INTEGRATING WAVELET ENTROPY AND BINARIZED STATISTICAL IMAGE FEATURES TO IMPROVE FINGERPRINT INTEROPERABILITY

by

Zachary Chapman

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Computer Science

Charlotte

2016

Approved by:

_____

Dr. Bojan Cukic

_____

Dr. Richard Souvenir

_____

Dr. Mohsen Dorodchi

©2016
Zachary Chapman
ALL RIGHTS RESERVED

ABSTRACT

ZACHARY CHAPMAN. Integrating wavelet entropy and binarized statistical image features to improve fingerprint interoperability. (Under the direction of DR. BOJAN CUKIC)

Biometric systems are widely deployed in governmental, military and civilian applications. There are a multitude of sensors and matching algorithms available from different vendors. This creates a competitive market for these products, which is good for the consumers but emphasizes the importance of interoperability. Interoperability in fingerprint recognition is the ability of a system to work with multiple fingerprint scanning devices. Assuming that the same sensor or vendor will always be available during the lifetime of an automatic fingerprint recognition system is unrealistic. Typical variations induced by fingerprint sensor diversity include image resolution, scanning area, gray levels, etc. Such variations can impact (i) the quality of the extracted features and (ii) cross-device matching performance. In order to enhance interoperability, previous research has proposed a variety of fingerprint feature representations as well as a classification scheme to improve match rates across devices; however, implemented systems are not good enough to be operative. In this work, we propose a learning-based compensation scheme based on features derived from the discrete wavelet transform (DWT) and binarized statistical image features (BSIF) of captured fingerprint images. In particular we are interested in DWT for its capability to preserve spatial information of an image when performing frequency analysis while BSIF has shown to be effective in texture recognition tasks as a local descriptor. Experiments are carried out on a data set consisting of fingerprints obtained from 494 users acquired using four different optical devices. Results show reduced error rates compared to the baseline as well as improved performance compared to previous research.

ACKNOWLEDGEMENTS

Thank you to my committee chair, Dr. Bojan Cukic, who has served as a mentor to me not only as my thesis advisor, but also in my role as his teaching assistant throughout my Master's program. He has motivated me to complete this work, something I never envisioned being a part of coming into the program.

I would like to also acknowledge Dr. Richard Souvenir and Dr. Min Shin whose classes in Machine Learning and Digital Image Processing pushed me academically and inspired me to enroll in this project. Additionally, I would like to acknowledge Dr. Mohsen Dorodchi who encouraged me to take on a thesis project to finish my program.

I am incredibly grateful toward Dr. Emanuela Marasco who taught me an incredible amount about biometrics and fingerprints in such a short time span. Her guidance was instrumental to completing this work.

DEDICATION

I would like to dedicate this thesis to my wife, Norma Chapman, who has been fully supportive of my efforts to return to higher education. Without her enduring the burdens she has, I do not believe I would have been able to perform to my fullest potential in this Master's program.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Motivation

Methods of authentication based on fingerprint recognition are widely used across a variety of governmental, military and commercial applications. Over the life of any such system, it is possible that the fingerprint acquisition device will change from the one originally implemented. With many vendors and different products on the market, the need to ensure interoperability in the system is important. Interoperability in fingerprint recognition is the ability of a system to integrate different acquisition devices without a significant degradation in matching performance.

Even among devices that make use of the same sensing technology (e.g., optical, capacitive), variations are introduced that create differences in raw data for the same user. Figure 1.1 shows the variations between fingerprint images obtained using different devices for a single individual. The first four devices (D0-D3) use optical sensors for fingerprint acquisition while the fifth (D4) is the scan of an ink-based ten print card.

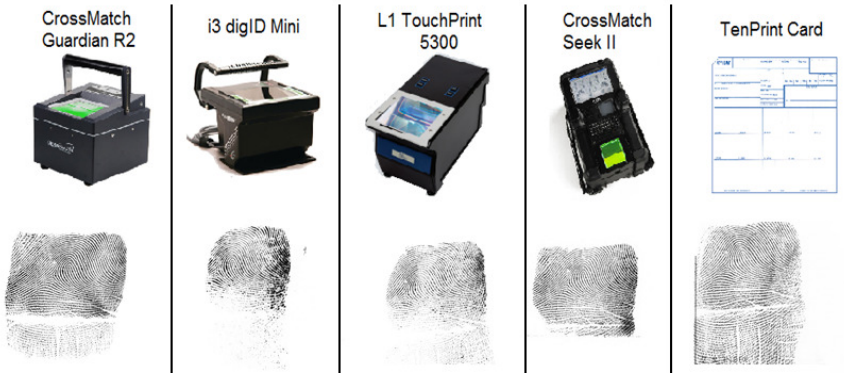
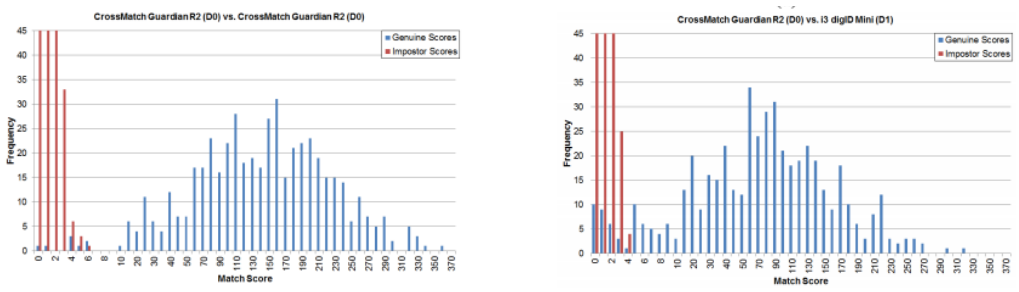


Figure 1.1: The same fingerprint captured across five different devices (four optical, one ink-based), from [1].



(a) (b)

Figure 1.2: Performance of traditional match score in both the intra and inter-device scenarios, from [1]: (a) Matching performance where gallery and probe images are obtained with the same device; (b) Matching performance where gallery and probe images are obtained with different devices.

Fingerprint matching requires two image samples. The first is a gallery image taken when a user enrolls into the system. The second is a probe image taken when a user attempts to access the system. An automated algorithm produces a score for the two fingerprints and compares the score to a set threshold that determines whether the fingerprints are a match or not. When the device used to acquire gallery and probe images is the same, matching algorithms perform with high accuracy. However, in the cases where gallery and probe devices are different, we see a greater overlap in genuine and impostor scores, leading to more questionable matches. Figure 1.2 shows the distribution of genuine and impostor scores in these two scenarios. In this study, we propose a model that mitigates errors as a result of inter-device matching.

1.2 Contribution

The main contribution in this work is the exploration of new features in order to improve interoperability in systems that make use of multiple fingerprint scanning devices. In particular we investigate features based on the discrete wavelet transform for its capability in maintaining spatial information while performing frequency analysis. We also evaluate the use of binarized statistical image features, a robust local image descriptor that has shown success in texture and face recognition tasks, on captured fingerprints. To our knowledge, neither of these image processing methods have been

leveraged in the area of fingerprint matching. These new features are incorporated with previously proven measures of quality and match scores into a proposed scheme that is shown to be more effective in improving fingerprint matching with images acquired through diverse sensors than using match score alone.

1.3 Organization

The rest of this work is organized as follows. Chapter 2 presents a review of previous research in fingerprint interoperability. Chapter 3 contains an overview of the architecture and features that comprise our approach in improving interoperability. Chapter 4 describes our method of experimentation and the results. Chapter 5 serves as a conclusion of this work and considers future direction in improving fingerprint system interoperability.

