Generating Realistic Ghost Fingerprints by Combining Real Fingerprint Images

J.E. Huiden - s1828673
Bachelor assignment committee:
Dr.ir. L.J. Spreeuwers , Prof.Dr.Ir. R.N.J. Veldhuis and Dr. C.G. Zeinstra
Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS)
Datamanagement & Biometrics (DMB), University of Twente, 7500 AE Enschede, The Netherlands
June 28, 2019

***Abstract*—Fingerprints are widely used biometrics in security systems and law enforcement. It is essential that possible distortions in the measurement of fingerprints are detected before the images are saved or used, to prevent incorrect matches. One possible distortion in fingerprints is called the ghost fingerprint. To train and test ghost fingerprint detection algorithms, a large amount of data is needed. Obtaining this data is not always possible due to cost, time and privacy issues. This paper proposes a method to create realistic ghost fingerprints by combining real fingerprint images which can be used as data for ghost fingerprint detection algorithms. This method uses masks to determine the region of overlap and threshold binarisation to replicate the overlapping lines as seen in real ghost fingerprint images. To create variation in the created fingerprints different transformation such as scaling, rotation, position, and changing line thickness are, within a range, randomly applied. The result is a system that is able to generate a large variety of realistic looking ghost fingerprints. Three databases are created with different classes separated by the percentage of ghost presence, the intensity of the ghost, and a combination of both. These databases are fed into two ghost fingerprint detection algorithms; one based on the local binary pattern, and one on frequency estimation. Comparing the results with real ghost fingerprints shows that the accuracy of the detection of the artificial ghost fingerprints is lower than the detection of real ghost fingerprints. The proposed separation parameters for the classes have effect on the accuracy of detection, however not as much compared to the different classes of the real ghost fingerprints.**

I. Introduction

Personal identification is currently widely used. Passwords and cards are used to gain access to different websites, buildings, money, and other personal information or property. These methods are fast and easy to use. However, they are not the safest options. Passwords can be cracked, and cards can be stolen and used by other people. A safer option for personal identification is the use of biometrics. Biometrics are any human behaviour or physical characteristics with the following properties: universality, uniqueness, permanence, and collectability [1]. Examples of biometrics are fingerprints, finger veins, irises, voices, and faces. Since biometrics are characteristics people are born with, it is hard to fool a personal identification system based on biometrics by other people. This security, together with the uniqueness is the reason why biotmetrics are safe identification methods used in security systems and besides that also used as proof in law enforcements.

It is important that the measurement of the biometrics is done carefully such that no contamination is present that might influence the detection. In fingerprint images one possible contamination is called the ghost fingerprint. When a ghost fingerprint is present a second non-wanted fingerprint overlaps part of the wanted fingerprint image. Figure 1 shows an example of a ghost fingerprint. Ghost fingerprint occurs in latent fingerprints, which are fingerprints left on a surface [2] but also on fingerprint scanners which are not properly cleaned.

Fig. 1: Example of an overlapping fingerprint.

Fingerprint recognition software can be split into three different groups; Correlation based matching, Minutiae-based matching, and Non-Minutiae feature-based matching [3]. All these methods use either the ridge shape, frequency characteristics, Minutiae locations which are small ridge characteristics such as ridge endings, bifurcations and cores [3], location or a combination of these characteristics. Since all of these groups are affected by the presence of a ghost in the fingerprint the recognition software will have more difficulty matching the fingerprints. It is therefore important to detect ghost fingerprints before the image is placed in a database or used by recognition algorithms. Previous research has been done to successfully separate ghost fingerprints to still be able to use the images in algorithms [4]. This research,

however, assumes that the masks that show which part of the image is overlapping and what not is marked manually. For an automated process and immediate detection of ghost fingerprints, it is important that the creation of these masks is also done automatically. Two researches have been found which proposes a method to mask ghost fingerprints. The first research looks to the intensity of the blurred ghost fingerprint and uses a threshold to separate the different intensities in different image parts [5]. Although this method seems to work, it relies heavily on the fact that there is an intensity difference between the overlapping and non-overlapping part of the image. Besides that, it assumes that there is no intensity change in the wanted fingerprint as well as in the overlapping fingerprint, which is not true in most cases. The second research proposes two methods to detect ghost fingerprints. Firstly, a method based on the Local Binary Pattern (LBP) is used, and second, a method based on the frequency estimation is used [6]. These methods are able to detect the location of a ghost fingerprint with an Equal Error Rate (EER) of 0.319 for the LBP and 0.261 for the frequency method. This means that there is still improvement left in the detection and masking of ghost fingerprints.

