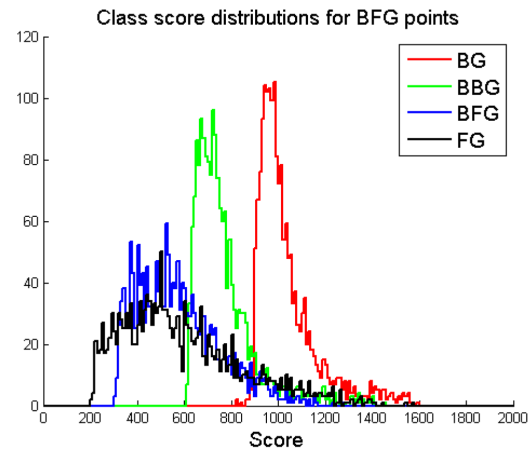
At the opposite, the same distributions are shown in Figure 4.7 for an image of set A called *sampleA* displayed with its default classification in Figure 4.8. One can see that the classification is quite accurate, except for the neighbourhood of the light device. As far as the histograms are concerned, the separation is also visible on all graphs. In case of a good contrast, the tests capture the difference between BBG and BG but the histograms of FG and BFG are closer due to their similar nature. The scores distributions are ordered according to the nature of the pixel: the ordering of their peak values is $\{BG, BBG, BFG, FG\}$ in case of BG pixels in Figure 4.7b and becomes $\{FG, BFG, BBG, BG\}$ in case of FG pixels in Figure 4.7b. One can also notice that the scores of FG are as low as 200 *sampleA* whereas they were at least taking values of 600 for *sampleD*. The lower the scores, the more certain is the classification.



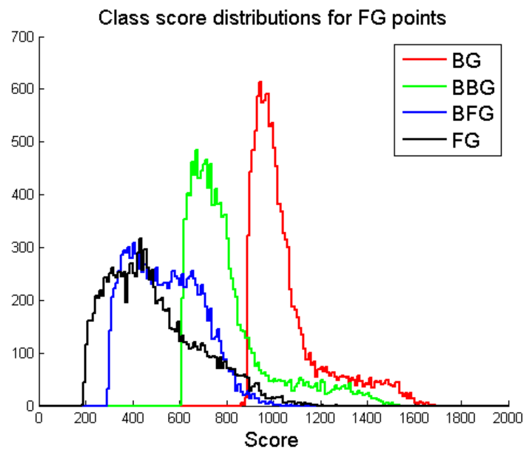
(a) Scores distributions for BG pixels



(b) Scores distributions for BBG pixels



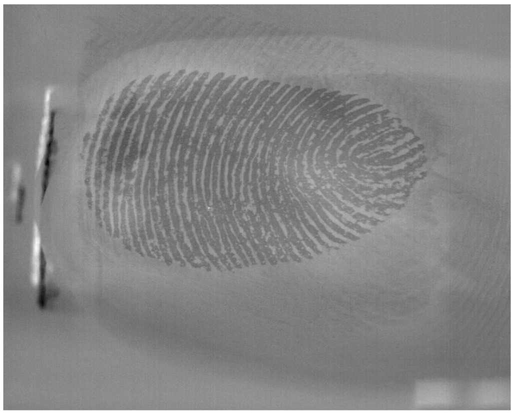
(c) Scores distributions for BFG pixels



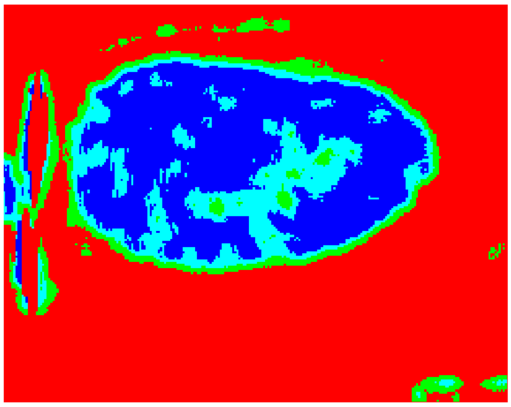
(d) Scores distributions for FG pixels

Figure 4.7: Distributions of scores for *sampleA*.

The next step is to adjust the distributions with appropriate shifts. The idea is the following: in the idealized case of two distinct peaks for each distributions, corresponding to the tests capturing the different cases, one shoud have the peak corresponding to the true class to the left of a certain value and all the other ones to the right. Therefore, all the scores are redefined as follows. This step is a part of the post processing since the whole distributions of scores of the image must be known. Each score s_{c_i} of a pixel now takes the value $s_{c_i} - \bar{s_{c_i}}$ where $\bar{s_{c_i}}$ is the mean of s_{c_i} over all evaluated pixels.



(a) *sampleA*



(b) Classification

Figure 4.8: *sampleA* and its default classification.

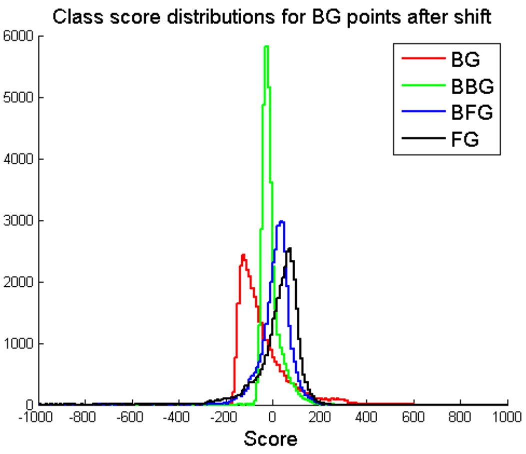
The new values of with the shift of the values are available in Table 4.4. One can see that they are drastically better than previously, especially for categories C and D. Yet, the results in D are still far from perfect. SUM still provides the best values of TPR and FPR, and is shifted according to the mean of the summed distributions. The latter classification will be the one used in practice.

Category	Default Classification	Sum Classification
A	0.98/0.011	0.987/0.007
B	0.943/0.026	0.954/0.021
C	0.929/0.038	0.944/0.031
D	0.711/0.054	0.732/0.051

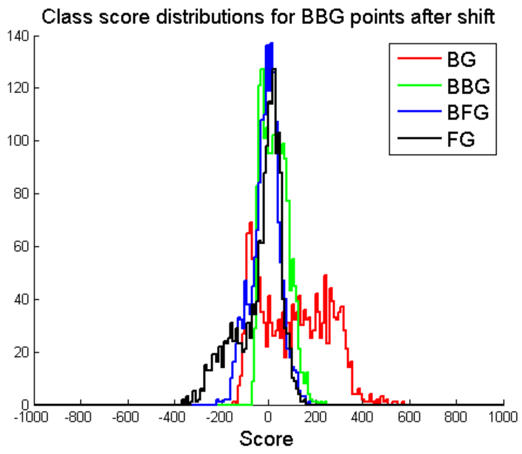
Table 4.4: Mean value of FPR and TPR of the different image categories with the score shift. First entry is FPR and second entry is TPR.