CHAPTER 2: RELATED WORK

2.1 Fingerprint Interoperability

In 2004, Jain and Ross explored the subject of fingerprint interoperability by comparing the match capabilities of optical sensors and capacitive solid-state sensors in both the intra-device and inter-device scenarios [2]. Both devices capture fingerprint images at a resolution of 500 dpi, but feature different scanning areas. Cross-sensor performance reported an Equal Error Rate (EER) of 23.13%, a significant decrease in performance from the intra-sensor results of 6.14% for the optical device and 10.39% for the capacitive one. In 2006, Alonso-Fernandez *et al.* noted that minutiae-based matchers outperform ridge-based matchers on the same data sets, but that both experience a large decrease in performance when matching between fingerprints taken with different sensing technologies [3]. Human interaction with fingerprint sensors has also been noted to affect matching performance, especially in systems that make use of different sensing devices [4] [5].

In 2006, Ross and Nadgir developed a non-linear calibration scheme based on Thin-Plate Splines to compensate for sensor diversity [6] . Relative distortions are modelled using manually selected control points. Although not fully automated, this approach showed a significant improvement in inter-sensor matching performance. In 2010, Poh *et al.* modelled a Bayesian Belief Network (BBN) in order to mitigate the effect of device acquisition mismatch when comparison is being performed with no a priori knowledge about the device a biometric sample is taken from [7, 8]. In their experiments, acquisition device is inferred from quality measures extracted from query images. Furthermore, the proposed BBN modelled how quality measures effect performance in determining thresholds for match score classifications.

In 2013 Lugini *et al.* performed a large scale statistical analysis of how match score changes across different optical devices [9]. Kendall's rank correlation test results point to a significant difference between sensor pairs and that change between devices is not symmetric. Most recently, Marasco *et al.* proposed an architecture for fusing fingerprint match scores with features extracted specifically to improve interoperability and using a machine learning algorithm for genuine and impostor classification [1].

CHAPTER 3: THE PROPOSED APPROACH

The goal of the system we propose is to determine if two fingerprints belong to the same identity by exploiting features that compensate for variations in the images related to acquisition device. Figure 3.1 illustrates the architecture we propose. Matching is performed on pairs of fingerprints, one of which represents the gallery and another which represents the probe. Preprocessing is performed on images that enhances ridge regions and normalizes images so that these regions have zero mean and unit standard deviation. Features are extracted from preprocessed images that make use of the spatial and frequency information provided through discrete wavelet transforms and binarized statistical image features and combined with traditional quality and match features obtained from raw images. Our combined feature set is used as input to train a binary classifier through supervised learning to generate a classification between a pair of fingerprints as either a genuine match or an impostor.

3.1 Image Pre Processing

All fingerprint images go through a series of preprocessing steps before feature extraction. First, images are normalized to have a mean of zero and unit standard deviation. Ridge regions are then detected by breaking each image into 16x16 blocks and comparing the standard deviation of the each block to a threshold value of 0.1. Blocks above the threshold are considered to be part of the fingerprint. Images are cropped around the detected ridge region and normalized once again to have a mean of zero and unit standard deviation. Figure 3.2 shows an example of a fingerprint image before and after preprocessing. Code for ridge detection and normalization was obtained from [10].

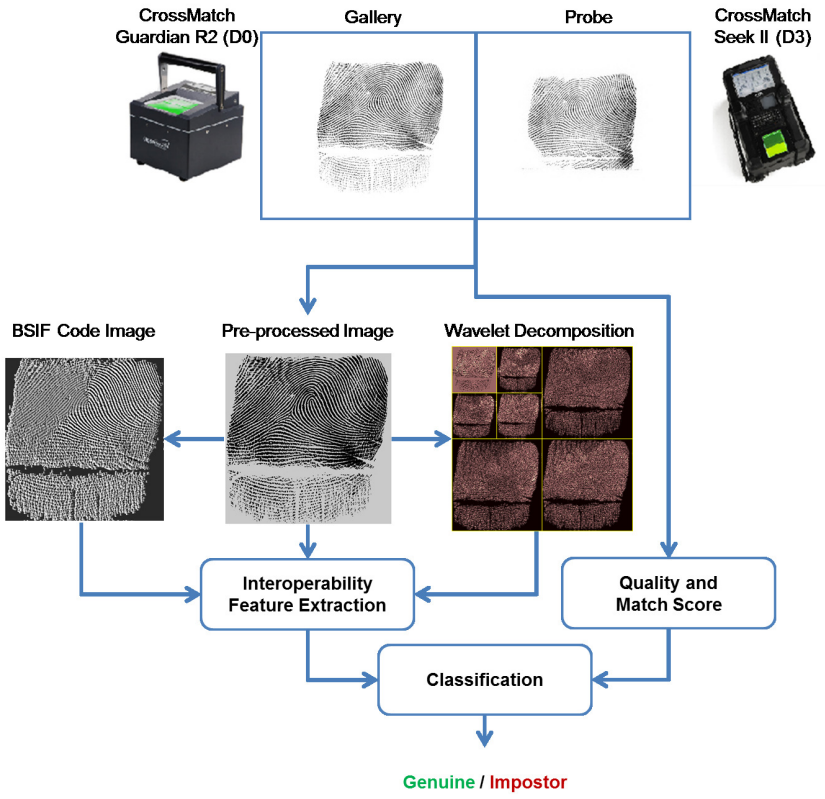


Figure 3.1: Proposed architecture in which interoperability features are extracted from pre-processed images and fused with quality features and match score from raw images. A classifier is trained with the resulting feature set to distinguish between genuine matches and impostors.

3.2 Feature Extraction

Wavelet Entropy. The discrete wavelet transform (DWT) in signal processing is used in many applications including compression, object recognition and numerical analysis [11]. Wavelet Transforms are of interest in this study for several reasons. First, its good time-frequency location abilities allows for analysis of non-stationary signals (including natural images) without losing spatial information related to discontinuities and edges. Secondly, it allows for representation of resolution at different scales, offering a hierarchical view of information. Lastly, is is easily realizable using a filter bank based on a mother wavelet and scaling function.

Wavelet decomposition via DWT for images is performed by passing an input image

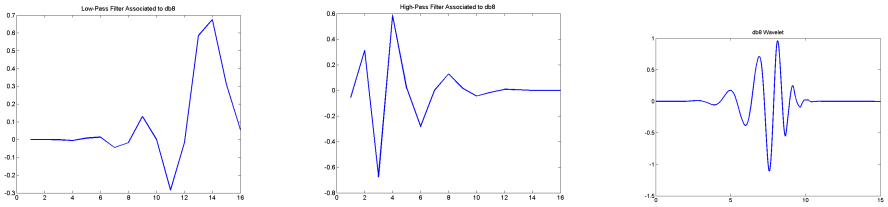
(a) (b)

Figure 3.2: Image Preprocessing: (a) Raw image obtained from D0; (b) Image after preprocessing.

through a low-pass and a high-pass filter and down-sampling, first along rows and then columns. This produces an approximation of the original image as well as horizontal, vertical and diagonal signal details. Decomposition can be performed at multiple resolutions by passing the approximation image at each level through another set of filters to produce the next level of approximation and detail. Approximations and detail signals are coefficients that represent how closely the wavelet is correlated with each specific section of the image.

Entropy measures the uncertainty associated with a random variable [12]. The random variables we are interested in represent the average amount of information generated by the distributions of the wavelet coefficients related to the approximation and details at different decomposition levels. Wavelet entropy measures represent the complexity of the signal in both the time and frequency domains [13]. In this study, we investigate the Shannon Entropy (SE) as well as the Log Energy Entropy (LEE) of the wavelet coefficients C produced by DWT at 6 levels of decomposition using the Debauchies db8 mother wavelet. Figure 3.3 illustrates the db8 mother wavelet along with the associated low- and high-pass filters. Figure 3.4 shows the first level of decomposition of a fingerprint image.