A problem, however, in testing these algorithms is the amount of data that is needed to verify the effectiveness of the algorithms, especially when the errors are small. For example, at least one million images are needed to be able to claim, with an accuracy of 95 %, that the FAR is between 0.006 and 0.014%. Meaning that one image in 7,000 until 16,000 images is falsely classified [7]. However, getting this amount of images is difficult due to the amount of time and money it takes. Additionaly, since fingerprints are personal information, privacy rules also get into play when gathering the images. There are already programs which are able to create synthetic fingerprints. Such as the SFinGe method which is able the generate realistic looking fingerprint from scratch [7]. However, no method has been found that is able to create realistic looking ghost fingerprints.
The goal of this paper is the propose a method to create realistic looking ghost fingerprints which can be used to test, optimise, and train algorithms to classify and mask ghost fingerprint images. The ghost fingerprints will be created by combining real fingerprint images, which has the advantage that noise and distortion are already present. An advantage of creating the ghost fingerprint is that the location of the overlap is known. This means that testing algorithms will be easier, but also that it can be used by learning algorithms to classify the images. Moreover the percentage of overlap, intensity, and the thickness of the lines in the ghost source, can be manually adjusted to create different groups separated on the influence it will have on the fingerprint recognition algorithms.

The first section of this paper will explain the steps and the theory behind the creation of the ghost fingerprints, after which the results are shown. The second section will briefly explain the LBP and the frequency estimation to compare the created images with the real images. The result of applying these methods will also be shown in that section. Lastly, a conclusion will be drawn and some recommendations will be given for future work.

II. Ghost fingerprint generation

This section will focus on the generation of the ghost fingerprint. First, in the method, the steps that are taken to reach the final images are explained. The proposed method will be tested and its results will be shown and compared to real images in the last part of this section.

A. Method

To create realistic looking ghost fingerprints it is essential to look at real ghost fingerprints. As can be seen in Figure 1 in the overlapping part of the fingerprint the black lines are interrupted by the lines of the ghost. The colour of the lines does not change. Meaning that the grey/black part of the ghost source image does not influence the grey/black part of the ghost fingerprint. Besides that it is important to note that the only place where to ghost can be seen is in the overlapping region. The other part of the ghost fingerprint is not shown. The steps the proposed method takes to recreate these observations are shown in figure 2 and are:

1) Source selection,
2) Geometric transformation,
3) Standardize size,
4) Masking,
5) Changing line thickness,
6) Threshold binarization,
7) Combining the images.

The generation of the ghost fingerprint is done in Mathworks' Matlab. The previous listed steps will now be explained in more detail.

1) Source selection: The source images that are used to create the ghost fingerprints are real fingerprints. A database is used provided by the Dutch police. The database contains 4,650 images separated into seven different classes. Class 0 contains 791 fingerprint images that do not contain ghosts. Some of these images did contain some noise and distortion and were left out, resulting in 699 images that can be used as source images. Furthermore, class 1 until 5 contains ghost fingerprints separated on the expected amount that the ghost affects automated fingerprint matching. The class 1 images are expected to affect the matching the most while class 5 is expected to affect the matching the least. These are in total 3,738 images. Class 6 are 121 fingerprint images that contain distortions other than ghosts. Both the base source as the ghost source are selected from the remaining class 0 images. The decision which image is used is made using a uniformly distributed random function. This means that every fingerprint in class 0 will has the same porbability to be used.

2) Geometric transformation: Before the mask of the fingerprints are created, the changes in geometry are made. After the study of the real ghost fingerprint images it was found that the base fingerprint does not have changes in rotation. It was

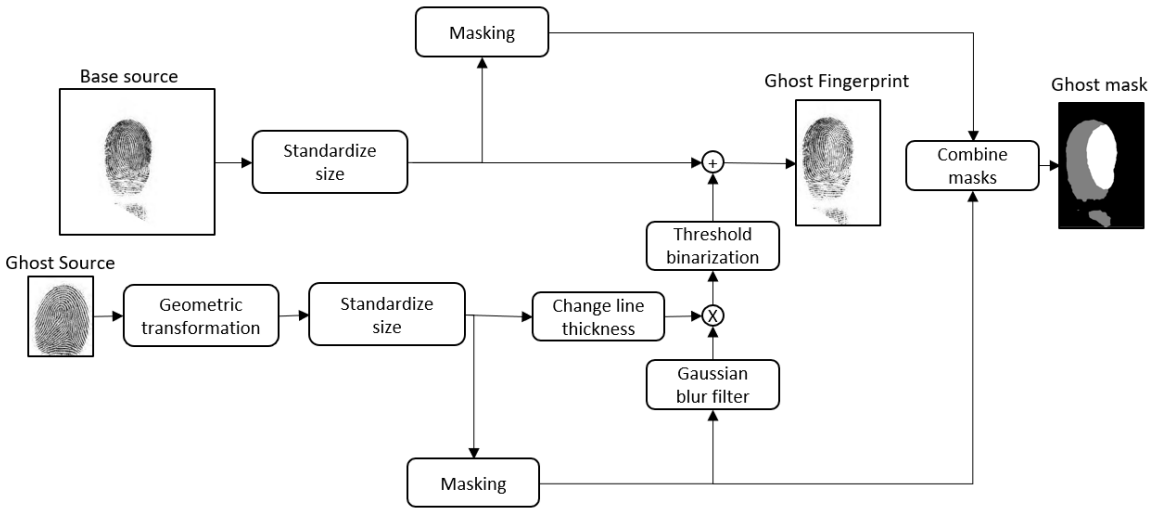


Fig. 2: Functional block diagram of the creation of artificial ghost fingeprints.

therefore decided to only apply the geometric transformation to the ghost source. The transformations that can be applied applied are:

- Scaling,
- Rotation,
- Location,
- Horizontal flip.