In order to compare the results, one can find in Figure 4.9 the shifted histograms of *sampleD* and the corresponding classification in Figure 4.10b. One can see on latter that the contrasted area in the contact zone is indeed classified as part of the fingerprint. However, in addition to the mistakes in the BG at the borders of the fingers and at marking points, the very low contrast at the center of the fingerprint is considered as background. When looking at the histograms, one can see in Figure 4.9d that the FG is now dominant on the left part of the scores, which was not the case in 4.5d. This step has therefore promoted the FG, in spite of its high scores in absolute. Despite that, the border classes are much less dominant, which corresponds to the re-evalution of the uniform prior hypothesis. However, they only are present in a very small transition band between FG and BG, in comparison with r_{Border}.

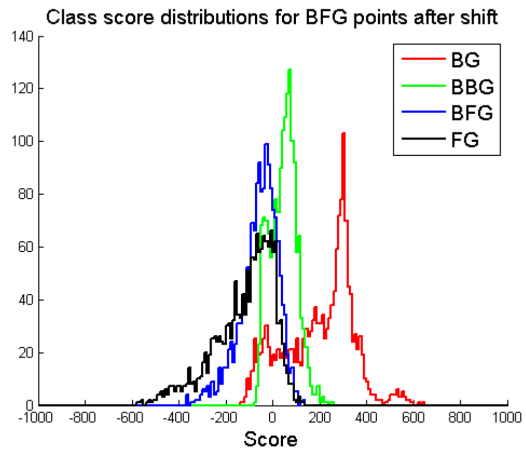
Finally, the recapitulating Figure 4.11 shows the couples (FPR, TPR) for all the images in GT, with and without shift. One can clearly see the impact of the category, as well as the one of the shifting of the scores. The last step is to apply morphological operations to the obtained masks to decrease FPR and increase TPR.



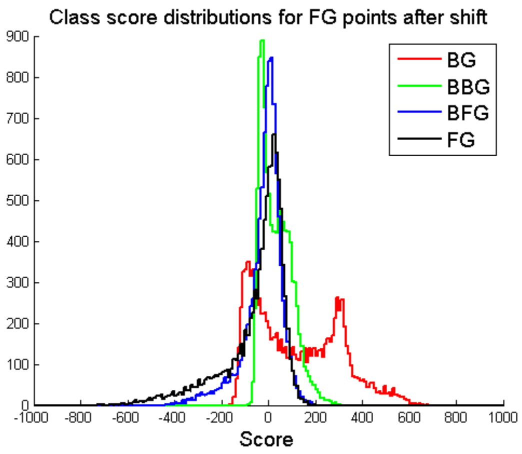
(a) Scores distributions for BG pixels



(b) Scores distributions for BBG pixels

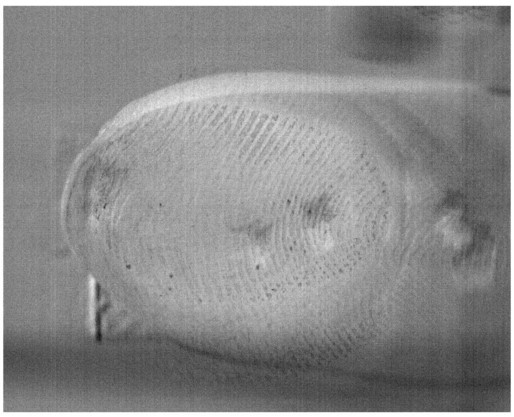


(c) Scores distributions for BFG pixels

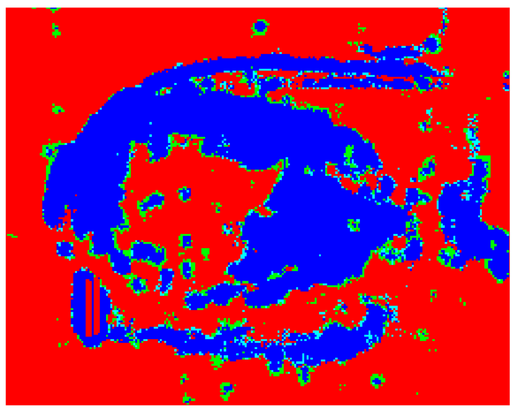


(d) Scores distributions for FG pixels

Figure 4.9: Distributions of scores for *sampleD* after the shifting of the scores.