Entropy calculations are performed as follows, where i and j index C, the wavelet



(a) (b) (c)

Figure 3.3: Debauchies db8 analysis filters: (a) low-pass filter and (b) high-pass filter create the transformation performed in this study. The filters are designed for 16 coefficients. X axis indicates the coefficient number, while the Y axis indicates the corresponding value for that specific coefficient number. These filters are associated to the Debauchies wavelet function shown in Figure (c), the mother wavelet selected in this work.

coefficient matrix:

Shannon Entropy of C

$$SE(C) = -\sum_{i,j} C_{i,j} log_2(C_{i,j} + \epsilon) \tag{3.1}$$

Log Energy Entropy of C

$$LEE(C) = \sum_{i,j} (log_2(C_{i,j})) \tag{3.2}$$

In our final feature set, entropy calculations between fingerprint pairs for each coefficient matrix at each level are represented as unsigned differences. Figure 3.5 shows the SE value difference distributions across genuine and impostor pairs acquired from the level one decomposition and Figure 3.6 shows the LEE value difference distributions. Using entropy measures beyond the first level of resolution did not show any improvement in the separation of genuine and impostor pairs. By focusing on the first level of the DWT, the proposed system requires less processing time for feature calculation.

Binarized Statistical Image Features. Local image descriptors have become a standard tool for providing image features in computer vision tasks such as object and

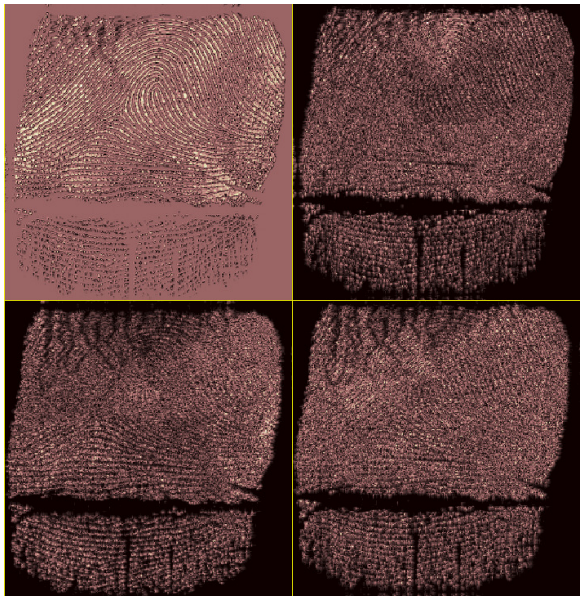


Figure 3.4: Single level decomposition of a fingerprint image using a db8 mother wavelet. In the top left is an approximation of the original image at a lower resolution while the remaining images represent signals in the horizontal, vertical and diagonal directions.

texture recognition. Methods of local description such as local binary pattern (LBP) [14] and local phase quantization (LPQ) [15] calculate a binary code to represent each pixel in an image. This code is a description of a pixel's neighborhood obtained by convolution with a set of manually predefined linear filters [1].

In 2012, Kannla and Rahtu introduced binarized statistical image features (BSIF) as a method of local image description which uses linear filters learned from a training set of natural images to produce a binary code for each pixel [16]. Contrasted to the manually predefined filters used in LBP and LPQ, the filters in BSIF are learned using Independent Component Analysis (ICA) in order to maximize statistical independence of the the filter responses. Experiments have shown that BSIF performed generally better than the previous methods at facial recognition and texture matching tasks [16], as well as in detecting spoofed fingerprints [17]. While not designed to counteract the effects of blur and rotation, BSIF proved to be comparably ro-

[1] http://www.ee.oulu.fi/ jkannala/bsif/bsif.html

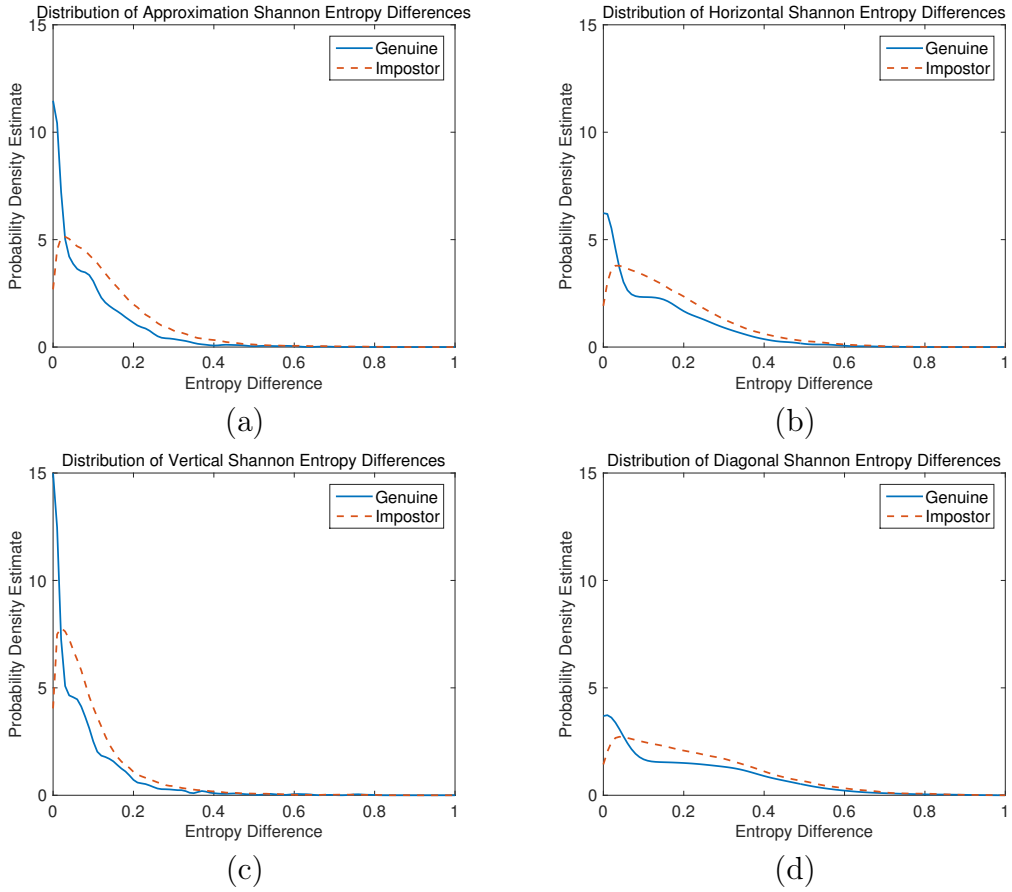


Figure 3.5: Shannon Entropy value differences acquired from the level one wavelet decomposition coefficients: (a) SE from the approximation image; (b) SE from the horizontal detail coefficients; (c) SE from the vertical detail coefficients; (d) SE from the diagonal detail coefficients.