Before these transformations are applied the image is first extended on all sides to a 1,400 · 1,400 pixels image. The fingerprint will be placed the center of this image. The added pixels are white and thereby do not influence any of the other operations. This extension is done to avoid artifacts from the transformations after cropping the image.

In order to generate a database, these transformations must be randomised to get diversed images. This randomisation is done for each transformation with a specific range. The scaling is done with a factor from 0.75 until 1.25. The amount of the rotation is between -60 and 60 degrees. The change in location is between -250 and 250 pixels for the vertical axis and between -300 and 300 pixels for the horizontal axis. Finally, the horizontal flip can be performed. The probability of these random functions are uniformly distributed. These ranges have been chosen to get a large amount of variety in the images but still create realistic looking ghost fingerprints.

3) Standardising size: Since combining images is easier when the images are the same size both the size of the ghost source and the base source will be standardised to 600 · 400 pixels. A smaller image will reduce image processing time, however, information might get lost due to clipping. After some testing it was found that a size of 600 · 400 pixels results in small images without significant clipping.

4) Masking: A problem with the database images is that the background is white. This means that when the images are added the background of both images will overlap the black ridges of the fingerprints. This would mean that the final image will only contain the overlapping part. To make sure this does not happen the fingerprints should be separated from the background. Thus a mask must be created. This is done using the mean and variance based method. This method works well in fingerprints with light background [8] which is the case in the used dataset. This method works by splitting the image up in smaller non-overlapping blocks. The standard deviation is being calculated for these block and compared to a threshold. This threshold divides the blocks between background and fingerprint. In this research instead of the non-overlapping blocks a moving standard deviation is used with a block size of 9 · 9 as described in Equation (1) for image I(i,k) with moving average M(i,k) described in Equation (2).

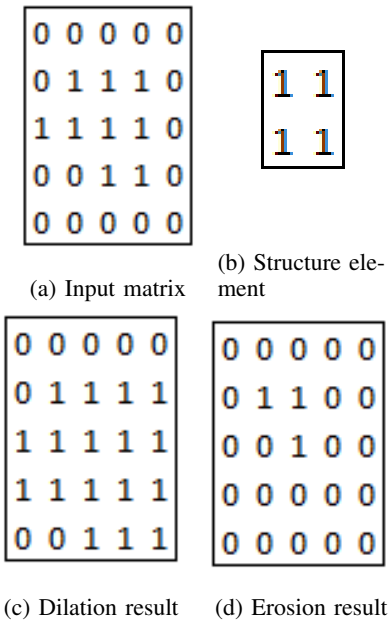
$$std(i,k) = \sqrt{\frac{1}{9^2}\sum_{j=-4}^{4}\sum_{l=-4}^{4}(I(i-j,k-l)-M(i,k))^2} \quad (1)$$

$$M(i,k) = \frac{1}{9^2}\sum_{j=-4}^{4}\sum_{l=-4}^{4}I(i-j,k-l) \quad (2)$$

The advantage of using a moving standard deviation instead of the non-overlapping blocks is that the created mask will be smoother. However, it does cost more processing time. Before the moving standard deviation can be performed the image first need to be extended with four white pixels on all sides to also be able to operate on the edges. The threshold value has been determined to be 0.2 after some observation of both lighter and darker fingerprints. This means that above a standard deviation of 0.2 the mask is set to 1, meaning there is a fingerprint. Below 0.2 the mask is set to zero, implying background. Since

this masking is not ideal and has trouble with light or dark regions, the mask must be closed and the holes must be filled. The last operation that must be performed is to make sure the mask is not bigger than the fingerprint. Therefore the mask must be made smaller. This is done by applying a Gaussian smoothing filter and applying a threshold to set the values back to zero or one. The standard deviation used for the filter was four and the threshold was set to 0.85.

5) Changing line thickness: To create more variation in the images, the line thickness of the ghost source can be adjusted. This operation is performed after the creation of the mask since, especially when the lines become too thick or thin, it affects the accuracy of the masking. Changing the line thickness is done using dilation and erosion, which is also used in the SfinGe method to simulate dry or wet skin [7]. The dilation operator is used to thicken objects. In a binary image this is done by replacing every pixel with value one with the structure element. The erosion operator is used to make object thinner. A pixel is being replaced with the lowest value in the range of the structure element [9]. An example of both the dilation and the erosion operation on a matrix can be seen in Figure 3. The reference pixel in the structure element is the top left pixel.