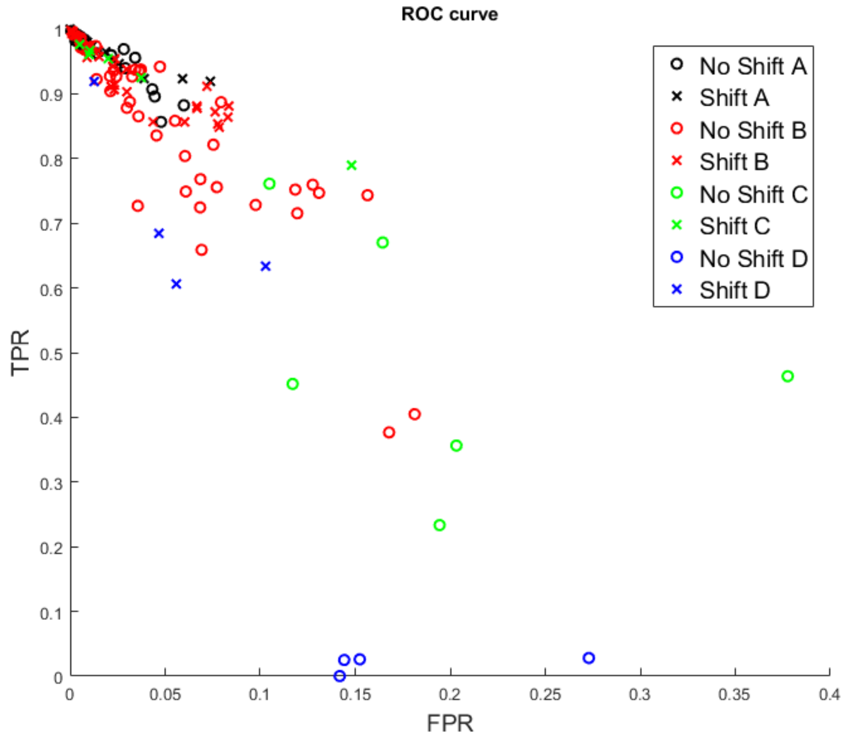


(a) *sampleD*

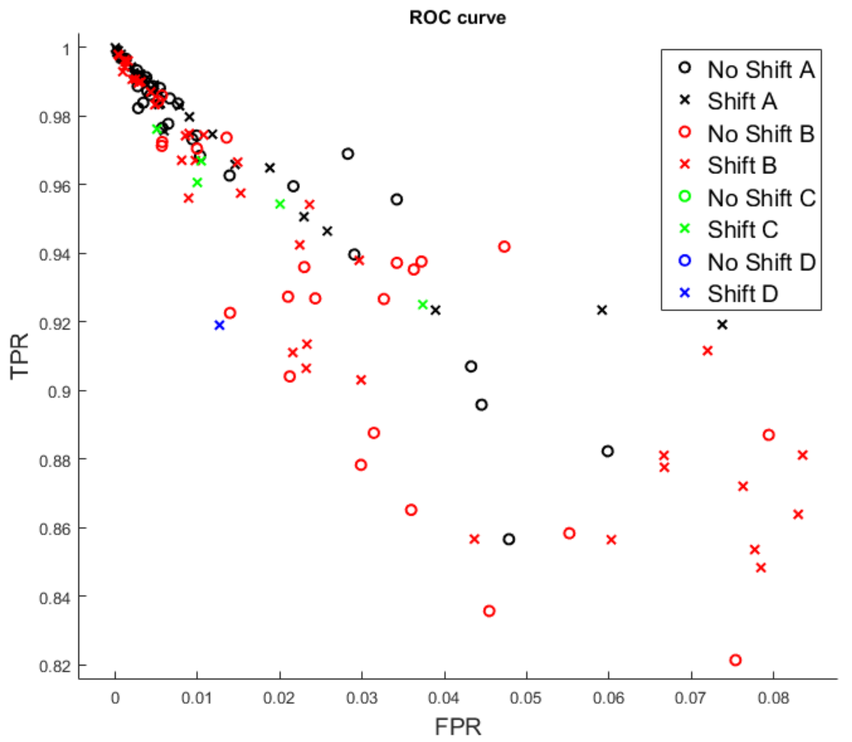


(b) Classification with shifted scores

Figure 4.10: *sampleD* and its classification after the shifting of the scores.



(a) ROC Curve



(b) Zoom on $(FPR, TPR) = (0, 1)$

Figure 4.11: ROC curve FPR vs TPR for all GT. Each color is a category of images.

The morphological operations are for different purposes. First, they allow to have clustered fingerprints, which is necessary, in most cases, to fit the physical reality. To do so, one will only save the largest connected component of the image that is classified as FG. Sometimes unfortunately, the largest component is not corresponding to the real contact area, notably for poorly contrasted pictures. However, this will still be used in practice because of the large improvement it creates in most classifications, especially with respect to the feasibility of the CA.

Second, they allow to have contact areas without any holes inside of it, which is most unlikely in reality. Finally, they allow to make the FG smoother since a fingerprint has curves but is not angular, with a closing for instance. Also, it prevents the elliptic shape of the CA to have peaks following some fingerprint trail or other spurious fingerprint with an opening for instance, or an erosion.

The following morphology is used to generate results of Table 4.5:

- 1. Erosion with a disk of radius=15
- 2. Saving the largest connected component
- 3. Fill holes
- 4. Opening with a disk of radius=30
- 5. Saving the largest connected component
- 6. Closing with a disk of radius=30

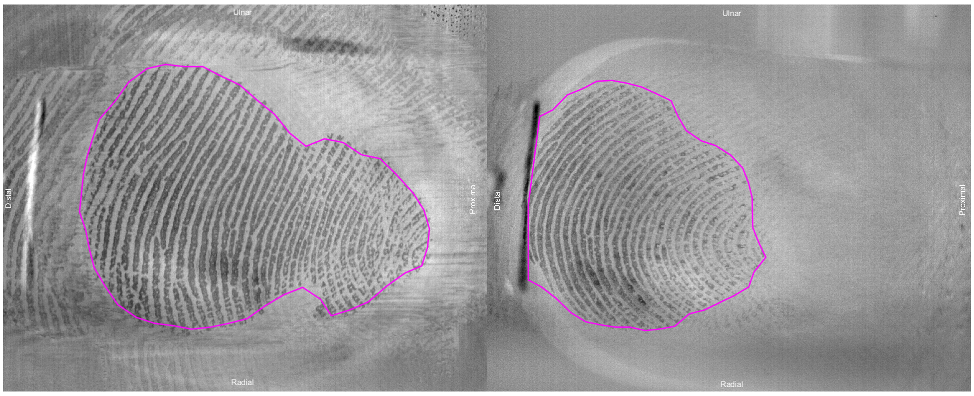
The first step is heuristic, the experience showed that the masks are a little larger than the CA and therefore one should take profit of it to erode and break bridges with false positive areas in the background. Therefore, the second step removes all isolated points of the mask and patches that can be due to the light device, the border of the finger, or spurious fingerprints. The third step is obvious but can be a cause of trouble if a large area of the BG is enclosed in the mask. The three last steps form one loop, of parameter $radius = 30$. Once more, they first remove connected false positive areas in steps 4 and 5 to finally smooth out the results in step 6. The utility of the different steps is illustrated in Figure A.8 and Figure A.9.

When studying the entries in Table 4.5, one can notice that the results are still correct, especially for the categories A and B, but most of all that all the values are smaller than the one in Table 4.4. This may be surprising at first sight, however the pixels are classified independently from the global FG shape and the fact that. Also, the morphology must not discard too many pixels in case of low-contrasted images nor too few in case of high-contrasted images. As the variety in the images is large, the morphology presents limitations described below. However, it is important to note that the values are not reflecting the fact that the FG is physically possible, which is here expected to be ensured with mathematical morphology.

Category	Sum Classification
A	0.937/0.029
B	0.888/0.048
C	0.81/0.106
D	0.576/0.081

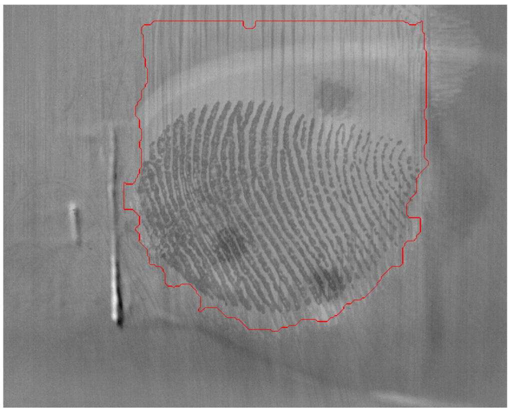
Table 4.5: Mean value of FPR and TPR of the different image categories with the score shift, after morphology. First entry is FPR and second entry is TPR.