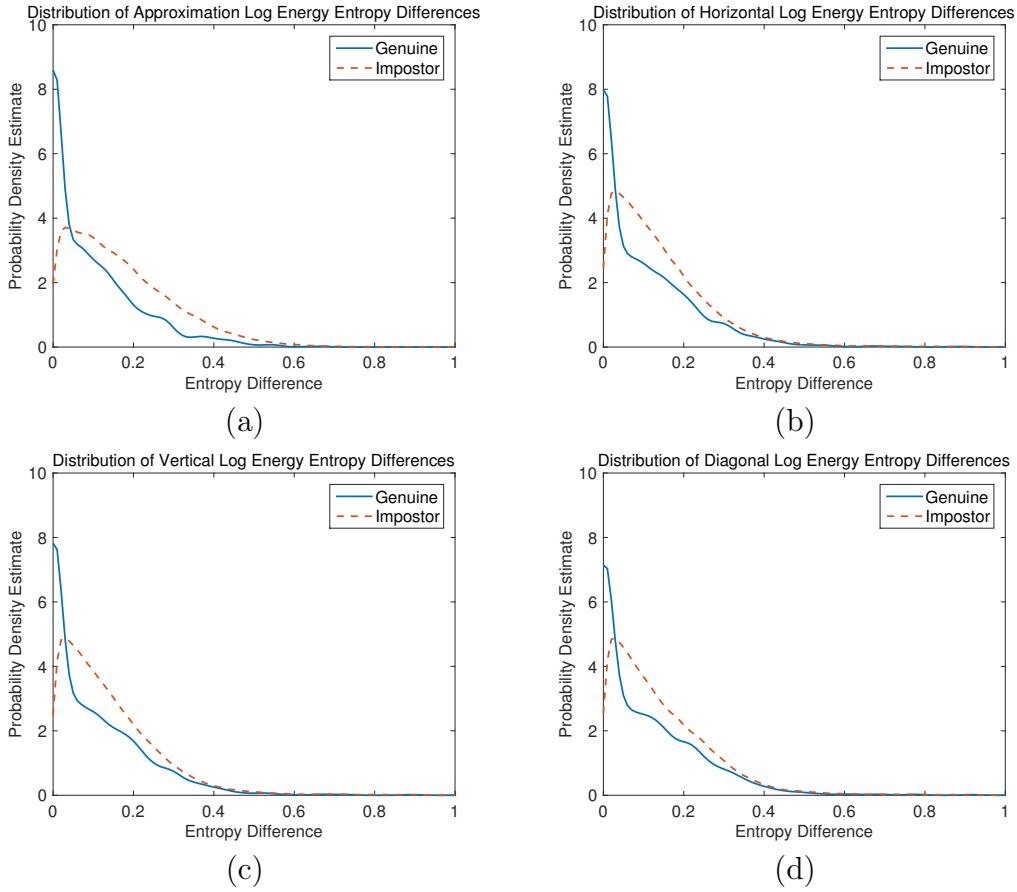


Figure 3.6: Log Energy Entropy value differences acquired from the level one wavelet decomposition coefficients: (a) LEE from the approximation image; (b) LEE from the horizontal detail coefficients; (c) LEE from the vertical detail coefficients; (d) LEE from the diagonal detail coefficients.

Figure 3.7: Result of convolution with each filter in the pre-learned 5 bit 11x11 filter set.

bust in both categories to previous methods specifically developed to mitigate such variations.

The number of bits used to represent each pixel is determined by the number of filters in the chosen set. The value of each bit, b_i, is calculated as a binary response of its associated filter, W_i, to the surrounding image region, X, at a threshold of zero.

$$s_i = \sum_{u,v} W_i\left(u,v\right) X\left(u,v\right) \tag{3.3}$$

The binarized feature $b_i = 1$ when $s_i > 0$, and $b_i = 0$ otherwise. Figure 3.7 shows the result of each filter in the provided 5 bit, 11x11 set on a single fingerprint image.

Resulting coded images are represented as a normalized histogram with a number of bins equal to the number of possible values encoded by the number of bits at each pixel. This can result in large feature vectors even when using as little as 5 bits, the smallest pre-learned filter sets provided. For this reason we investigate more compact representation of the histograms. We consider both the differences between statistical summaries of the histograms for each image as well as the Euclidean difference between histograms as methods of comparison. Summary measures used include minimum and maximum frequency values and the standard deviation. Figure 3.8 shows the distribution of these values across genuine and impostor pairs.

Features obtained from several filter sets (ranging from 5 bit, 3x3 to 9 bit, 15x15) were evaluated with best results being obtained from the 5 bit, 11x11 filter set.

Image Quality. Measures of quality represent the suitability of a biometric sample

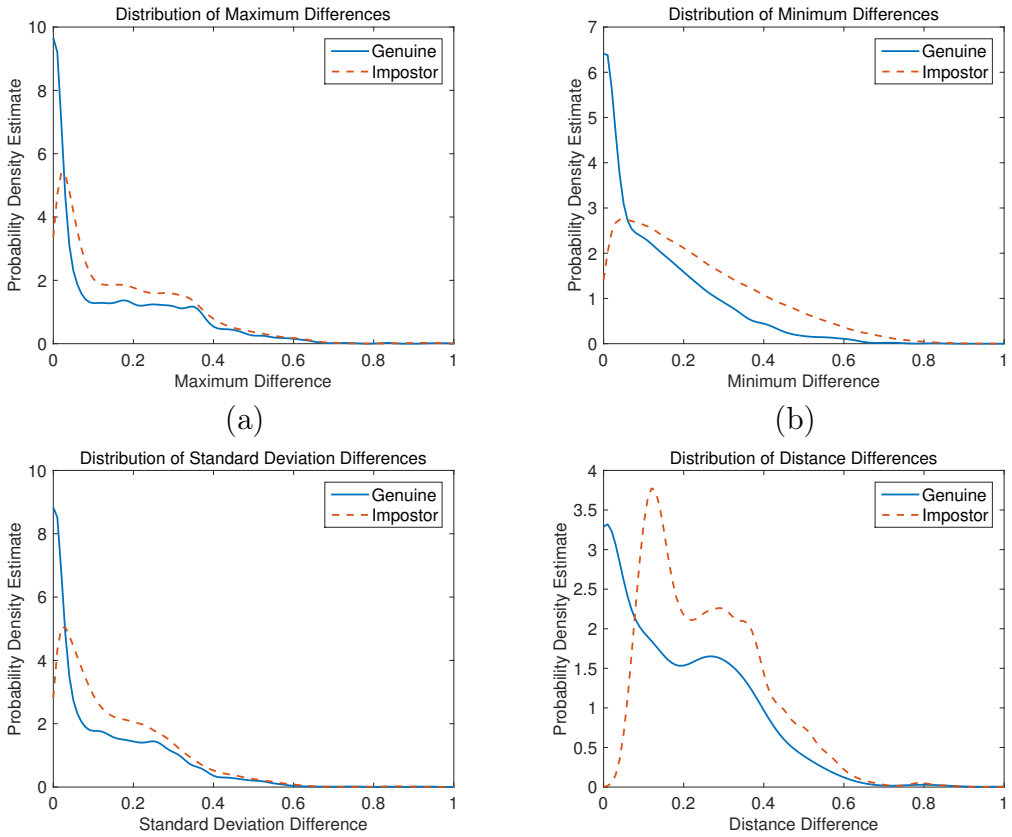


Figure 3.8: Features that summarize BSIF histograms: (a) Maximum frequency of BSIF histograms across devices ; (b) Minimum frequency of BSIF histograms across devices ; (c) Standard deviation of BSIF histograms across devices ; (d) Euclidean distance between image histograms.

for automated matching. Grother and Tabassi formalize the concept that biometric quality measures should predict matching performance [18]. These measures should not be based on human perception which is subjective and may not agree with an automated matching system.

For this work, we assess fingerprint image quality using the NIST Fingerprint Image Quality algorithm (NFIQ) [19], an open source tool provided in the NIST Biometric Image Software (NBIS) distribution [2]. The NFIQ score is calculated for each image (gallery and probe) separately and is represented as an integer in the range [1, 5] where 1 indicates the highest quality and 5 the lowest. NIST provides a recommendation that fingerprint images should be acquired up to four times until an NFIQ of 3 or less is obtained in order to obtain best match results [20]. Lugini *et al.* evaluated the frequency of low genuine match scores as related to image quality and found that in the inter-device scenario even an NFIQ score of 3 leads to reduced matching performance [9]. In this work, we include images with scores in the full range of [1, 5] without excluding or reacquiring images that fall outside of the recommendations. Figure 3.9 shows how NFIQ scores are distributed across devices.