(a) Input matrix (b) Structure element

(c) Dilation result (d) Erosion result

Fig. 3: Simple example of applying erosion and dilation to a matrix.

To create variation, the erosion and dilation function are applied at random. The possibility that one of these functions is applied is 50%, and the probability of choosing between the erosion and dilation function is equal. The structure element that is used is a 2x2 matrix as shown in Figure 3b.

6) Mask blurring: Before the transformed ghost source is combined with its mask, the mask is first blurred. This will help to shape the mask more around the fingerprints when binarisation is applied. The blur filter is a Gaussian filter with a standard deviation of 16.

7) Threshold binarisation: The blurred mask and the transformed ghost source are multiplied together. The result is the transformed ghost source with a black background and blurred edges. As described before the grey-level of the ridges does not change in the overlapping and the non-overlapping part. This means that the black/grey ridges of the ghost source do not influence the ridges in the created ghost fingerprint. To simulate this, the ghost source can be binarised. This means that below a certain threshold the values will be set to zero and above that threshold the values will be set to 1. The fingerprints in the database have different grey-value. Choosing a threshold value that works on all the fingerprints is not possible. Therefore, the threshold will be determined for each fingerprint separately. The average value of the overlapping part of the ghost source is determined before changing the thickness of the ridges. This value is then added to a random value ranging between -0.2 and 0.2 to get more possible variation in the final images. The resulting value is used as threshold for that image. The created image is blurred with a Gaussian filter with standard deviation 1 and multiplied with a value of 0.8 to make the white parts more grey. This last step is done to remove hard edges making the image more real.

8) Combining the images: To create the final image the in the previous step created image should be added to the standardized sized base source. The last adjustment that must be done is to clip the pixel value to one. Thus all the pixels that have a value bigger than one will be set to one. This must be done to plot the final image. Furthermore, a ghost mask image can be made where it is easy to see which part of the fingerprint contains overlapping. This mask image is created by combining the previously generated masks. The masks are first multiplied and its result is added to the mask of the base. Then the result must be divided by two. An image is created in which black is the background, grey is the base fingerprint with no overlap, and the white part is where the overlap takes place.

B. Results and discussion

To see if it is possible to create real looking ghost fingerprints. Four different real fingerprints from the dataset have been recreated with the method described above. To be able to recreate the image as best as possible the parameters were manually chosen instead of using a random function. The used images were also hand picked from the dataset. The results are shown in Table I. The first and the second column are the fingerprint images that need the be combined after the standardisation of the size. The third column is the created ghost fingerprint, and the fourth column is the real ghost fingerprint that is recreated. The different rows are different attempts to recreate a ghost fingerprint. In Table II the parameters that were used for each row are shown. A few randomly generated ghost fingerprints are shown in the appendix.

TABLE I: Recreation of a real ghost fingerprint from fingerprint images

Base source	Ghost source	Created ghost fingerprint	Comparable real ghost fingerprint
			
			
			
			

TABLE II: Parameters used to create the images shown in Table I.

Type of Transformation	Row 1	Row2	Row3	Row4
Scaling factor	1.2	1.1	1	1
Rotation (degrees)	0	-15	25	20
Vertical movement (pixels)	100	-80	60	-90
Horizontal movement (pixels)	-120	100	200	-225
Image Flip	True	False	False	True
Erosion or Dilation	Erosion	Dilation	Erosion	None

Table I shows that different types of ghost fingerprints can be created. The high contrast ghost fingerprint as seen in the second row creates hard edges between the overlapping part and the non-overlapping part. These can be seen in both the artificial and the real ghost fingerprint. More subtle ghost fingerprints can also be replicated as shown in the third row. Overlapping locations at the top of the image can be made as shown in row four. The parameters that were used for these images are in the range for the random function. These images can thus be randomly generated by the described method.

A problem, however, in the system is the masking. When the source image is too light or dark the masking algorithms do

not mask properly. In case of light ghost source fingerprints, this can cause hard edges, for a darker source image this is less of a problem.
Another problem that might occur is that the erosion operator makes the light images even more lighter and the dilation operator makes dark images darker. This results might result in very light overlapping parts or parts where according to the mask overlap should take place but is not present. This is mostly solved by the threshold binarisation, however in some rare cases this problem might still occur.

III. Applying LBP and Frequency estimation algorithms

The artificial ghost fingerprint looks comparable to the real. However, when algorithms do not react the same way to the the images the artificial ghost fingerprints cannot be used in testing or training. In this section a database will be created according to the previous section which will be applied to two different ghost fingerprint classification methods. The method will also be applied to the real ghost fingerprint from the dataset to compare the results.