To conclude this chapter, let us review some resulting segmentations in order to determine the limitations of the method. First, two accurate masks are displayed in Figure 4.12 in case of spurious fingerprints and of a CA next to the light of the device. Afterwards, and most importantly, one can see in Figure 4.13 that the morphology does not perfectly complete the task. In Figure 4.13a, one can see the most problematic situation throughout the results: if a clear trail of the fingerprint is directly next to the fingerprint, then it is classified as FG and the morphology is powerless because of the force of the connection between them. In Figure 4.13b, one can observe that for a disparate fingerprint, the masks will be as well and the opening operation separate the CA into separate entities. As one takes the larger connected area, the algorithm looses a great part of the mask when performing morphology. Finally, Figure 4.13c shows that the hypothesis that the largest component corresponds to the FG is not true. Worse, all the CA is discarded if it is indeed incorrect.

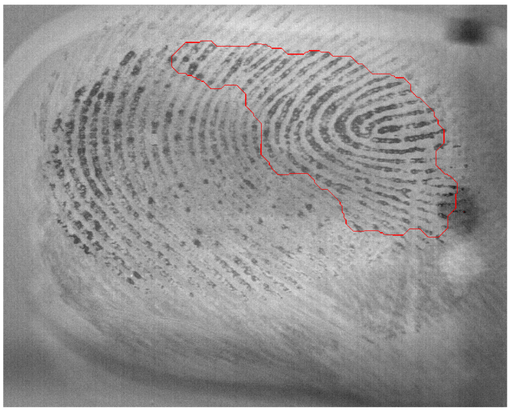


(a) Spurious fingerprints are not selected (b) Light is handled

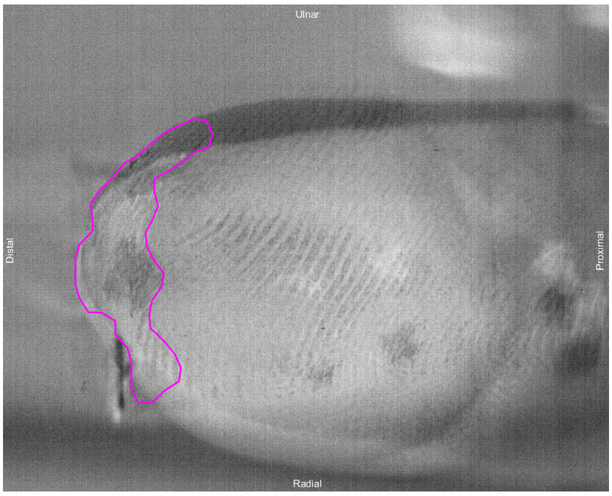
Figure 4.12: Illustration of the algorithm performance.



(a) Clear trails directy next to the CA



(b) Disparate contrast in the fingerprint



(c) Wrong selection of the larger connected component of FG

Figure 4.13: Algorithm limitations and morphology.

Chapter 5

Conclusions and perspectives

This thesis had for objective the correct segmentation of the contact area of the finger tip of a patient with a glass plate, potentially lubricated, despite the numerous states of the latter in term of contrast and perturbations in the rest of the image. Moreover, this work has contented itself to the use of random ferns to study their utility in this particular context. It also uses four classes instead of two in order to measure the utility and the precision of the ferns as well as reaching a deeper insight of the texture behaviour at the borders of the contact area.

Random ferns are a tool that capture the texture based on several binary tests on points taken at random in a direct neighbourhood. They are first trained on the ground truth in order to learn distributions of the fern values, and second evaluated when classifying a new image, generating for each pixel scores for each class based on the conditional probabilities of fern values for each of them, computed in a semi-naive Bayesian way.

The binary tests relay on useful discriminating features. The ones that have led to the best separation of the different classes are based on absolute deltas of intensities and on coherence. According to pivotal values in the histograms of the different features class distributions, the tests have been designed to separate the fern value distributions of FG and BG as well as those of BFG and BBG as far as possible, in terms of distances of Bhattacharyya.

A comparison with local features like ULBP teaches that the separation is greater when considering broad-area features. This is due to the ridge-valley structure of the fingerprint that needs large considered area to be efficiently captured without too many false positives. However, one should not take too large areas if he wants not to lose the pixel-wise precision.

Here, the main criterion to select the tests have been the distance between FG and BG. Despite the fact that the border classes are indeed leading to results where they capture the expected information, confirming at the same time that the tests are able to separate all classes, the four classes may not be necessary. In particular, BFG is very similar to FG. A further work would be to remove BBG and BFG, or maybe to do a single border class that contains them both, or just keep BBG. In the end, the classes are summed to obtain the best results. The multiplication of classes is therefore probably not necessary, even with maintaining the same set of tests.

A few elements have been used to harmonize the data captured by the ferns and to improve the results. First, the patches have all been rotated according to their main gradient towards the same direction, leading to better results. Second, a recurring misleading element, the light produced by the device holding the finger still, is handled before training or evaluating an image and the corresponding pixels are classified as BG. In order to correct the hypothesis that the prior on the different classes is the same, the class scores obtained for each pixel are shifted left by their corresponding mean value. This dramatically improves the results. Finally, morphological operations are performed to improve the resulting segmentation.

The method has proved to be accurate for more than 90% of the pixels in average for high-contrasted images. Even for terribly bad-contrasted images, it is accurate for about 50% of the pixels in average after morphology and 70% of the pixels before that. These results are appreciable since the training set only contains 38 images out of 80 with 6000 pixels trained for each class and for each image. This corresponds to a small 0.87% of the total number of pixels in the GT.

However, the method presents several limitations. First, the tests are quite time-consuming due to the selected features that require several operations each. Second, difficulties are encountered when doing the morphology that are due to the nature of the images and the limitations of the data set. Nevertheless, it is necessary to perform this final post-processing as the resulting masks must be physically meaningful. This step is maintained, even though the segmentations are better in terms of TPR and FPR without it.

During the realization of this work, a large amount of leads towards a better solution were noticed as the results were growing in numbers. Some of them are introduced below.