Minutiae Count. Fingerprint minutiae are ridge endings or ridge bifurcations extracted from a fingerprint image [21]. Minutiae points are typically represented by their location and direction, expressed as a triplet, $\mathbf{m} = [x, y, \theta]$. The number of minutiae extracted from an image can be seen as a measure of image quality and may vary based on human interaction with a sensor. For this work, we use MINDTCT, a program from the NBIS distribution to obtain a minutiae count from each image. Figure 3.10 shows the minutiae count across different devices.

Match Score. Automated matching algorithms are used to indicate the degree of similarity between fingerprints. The score is computed as a function of the number of corresponding minutiae. Minutiae points from two images are paired when they

[2] http://www.nist.gov/itl/iad/ig/nbis.cfm

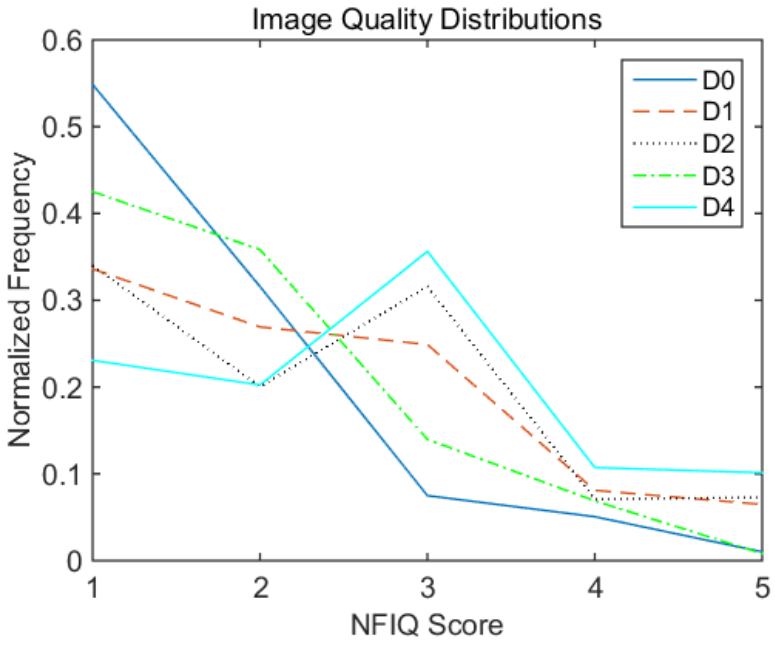


Figure 3.9: Distribution of image quality scores (NFIQ) across each device.

fall within a predefined distance and angle threshold. For this work, match scores are generated between image pairs using the Identix BioEngine Software Development kit. As well as being included in our proposed feature set, these scores act as a baseline for comparison with our results.

Alignment. Placement of a user's finger on a sensing device can affect the location and orientation of minutiae points. By geometrically transforming two sets of minutiae points to the same coordinate system, we are able to relate two impressions of a finger. We use the Generalized Hough Transform detailed in [1] which takes two sets of minutiae points (one from the gallery image and one from the probe image) as input to produce three transformation parameters, Δx, Δy, and $\Delta\theta$, as output.

Pattern Noise. When a fingerprint image is captured, noise can be introduced from various sources [22]. Two main components for such noise include random *photonic noise* and deterministic *pattern noise*. Pattern noise corresponds to a systemic distortion and is present in every image acquired by a particular sensor. The dominant part of the pattern noise is the Photo-Response Non-Uniformity noise (PRNU),

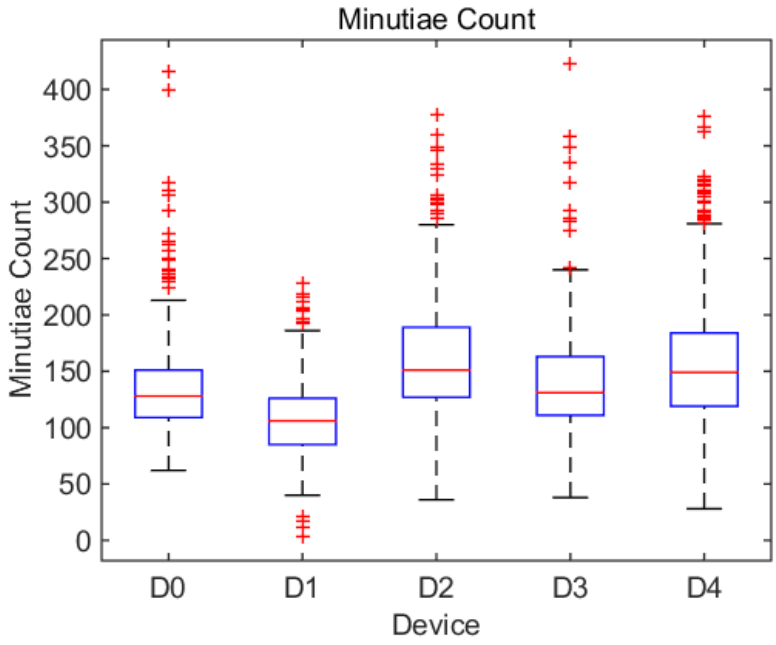
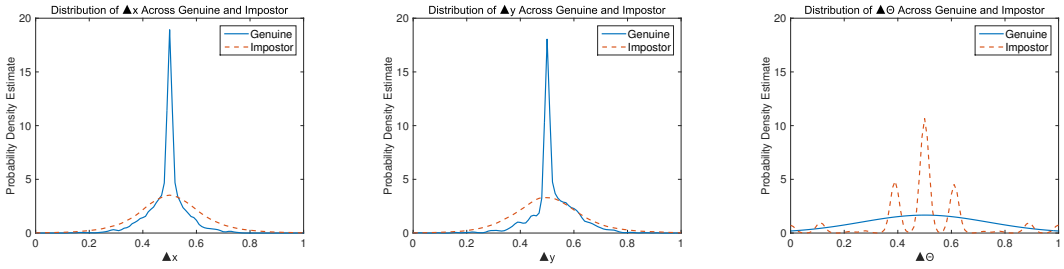


Figure 3.10: Box plots of minutiae count for each device.



(a) (b) (c)

Figure 3.11: Alignment parameters obtained from Generalized Hough Transform of fingerprint image pairs: (a) Δx ; (b) Δy ; (c) $\Delta\theta$.

caused by pixels exhibiting different sensitivity to light. After computing an approximate reference PRNU pattern for each sensor by averaging the residual PRNU of all images taken by each sensor, we obtain the correlation between the reference pattern and the noise pattern extracted from the image as:

$$\rho = corr(\mathbf{n},\mathbf{r}) = \frac{(\mathbf{n} - \mu_n)(\mathbf{r} - \mu_r)}{\|\mathbf{n} - \mu_n\| \, \|\mathbf{r} - \mu_r\|} \tag{3.4}$$

where $\mathbf{n}$ is the residual noise of an image and $\mathbf{r}$ is the reference PRNU pattern for sensor the image is captured with. Figure 3.12 shows the difference of these correlation

values between fingerprint pairs.