A. Method

The methods that will be used are the Local Binary Pattern and the frequency estimation. Both methods are explained in detail in the research by Holland, Oonk, Spaan & Zonneveld (2019) [6]. Since the real ghost fingerprints are classified in five classes depending on its influence on fingerprint matching algorithms. The created ghost fingerprints will also be classified in different classes. Three different classification characteristics have been used: percentage of overlap, average grey-level of the overlap, and the combination of the previous. These three different classification characteristics will all be applied to the classification methods and compared. The comparison will be made using the Receiver Operating Curve. In a ROC curve the true positive rate (i.e. percentage ghost fingerprint of the mentioned class classified as ghost fingerprint) is plotted against the false positive rate (i.e. percentage non-ghost fingerprint classified as ghost fingerprints) for different thresholds. The ROC curve will thus be made using one of the classes for the true positive rate and class 0 of the database for the false positive rate. A higher area under the ROC curve means a better classification algorithm and a diagonal line from point (0,0) until (1,1) shows a classifier with random performance [10]. The Equal Error Rate (EER) can be read off from the ROC curve. It corresponds to the point of the ROC curve where the false negative rate (1- true positive rate) is equal to the false positive rate. A lower EER value means a better classifier performance. In this case the EER means the probability that a non-ghost fingerprint is predicted as ghost and a ghost fingerprint is predicted as non-ghost fingerprint.

1) LBP: The local binary pattern is a method widely used as texture descriptor [11]. A pixel is given a value based on its own and its surrounding values. This is done by comparing each pixel with its eight neighbours. When the neighbour pixel has a higher value than the centre pixel a one will be placed in a $3 \cdot 3$ matrix on the same location. By concatenating all these ones or zeros in clockwise direction starting at the top left a binary number can be read off. This binary number is converted to a decimal number and placed in a new image on the same location as the centre pixel [12]. This method is applied on all the pixels in an image and the result is a new image with the LBP value of each pixel on the corresponding place. An example of how LBP work is shown in Figure 4. Before the LBP is applied, the ghost fingerprint image is first

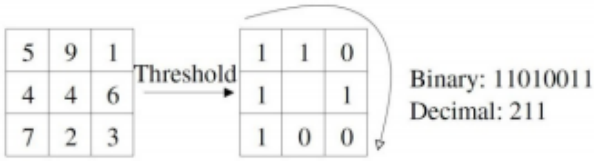
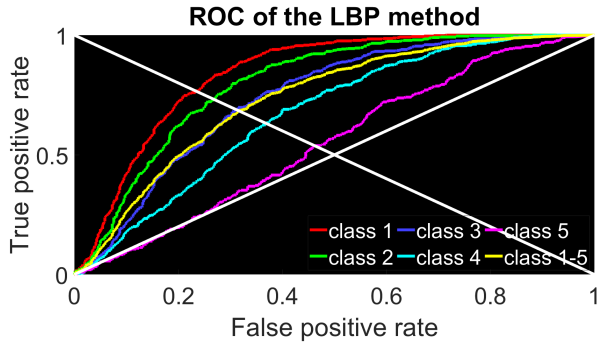


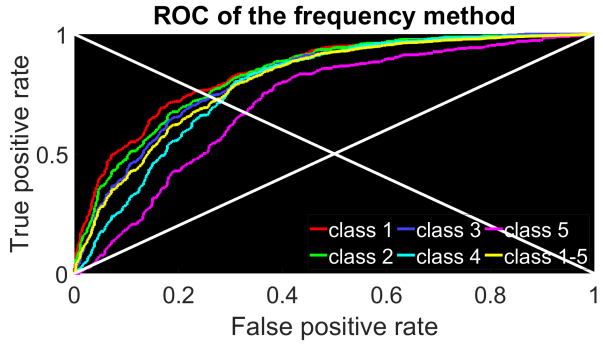
Fig. 4: An example of a local binary pattern on a pixel [12].

binarised and the holes are closed. The image is then being placed in a $750 \cdot 800$ pixels image such that the size is the same for every fingerprint. Since the ridges of the fingerprints are three pixels wide, applying the LBP as described above will not result in useful information. Therefore, the LBP will be performed on a $9 \cdot 9$ grid. The value of the centre pixel and the eight neighbours are averaged and are compared to the average of the $3 \cdot 3$ grid surrounding the centre $3 \cdot 3$ pixels. When this operation is executed for every pixel, the created image with LBP values is separated in blocks of $25 \cdot 40$ pixels. For each block the values are placed in a histogram. In blocks where overlap takes place there are more bins with smaller (between 1 and 20) heights. A score is being assigned to these blocks by counting the number of bins with heights between 1 and 20. A score higher than 60 is found to correspond to a ghost region. A score between 10 and 60 is a fingerprint and a score lower than 10 corresponds to a background region.
Using these assigned regions in the blocks, a percentage of predicted ghost fingerprint with respect to the complete fingerprint is calculated. This percentage is compared to the threshold used to create the ROC curve.