First, a multi-scale approach could be especially appropriate in a fingerprint context. Indeed, this could both reduce the computational cost and make the local features like the ULBP account for broad-areas characteristic behaviours of the contact area in a large scale, while performing a local study at a small one. A solution could also classify sub-blocks of areas of decreasing sizes based on the previous classification, in an iterative fashion while preserving the connectivity of the solution. This could for instance avoid a split of the FG as in Figure 4.13b.

Second, other features can be used, in particular the Frequency domain of a sub-area of the picture. [5] An illustration of the amplitudes in the frequency domain after the 2D Fourier Transform is available in Figure A.10 after that a band-pass filter has been applied. One can see different behaviours according to the noise in the image. The larger magnitude are concentrated in case of a clean picture, whereas they are spread in all directions for a noisy one.

Finally, the results before mathematical morphology are very good actually, especially for this relatively low number of pixels trained. This step can definitely be improved with a deeper study. For instance, one could use an energy minimization based on line integrals of scores [4] to differentiate between several large distinct areas in the FG. A last idea would be to perform a region-growing algorithm starting from the remaining point after drastic erosion of the mask, with a criterion based on the mean of the current region and of the initial mask. This could help solving the problem in Figure 4.13a in particular.

Bibliography

[1] A. Barrea, E. Jarocka, P. Lefèvre, J.-L. Thonnard, B.B. Edin, and B.P. Delhaye. "Human tactile afferent responses to skin strain patterns caused by fingertip shearing". *Program No. 151.02. 2016 Neuroscience Meeting Planner. San Diego, CA: Society for Neuroscience*, 2016. Online.

[2] A. M. Bazen and S. H. Gerez. "Directional Field Computation for Fingerprints Based on the Principal Component Analysis of Local Gradients". In *ProRISC 2000 Workshop on Circuits, Systems and Signal Processing, Veldhoven, The Netherlands*, 2000.

[3] A. M. Bazen and S. H. Gerez. "Segmentation of Fingerprint Images". In *ProRISC 2001 Workshop on Circuits, Systems and Signal Processing, Veldhoven, The Netherlands*, 2001.

[4] A. Browet, C. De Vleeschouwer, L. Jacques, N. Mathiah, B. Saykali, and I. Migeotte. "Cell Segmentation with random ferns and graph-cuts". In *IEEE International Conference on Image Processing*, 2016.

[5] H. Choi, M Boaventura, I. A.G. Boaventura, and A. K. Jain. "Automatic Segmentation of Latent Fingerprints". In *2012 IEEE 5th International Conference on Biometrics: Theory, Applications and Systems, BTAS 2012*, pages 303–310, 2012.

[6] H. Flayeh, D. Jomaa, and M. Dougherty. "Segmentation of Low Quality Fingerprint Images". In *2010 International Conference on Multimedia Computing and Information Technology (MCIT)*, pages 85–88, 2010.

[7] S. Greenberg, M. Aladjem, and D. Kogan. "Fingerprint Image Enhancement using filtering Techniques". *Elsevier Schience Ltd*, 2002.

[8] Michiel Hazewinkel. *"Encyclopaedia of Mathematics, Supplement Volume I"*, pages 121–122. Springer Netherlands, 1997.

[9] A. K. Jain and N. K. Ratha. "Object detection using gabor filters". *Pattern recognition*, 30(2):295–309, 1997.

[10] M. Kass and A. Witkin. "Analyzing oriented patters". *Computer Vision, Graphics, and Image Processing*, 37(3):362–385, 1987.

[11] T. Ojala, M. Pietikainen, and T. Maenpaa. "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns". *IEEE transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.

[12] M. Ozuysal, M. Calonder, V. Lepetit, and Pascal Fua. "Fast Keypoint Recognition using Random Ferns". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32:448–461, 2010.

[13] P. Perona. "Orientation diffusions". *IEEE Transactions On Image Processing*, 7(3):457–467, 1998.

[14] Université Catholique de Louvain. *Ma thèse en 180 secondes - Finale UCL 2016 - Allan BARREA*, March 2016. Online. url : https://www.youtube.com/watch?v=cjSz3IfWU9w.

[15] Chengming Wen, Tiande Guo, and Yuyang Zhou. "A Novel and Efficient Algorithm for Segmentation of Fingerprint image based on LBP Operator". In *International Conference on Information Technology and Computer Science*, volume 2, pages 200–204, 2009.

[16] F. Zheng and G.I. Webb. "A comparative study of semi-naive bayes methods in classification learning". In *In Proceedings of the Fourth Australasian Data Mining Conference (AusDM05)*, pages 141–156, Sidney, 2010.