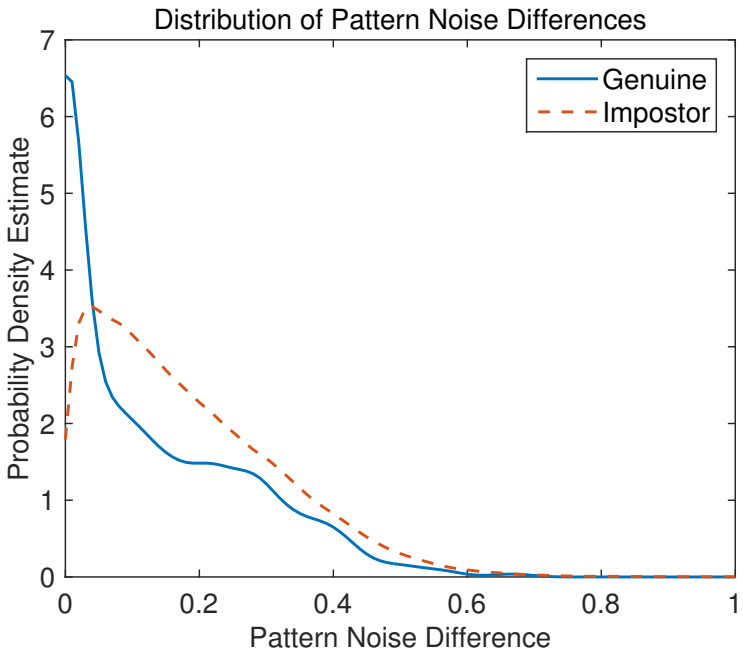


Figure 3.12: Pattern noise value differences between fingerprint pairs.

Image Gradient. Gradient of an image measures the rate of change in grey levels. We calculate image gradient as:

$$\nabla I = \begin{bmatrix} G_x \\ G_y \end{bmatrix} \tag{3.5}$$

where G_x represents differences in the horizontal direction and G_y represents differences in the vertical direction. Both G_x and G_y are matrices the same size as the original image. Magnitude of the gradient is then given by:

$$\nabla \mathbf{I} = [G_x^2 + G_x^2]^{1/2} \tag{3.6}$$

We use the difference of image gradient magnitude means of fingerprint pairs as one of our features. Figure 3.13 shows the distribution of the differences between fingerprint pairs.

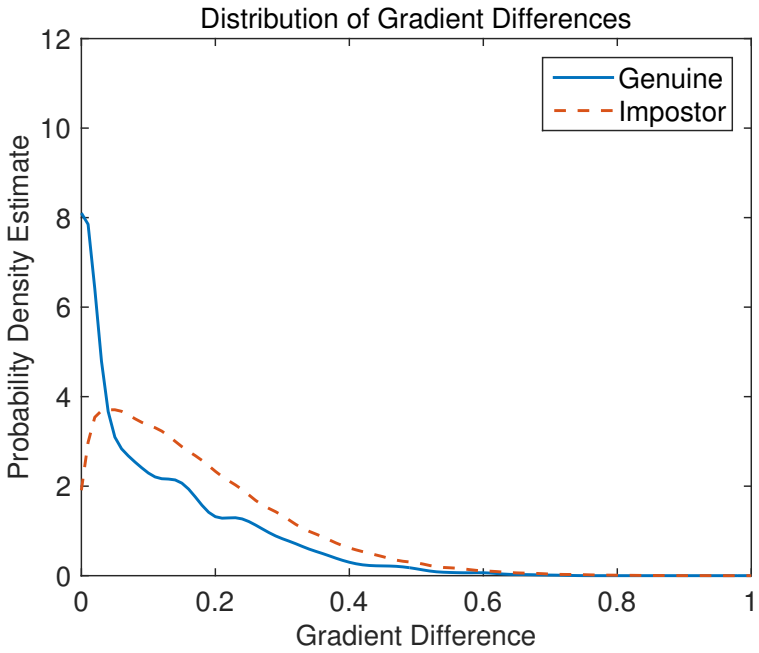


Figure 3.13: Mean gradient value differences between fingerprint pairs.

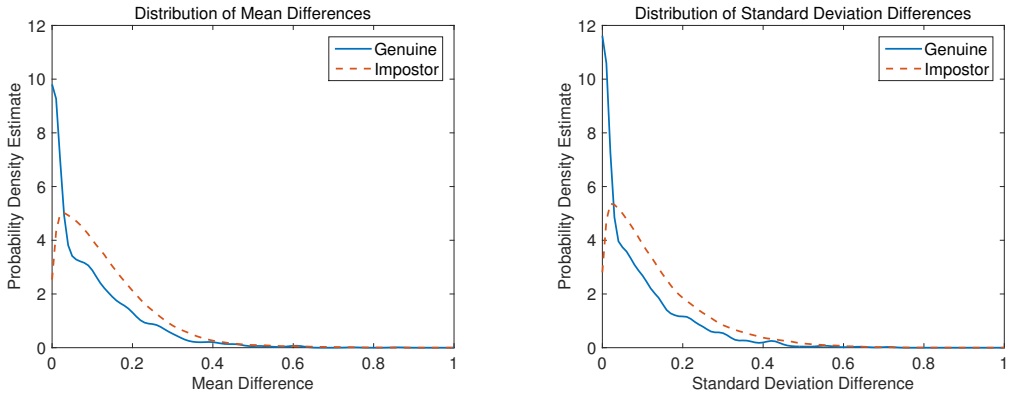
Intensity-based Statistics. For each pre-processed image we calculate the grey level mean and standard deviation. The mean is calculated as:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} A_i \tag{3.7}$$

where each image is represented as a vector, A. Standard deviation is calculated as:

$$S = \sqrt{\frac{1}{N}\sum_{i=1}^{N} |A_i - \mu|^2} \tag{3.8}$$

Figure 3.14 shows the distribution of the mean and standard deviation differences between fingerprint pairs.



(a) (b)

Figure 3.14: Gray level statistic differences between fingerprint pairs: (a) Mean; (b) Standard Deviation.

CHAPTER 4: EXPERIMENTATION

4.1 Data Set

The data set employed in this study consists of fingerprints collected from 494 participants at West Virginia University. Images obtained include individual rolled fingerprints on both the left and right hands as well as left and right thumb slaps. For this study, we only use the right index fingerprints from the full set. Quality measures and match scores were extracted from images at time of acquisition. In order to reduce the number of impostor examples, impostor match scores were generated by dividing users into groups of 100 and matching fingerprints within the same group.

Fingerprints were acquired from each participant by using four live-scan optical sensors (D0-D3) and one set of ink-based ten-print cards (D4). Each user provided fingerprints in the same order, with D4 being last so that ink would not affect the other devices. All sensors and ten-print card scans operate at a resolution of 500 dpi and are FBI certified. Details for each device are provided in Table 4.1.

Our data is separated into three sets for training, validation, and testing. Users are mutually exclusive to each set so that we can test how effective our method is at generalizing to new users. Both the training and validation sets are composed of

Table 4.1: Characteristics of the Live-scan devices used for the fingerprint acquisition carried out in this study.

	Manufacturer	**Model**	**Resolution (dpi)**	**Image size (pixels)**	**Capture area (mm)**
D0	Cross Match	Guardian R2	500	800 x 750	81 x 76
D1	i3	digID Mini	500	752 x 750	81 x 76
D2	L1 Identity Solutions	TouchPrint 5300	500	800 x 750	81 x 76
D3	Cross Match	Seek II	500	800 x 750	40.6 x 38.1
D4	Ten Print Scans	-	500	800 x 715	-

Table 4.2: Details for each sub-set of fingerprint pairs.

Set	Users	Genuine	Impostor
Training	125	1,875	131,250
Validation	125	1,875	131,250
Testing	244	3,660	248,025

fingerprint pairs between 125 users and the testing set is composed of the remaining 244. Further details of impostor and genuine splits are provided in Table 4.2.