2) Frequency estimation: The second method is based on the frequency analysis of the image. The frequency of the image can be seen as the distance between the ridges [13]. The frequencies of the image can be found by applying the Fast Fourier Transform (FFT). This is done after the image is enhanced as described in the LBP section, after normalising the image and after splitting the image into blocks of $25 \cdot 40$ pixels. It was found that when a ghost was present in a block the maximum frequency was higher. The maximum frequency of a fingerprint, however, is different for each fingerprint. Therefore the maximum value of a block is compared to the average maximum frequency value of all the blocks. When the maximum frequency is higher than the average maximum frequency the block is assigned a one and when below a zero. A mask of the complete fingerprint is created to separate the ghost blocks with the background blocks. As is done for

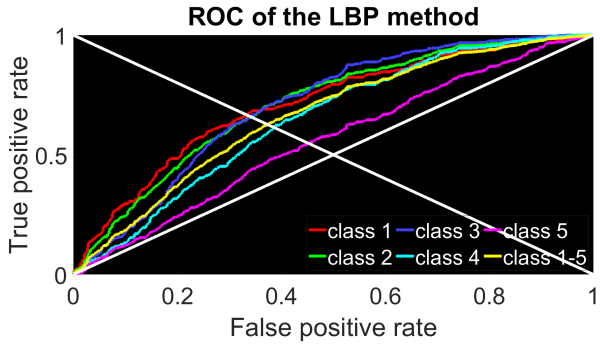


(a) ROC curve of the LBP method for the different classes and for a combination of all the classes.

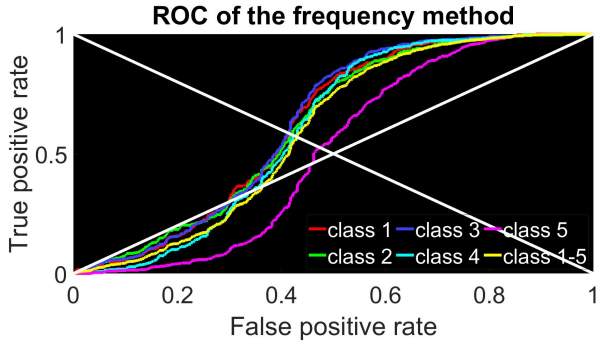


(b) ROC curve of the frequency method for the different classes and for a combination of all the classes.

Fig. 5: ROC curves of the LBP and frequency method on the real fingerprint data.



(a) ROC curve of the LBP method for the different classes and for a combination of all the classes.



(b) ROC curve of the frequency method for the different classes and for a combination of all the classes.

Fig. 6: ROC curves of the LBP and frequency method on the created ghost fingerprint with the percentage ghost as class differentiator. Class 1 0-20%, class 2 20-40%, class 3 40-60%, class 4 60-80%, class 5 80-100%.

the LBP method a percentage of overlap with respect to the complete fingerprint is calculated. This percentage is compared to the threshold to create the ROC curve.

B. Results and discussion

To compare the results of the created ghost fingerprints with the real ghost fingerprints both dataset are applied to the described methods. The real data contains five different classes. Class 1 is expected to have the most influence on fingerprint recognition software and class 5 the least. The ROC curve of the LBP and the frequency method on the real fingerprint images can be found in figure 5. Similar figures can be found in the research by Holland, Oonk, Spaan & Zonneveld (2019) [6] since the same dataset is used. In figure 6 the ROC curve of the LBP and the frequency method on the artificial ghost fingerprint can be seen. The different classes correspond to the different amount of overlap. Class 1 has a ghost that covers 80 - 100% of the total fingerprint and class 5 to 1 - 20% of the total fingerprint. The classes in between each have their own amount of overlap. The ROC curves of the dataset seperated based on the average grey-level and the a combination of the grey-level and the percentage is shown in the appendix. The dataset based on average grey-level is seperated in four classes with an average grey level between 0 and 0.25 for class 4, 0.25 and 0.42 for class 3, 0.42 and 0.58 for class 2, and 0.58 and 0.75 for class 1. The classes based on the combination of the intensity and percentage of overlap are separated on the average value of these parameters. Class 4 0-0.2, class 3 0.2-0.4, class 2 0.4-0.6 and class 1 0.6-0.8. The ROC curves are created with 500 images for each class except for class 5 of the real data and the percentage based separation containing 301 images and class 1 of the grey-level and combination separation contains 180 images. The false positive data is created using the 699 images from class 0 of the dataset. In Table III the different EER values can be found for the different classification parameters and data.