Appendix A

Additional Figures

A.1 Light filter parametrization

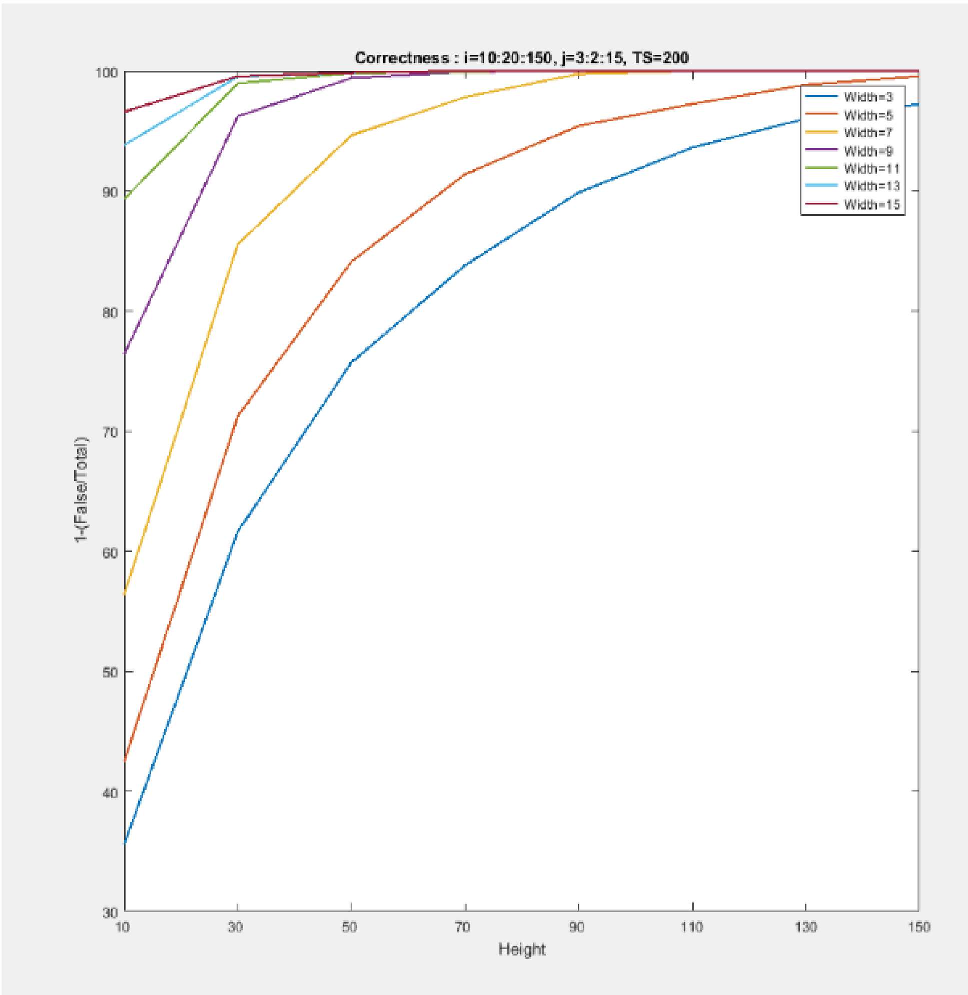


Figure A.1: Correctness in percents defined as $1-p$ where p is the proportion of pixels that are not in BG and selected as part of the light device, as a function of W and H values, parameters of the opening structuring element $H \times W$. i is the vector of H values tested. j is the vector of W values tested. $(H, W) = (90, 7)$ offers good results with $p \approx 0$

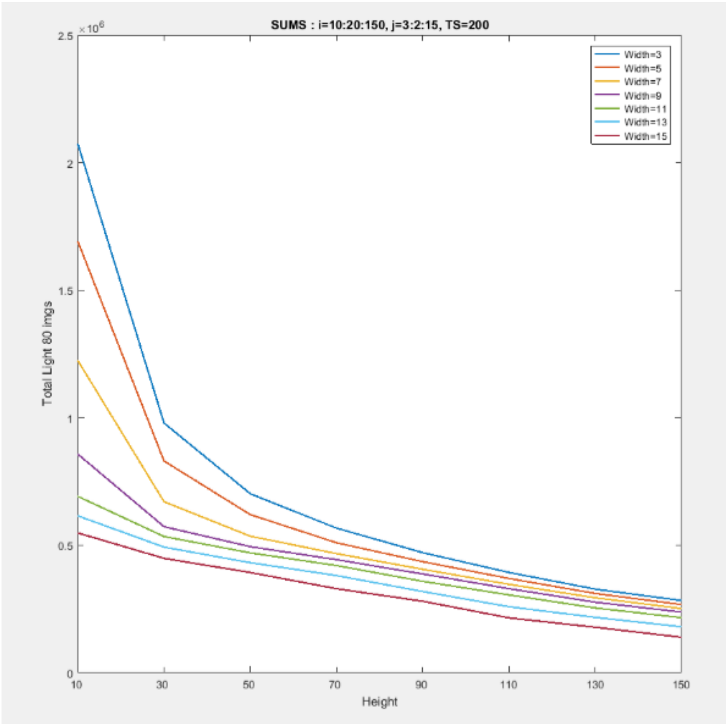


Figure A.2: Total number of pixels of light, as a function of W and H values, parameters of the opening structuring element $H \times W$. i is the vector of H values tested. j is the vector of W values tested. The number of pixels decreases fast at first then slowly decreases while H increases.

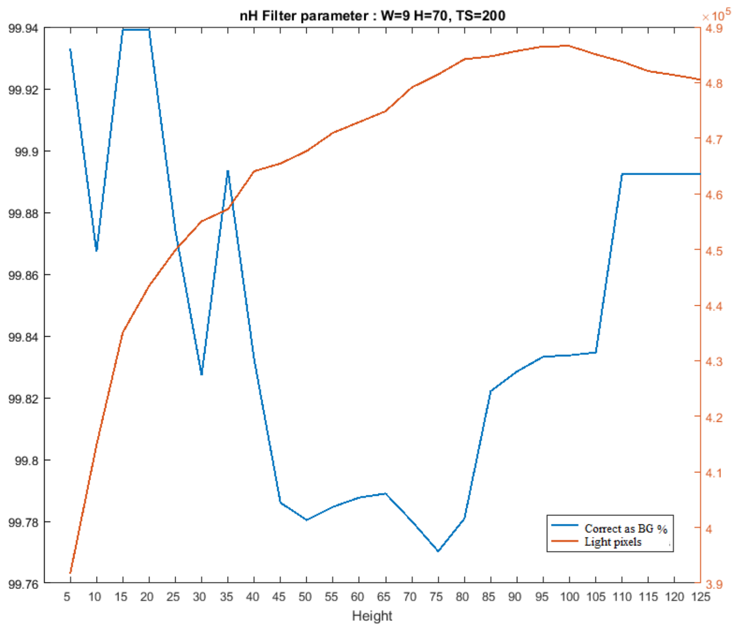


Figure A.3: Optimal height of the filter, i.e. number of slices nH. Total number of light pixels in red and correctness in percents defined as $1 - p$ where p is the proportion of pixels that are not in BG and selected as part of the light device, in blue. Results obtained for the set of parameters indicated.

A.2 Illustration of image categories

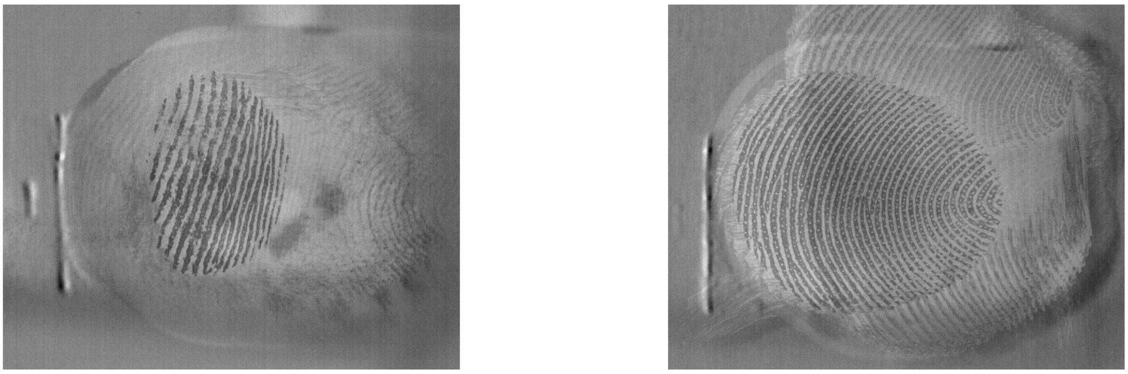


Figure A.4: Images from Set 2 and category A.