4.2 Classification

For this work, classification experiments were performed using the Random Forest classifier, a variation of traditional bootstrap aggregation (bagging) of weak tree-structured learners. Ensemble methods combine hypotheses obtained from multiple weak learners in order to obtain a final classification of input data [23]. Ensembles of weak tree-based learners show a general resistance to the effects of overfitting (even as the number of learners increases) and have been successful in scenarios where a decision boundary is too complex for a single learner.

Random Forests consist of a collection of k tree-based classifiers $h\Theta_1, ..., h\Theta_k$ where Θ_i is a subset of the full training data set (randomly sampled with replacement) used to train a decision tree, h_i [24]. When building each tree, a random subset of features is used at each node to determine the split. During classification, an input sample is entered as input to each of the k trees. The final class label for the sample is decided by a majority vote from each tree's output. Random Forests were recently used in [1] to show improvement in fingerprint interoperability.

There are two major parameters to select when growing out a Random Forest. The first is the number of trees that will make up the ensemble. In general, the main negative aspect of increasing the number of trees is an increase in computational time for vanishing improvements classification rates. The second parameter to tune is the

number of features randomly sampled at each split. For classification tasks, Breiman originally performed tests using a single randomly chosen feature and $log_2(M + 1)$ features where M is the total number of features for a data example. Many implementations of Random Forest (including the Matlab implementation we use) default instead to $\sqrt{M}$, which for our feature set ends up being very similar.

In selecting parameters for our model, we tune on a coarse selection of the number of trees: [20 100 500 2500]. For number of randomly sampled features we use the numbers [1 5 13 26], corresponding to a single feature, $\sqrt{M}$, $\frac{M}{2}$, and M (corresponding to traditional bagging methods) respectively. Table 4.3 shows the results of parameter tuning on our validation set of fingerprint pairs. Area under the curve of the detection-error tradeoff, in which lower values are better, was used to evaluate the effectiveness of each model. While three different models gave the same value for AUC, the one in which N = 500 and F = 13 was selected for evaluation of the test set for achieving the best overall error-rates in the fewest number of trees.

4.3 Results

For our results, we use False Match Rates (FMR) and False Non Match Rates (FNMR) to evaluate the performance of our proposed method. FMR corresponds to testing instances where an impostor is incorrectly labelled as a genuine match whereas FNMR corresponds to instances where a genuine match is labelled as an impostor. Figure 4.1 shows the results of the selected Random Forest classifier on the final set of features. Our proposed approach obtains a FMR of 0.004% and a FNMR of 4.454% when tested on fingerprint pairs from users not seen by the classifier during training. Figure 4.1 shows the DET obtained form our classifier compared to the DET of using the match score alone (serving as a baseline).

We additionally see an improvement over the results of [1] when we apply the previous method to a data set with fingerprint pairs that are mutually exclusive between training and testing. Figure 4.2 shows the DET obtained from our model

Table 4.3: Comparison of validation results on different Random Forest parameters where N is the number of trees, F is the number of randomly selected features at each split, FMR is the percentage of impostor examples misclassified, FNMR is the percentage of genuine examples misclassified, and DET AUC is the area under the curve of the Detection-Error Tradeoff.

Random Forest Validation Error				
N	**F**	**FMR (%)**	**FNMR (%)**	**DET AUC**
20	1	0.001	40.000	0.200
	5	0.003	3.733	0.019
	13	0.006	3.147	0.016
	26	0.009	2.987	0.015
100	1	0.000	51.520	0.258
	5	0.004	3.467	0.017
	13	**0.008**	**2.880**	**0.014**
	26	0.008	2.880	0.015
500	1	0.000	35.040	0.175
	5	0.003	3.627	0.018
	13	**0.005**	**2.880**	**0.014**
	26	0.013	2.987	0.015
2500	1	0.000	49.547	0.248
	5	0.003	3.573	0.018
	13	**0.005**	**2.880**	**0.014**
	26	0.008	2.987	0.015

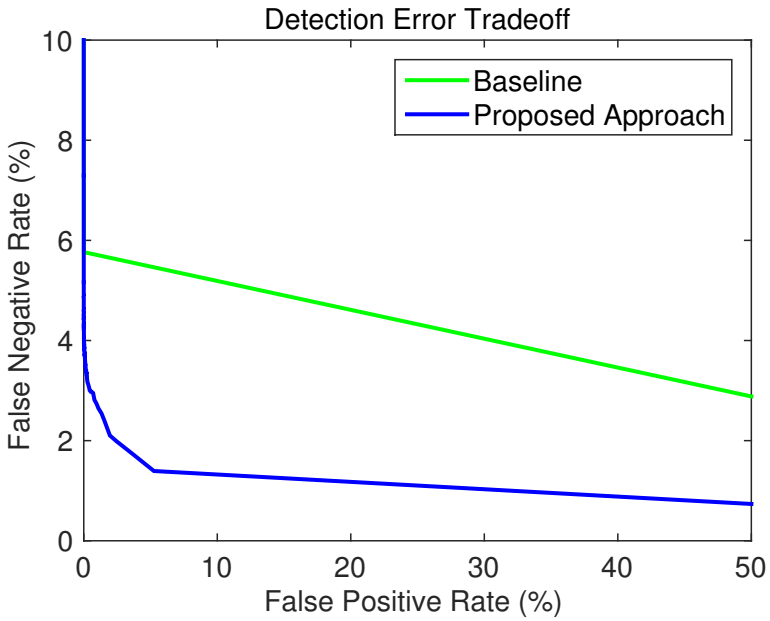


Figure 4.1: Comparison of the results obtained with the proposed approach with the Detection Error Tradeoff with the baseline of a thresholded match score.

Table 4.4: False Non-Match Rates obtained by each method when False Match Rate is held equal at 0.01%.

Method	**FNMR (%)**
Match Score	5.77
Marasco et al. [1]	4.73
Proposed Method	4.24

compared to the DET obtained by using the previously used feature set and classifier on the same training and testing split that we used.

We further illustrate improvements over previous methods by comparing FNMR when FMR is held constant. 4.4 shows the results for each of the previously discussed DETs evaluated at a FMR of 0.01%.

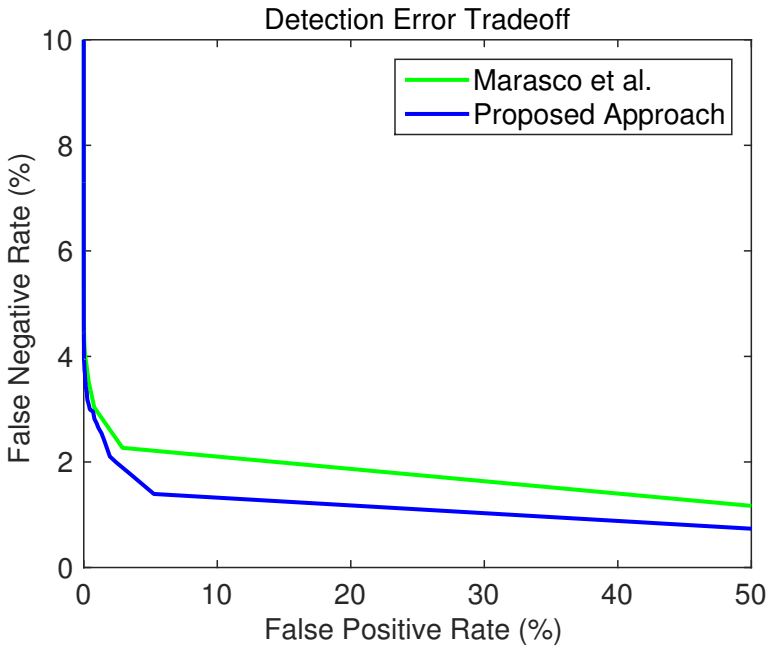


Figure 4.2: Comparison of the Detection Error Tradeoff curve obtained from the proposed method with that obtained by using the features and classifier presented in [1] on the same training and testing set.