In Figure 5 and 6 it can be seen that the methods can detect ghost fingerprints. The accuracy of the detection of ghost fingerprints from the artificial ghost fingerprints is lower than from the real ghost fingerprints as can be seen in Table III. Besides that the curves of the different classes are closer to each other in the artificial fingerprints. This means that the percentage overlap has influence on the detection of the ghost fingerprints, but it is not the only classifier used for the classes in the real ghost fingerprint. A possible explanation for the lower accuracy is that since percentage is probably not the only parameter to classify the ghost fingerprint in different classes, the other parameters are still randomly divided between the classes. This means that the classes are more averaged in

TABLE III: EER of the classification between ghost fingerprint or non-ghost fingerprint of for each class and method separately

Data	Method	Class 1	Class 2	Class 3	Class 4	Class 5	Class 1-5
Real	LBP	0.23	0.27	0.33	0.37	0.46	0.31
	Freq	0.24	0.26	0.26	0.28	0.32	0.28
artificial (percentage)	LBP	0.33	0.34	0.34	0.39	0.45	0.37
	Freq	0.41	0.42	0.41	0.43	0.47	0.43
artificial (gray-level)	LBP	0.37	0.36	0.35	0.42	-	0.37
	Freq	0.31	0.36	0.4	0.44	-	0.40
artificial (gray-level + percentage)	LBP	0.49	0.36	0.37	0.42	-	0.39
	freq	0.39	0.37	0.41	0.44		0.41

terms of difficulty to detect compared to the real data and classification method. This however does not explain the lower EER for the classes combined. This might be explained by the additional noise and distraction that is added by combining two different fingerprints. Both fingerprints have their own noise and distortion, when combined these will be added to each other possibly creating more noise than when the fingerprint come from one measurement as is the case with the real ghost fingerprints. Besides that the methods are optimized using the real ghost fingerprint data. This, however, should not be an issue since the artificial ghost fingerprints are supposed to look like the mentioned real ghost fingerprints.

It can be seen that the ROC curve of the frequency methods on the artificial data shows a s-curve while the ROC curve on the original data shows a regular curve. The s-curve can be explained by the difference in standard deviation and the mean of the predicted percentage of overlap of the non-ghost fingerprint data and the other classes. The standard deviation of the non-ghost data is higher and while the mean is still lower than the ghost fingerprint data at some point there are more non-ghost fingerprint above the threshold value of the ROC creating a s-shaped curve. In the real data, the difference in the mean between the class 0 and the other classes is higher compared to that difference in the artificial data. This results is a different shape of the ROC curve. The difference in mean and standard deviation have influence on the accuracy of detection therefore the shape of the ROC curve is related to the EER values which is on average 0.15 higher for the artificial data.

IV. Conclusion and future work

In this work, a method is proposed to generate ghost fingerprints. This method is able to generate realistic looking ghost fingerprints by combining real fingerprint images. Due to the different amount of geometrical transformations and the possibility to change the line thickness, a large variety of different ghost fingerprints can be created. Ranging from a small amount of overlap to a complete overlap and from a barely noticeable ghost till a high contrast difference between ghost and non-ghost. The detection of the artificial ghost fingerprints has a lower accuracy than the detection of the real ghost fingerprints. The EER of the real ghost fingerprints is 0.31 for the LBP method and 0.28 for the frequency method for the combination of all classes compared to an EER between 0.37 and 0.39 for the LBP method and between 0.40 and 0.43. The shape of the ROC curves of the LBP method of the real and created data looks similar. The shape of the frequency curves looks different, however, since this is a result of the detection difficulty it can still be concluded that the detection methods works on the created ghost fingerprints and that the created fingerprints can thus be used to train algorithms to detect ghost fingerprints. When this method is used, it should be taken into account that the created ghost fingerprints are more difficult to detect. Attempts have been made to create different ghost fingerprint classes based on the difficulty of detection as can also be seen in the real ghost fingerprints data. It can be seen that both the percentage of overlap as well as the intensity of the ghost influence the accuracy of detection. However, the proposed separation method does not create as much accuracy difference compared to the classes of the real ghost fingerprint, meaning that there is still some other parameter that influences the accuracy or that the used parameters should be combined differently.

In future work, the previously mentioned classes can be improved such that algorithms can also be trained to distinguish the different classes. The percentage of overlap and intensity of the ghost fingerprints influence the accuracy of detection meaning that these must be better combines or that another parameter yet unknown also has influence on the accuracy. The creation of the ghost fingerprint can also be improved. The problem with the masking of very light or dark parts can be fixed by possible using another method of masking or maybe by choosing better threshold values. The problem that images become too light or too dark due to the erosion or dilation operator can be solved by creating a decision system that detects the grey-level of the overlapping part before erosion or dilation is applied and can decide not to perform one of these operations when the image is too light or too dark. Additionally, the program currently creates a random ghost fingerprint image. To gain a bit more control over the created ghost fingerprint dataset, the method can be rebuild such that the user can ask for ghost fingerprints of a particular class, amount of overlap or other parameters. The program then creates a fingerprint corresponding to the demands of the user. Lastly, a program can be created that is able to create ghost fingerprints without the use of fingerprints images. This can be done by combining this method with the previous mentioned sFinGe method, which is able to generate fingerprints without the use of source material. It might be needed to add more noise and distortion to create realistic looking ghost fingerprints.