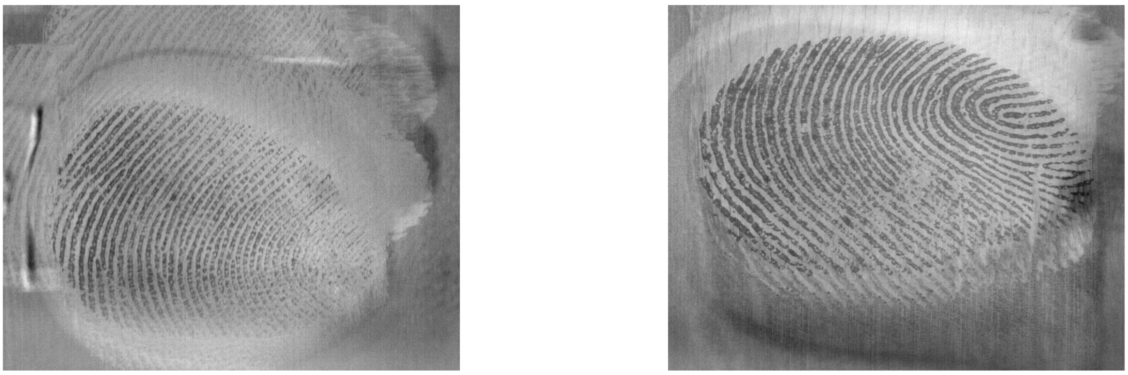


Figure A.5: Images from Set 2 and category B.

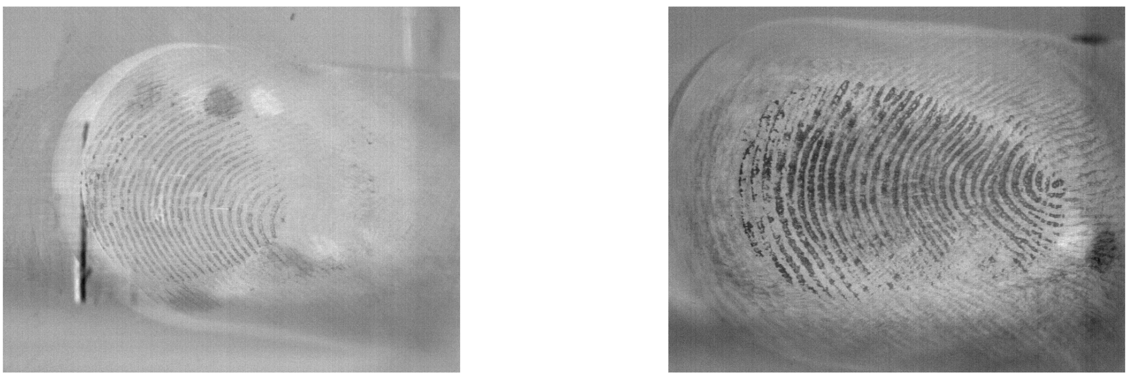


Figure A.6: Images from Set 2 and category C.

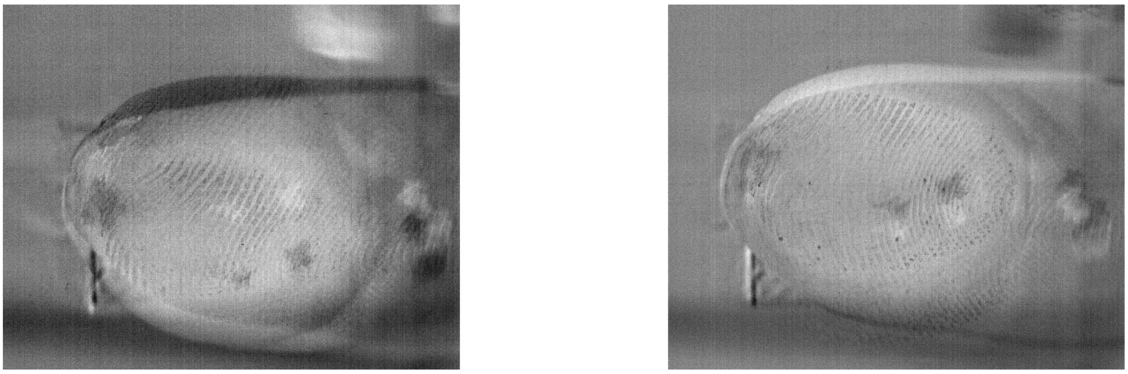
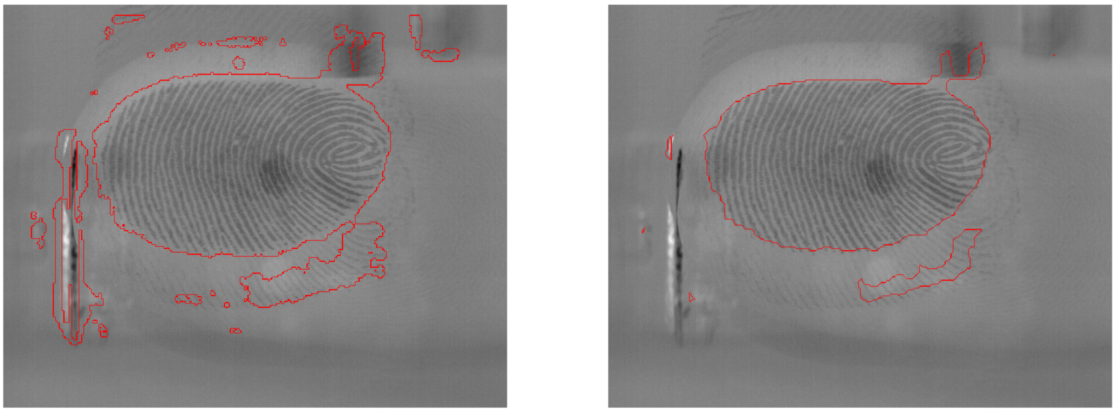