REFERENCES

[1] Anil Jain. *Biometrics : personal identification in networked society*. Springer, New York, 2006.

[2] M. Edwint O'Neill. The development of latent fingerprints on paper. *Journal of Criminal Law and Criminology*, 1937.

[3] Davide Maltoni. *Handbook of fingerprint recognition*. Springer, New York, 2003.

[4] F. Chen, J. Feng, A. K. Jain, J. Zhou, and J. Zhang. Separating overlapped fingerprints. *IEEE Transactions on Information Forensics and Security*, 6(2):346–359, June 2011.

[5] Tejas K, Swathi C, Aravind Kumar D, and Rajesh Muthu. Automated region masking of latent overlapped fingerprints. *CoRR*, abs/1710.09267, 2017.

[6] Holland M van, Oonk P, Spaan C, and Zonneveld T van. Detection of ghost fingerprints by local binary patterns and frequency methods. 2019.

[7] R. Cappelli, A. Erol, D. Maio, and D. Maltoni. Synthetic fingerprint-image generation. In *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, volume 3, pages 471–474 vol.3, Sep. 2000.

[8] M Akram, Anam Tariq, Shahida Jabeen, and Shoab Khan. Fingerprint image segmentation based on boundary values. pages 134–138, 01 2008.

[9] *Morphology*, pages 501–514. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[10] T. Fawcett. An introduction to roc analysis. *Pattern Recognition Letters*, 27, 112 2005.

[11] Ahonen T., Hadid A., and Pietikainen M. Face description with local binary patterns: Application to face recognition. 2006.

[12] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen. Local binary patterns and its application to facial image analysis: A survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 41(6):765–781, Nov 2011.

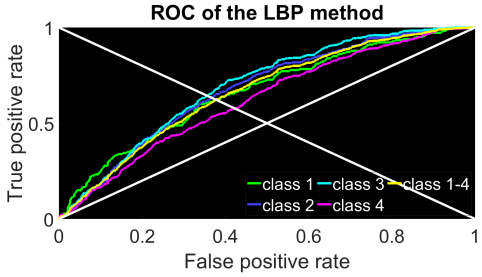
[13] Victor-Valeriu Patriciu and Stelian Spinu. Fingerprint ridge frequency estimation in the fourier domain. *Advances in Electrical and Computer Engineering*, 14:95–98, 11 2014.

APPENDIX

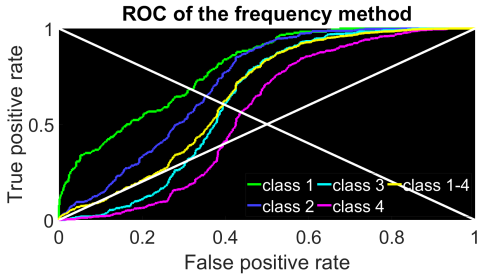
A. Generatad ghost fingerprints

Fig. 7: Generated ghost fingerprint image

B. ROC plots

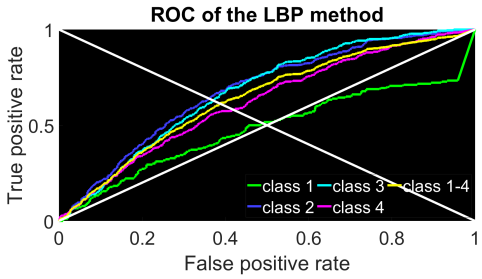


(a) ROC curve of the LBP method for the different classes and for a combination of all the classes

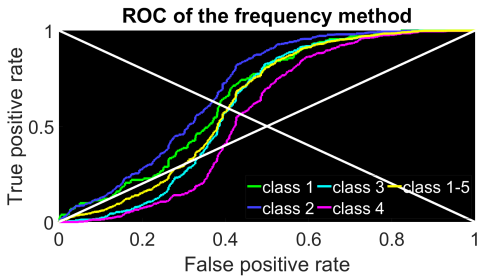


(b) ROC curve of the frequency method for the different classes and for a combination of all the classes

Fig. 8: ROC curves of the LBP and frequency method on the created ghost fingerprint with the average grey-level of the overlapping part of the ghost as class differentiator. Class 4 0-0.25, class 3 0.25-0.42, class 2 0.42-0.58, class 1 0.58-0.75



(a) ROC curve of the LBP method for the different classes and for a combination of all the classes



(b) ROC curve of the frequency method for the different classes and for a combination of all the classes

Fig. 9: ROC curves of the LBP and frequency method on the created ghost fingerprint with the percentage ghost and average grey level as class diferentiator. The combination is the average of the percentage and the averege gray level. Class 4 0-20%, class 3 20-40%, class 2 40-60%, class 1 60-80%