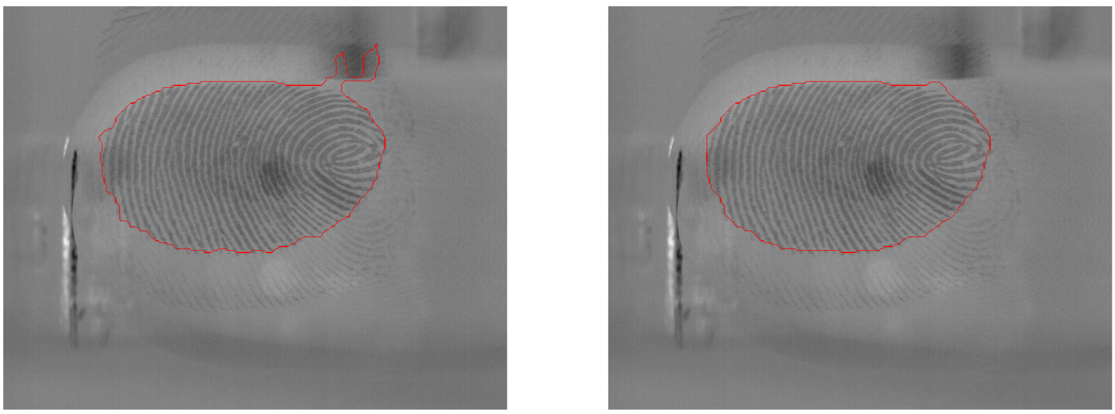
Figure A.7: Images from Set 2 and category D.

A.3 Morphology steps



(a) Initial mask

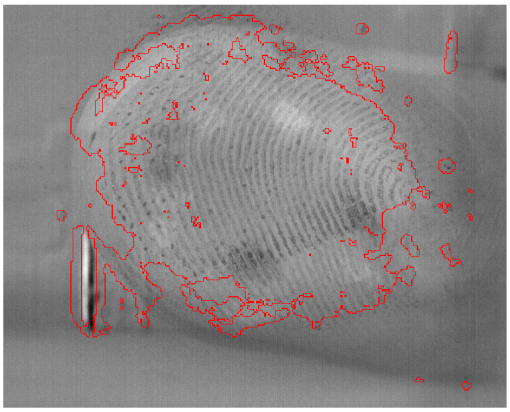
(b) After Step 1

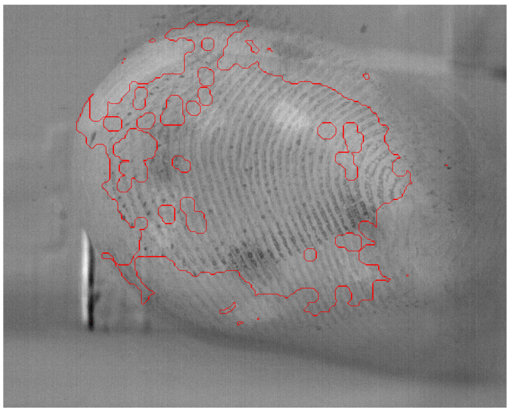


(c) After Step 3

(d) Final mask

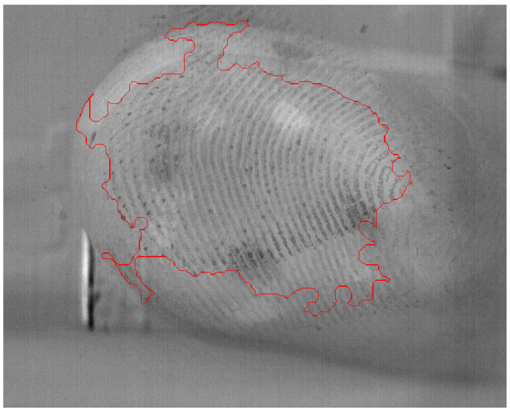
Figure A.8: Illustration of the steps of morphology on the FG mask.

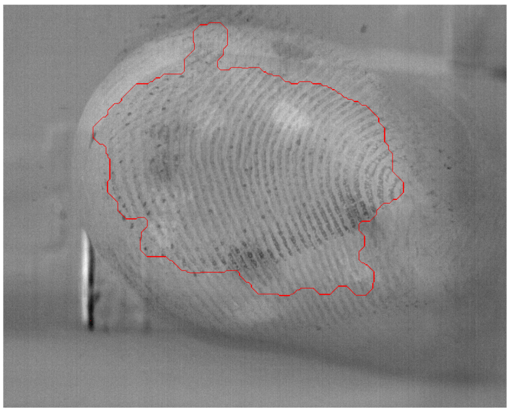




(a) Initial mask

(b) After Step 2



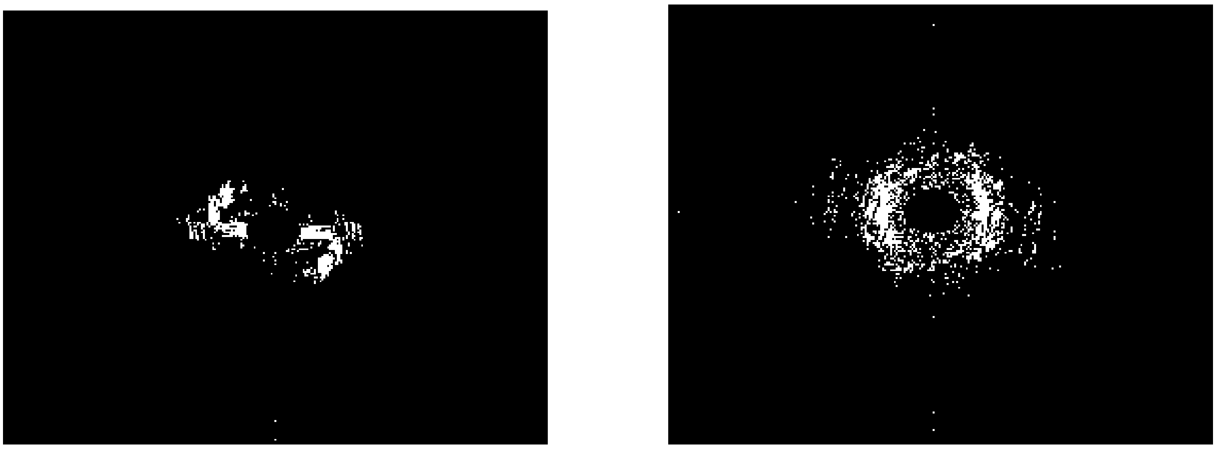


(c) After Step 3

(d) Final mask

Figure A.9: Illustration of the steps of morphology on the FG mask.

A.4 Frequency Analysis



(a) Clear, high-contrasted frame

(b) Noisy, low-contrasted frame

Figure A.10: Frequency analysis after 2D Fourier Transform after logarithmic transformation, band-pass filter and binarization over some threshold. Left, an image of excellent quality with light of the device visible. Right, a noisy image with spurious fingerprints but no visible device.

Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve www.uclouvain.be/epl