CHAPTER 5: CONCLUSIONS

5.1 Feature Importance

While we show improvements over previous methods, the primary purpose of this study is to explore Wavelet Entropy and Binarized Statistical Image Features as applied to the problem of fingerprint interoperability. To perform this evaluation in isolation of other changes over previous methods, we compare the DET obtained from our full feature set with the DET of three other sets. The first additional set does not include wavelet entropy measures, the second does not include binarized statistical image features, and the third includes neither wavelet entropy nor the binarized statistical image features. Figure 5.1 shows that there is very little change in the DET curves between these feature sets. Table 5.1 further illustrates that results when held at 0.01% FMR appear slightly worse when using the proposed features. This seems at odds with feature ranking methods, class correlation for instance, that included our new features high in rankings.

Despite the overall improvements, there is no definitive proof that our newly proposed features have a strong impact on fingerprint interoperability. Although we make use of features explored in previous work, there are several changes that may account

Table 5.1: False Non-Match Rates obtained by feature sets when False Match Rate is held equal at 0.01%.

Method	**FNMR (%)**
Proposed	4.24
No Entropy	4.13
No BSIF	4.18
Neither	4.10

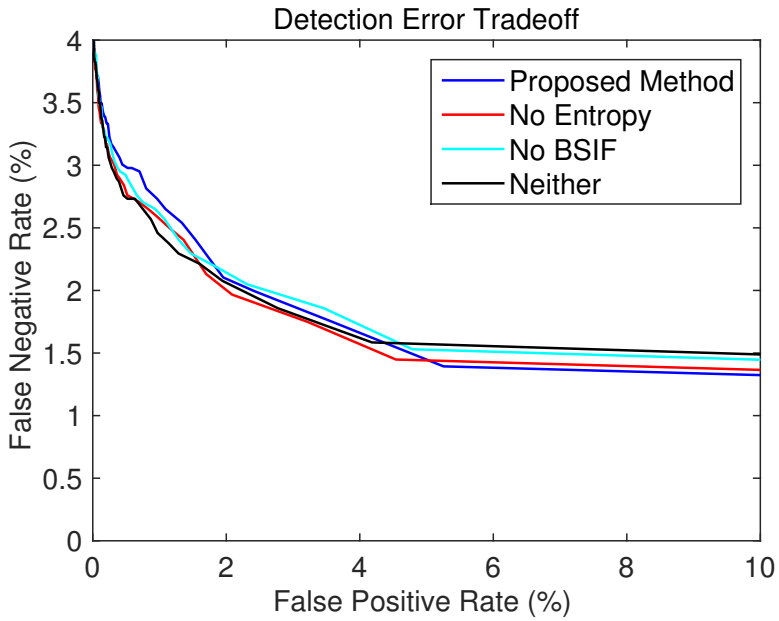


Figure 5.1: Comparison of the results obtained with the full feature set that includes Wavelet Entropy and BSIF measures with three feature sets: (1) one without entropy; 2 one without BSIF; (3) one with neither entropy nor BSIF.

for the difference in results. First, most of the features used in this study represent a pair of fingerprints by taking the unsigned difference between corresponding values. This choice was primarily made as a way to explore the distribution of genuine and impostor examples in a more intuitive manner. Second, our pre-processing method may have an effect; however, it is not known exactly what steps were taken during pre-processing in previous work. Lastly, it is possible that there are differences in both feature extraction and learning algorithm implementations between versions of software components used.

5.2 Future Work

Future directions to consider for the study of fingerprint interoperability include extending experiments to include additional sensors. Ideally the features selected to improve interoperability could generalize even to devices not seen during model training. Different sensing technologies could also be explored. This study focuses on the use of optical sensors, but other device types exist such as capacitive and ultra-

sound. It is important to note that the acquisition devices involved in this study are high end applicances and that attempts to extend research into types of sensors such as those present on consumer devices may require drastic changes to features and methodology. Further work may also include the exploration of any number of image pre-processing techniques, additional features and classification algorithms.

REFERENCES

[1] E. Marasco, L. Lugini, and B. Cukic, "Minimizing the Impact of Low Interoperability between Optical Fingerprint Sensors," *Biometrics: Theory, Applications and Systems (BTAS)*, pp. 1–8, 2013.

[2] A. Ross and A. Jain, "Biometric Sensor Interoperability: A Case Study in Fingerprints," *International ECCV Workshop on Biometric Authentication*, pp. 134–145, 2004.

[3] F. Alonso-Fernandez, R. N. Veldhuis, A. M. Bazen, J. Fiérrez-Aguilar, and J. Ortega-Garcia, "Sensor interoperability and fusion in fingerprint verification: a case study using minutiae-and ridge-based matchers," *2006 9th International Conference on Control, Automation, Robotics and Vision*, pp. 1–6, 2006.

[4] E. P. Kukula, S. J. Elliott, and V. G. Duffy, "The effects of human interaction on biometric system performance," *International Conference on Digital Human Modeling*, pp. 904–914, 2007.

[5] S. K. Modi, S. J. Elliott, and H. Kim, "Statistical analysis of fingerprint sensor interoperability performance," *Biometrics: Theory, Applications, and Systems, 2009. BTAS'09. IEEE 3rd International Conference on*, pp. 1–6, 2009.

[6] A. Ross and R. Nadgir, "A Calibration Model for Fingerprint Sensor Interoperability," *SPIE*, vol. 6202, 2006.

[7] N. Poh, J. Kittler, and T. Bourlai, "Improving Biometric Device Interoperability by Likelihood Ratio-based Quality Dependent Score Normalization," *First IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS)*, pp. 1–5, 2007.

[8] N. Poh, J. Kittler, and T. Bourlai, "Quality-based Score Normalization with Device Qualitative Information for Multimodal Biometric Fusion," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 40, no. 3, pp. 539–554, 2010.

[9] L. Lugini, E. Marasco, B. Cukic, and I. Gashi, "Interoperability in Fingerprint Recognition: a Large-Scale Study," *Workshop on Reliability and Security Data Analysis (RSDA), Budapest*, pp. 1–6, June 2013.

[10] P. D. Kovesi, "MATLAB and Octave functions for computer vision and image processing."

[11] M. Weeks, *Digital Signal Processing Using MATLAB & Wavelets*. Jones & Bartlett Learning, 2010.

[12] M. D. Marsico, M. Nappi, D. Riccio, and G. G. Tortora, "Entropy-based Template Analysis in Face Biometric Identification Systems," *Signal, Image and Video Processing*, vol. 7, no. 3, pp. 493–505, 2013.

[13] H. Zheng-You, C. Xiaoqing, and L. Guoming, "Wavelet Entropy Measure Definition and Its Application for Transmission Line Fault Detection and Identification," *International Conference on Power System Technology*, pp. 1–6, 2006.

[14] 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*, vol. 24, no. 7, pp. 971–987, 2002.

[15] V. Ojansivu and J. Heikkilä, "Blur Insensitive Texture Classification using Local Phase Quantization," *Image and Signal Processing*, pp. 236–243, 2008.

[16] J. Kannala and E. Rahtu, "BSIF: Binarized Statistical Image Features," *21st International Conference on Pattern Recognition (ICPR)*, pp. 1363–1366, 2012.

[17] G. L. M. F. R. L. Ghiani, A. Hadid, "Fingerprint Liveness Detection Using Binarized Statistical Image Features," *Sixth IEEE International Conference on Biometrics: Theory, Applicaitons, and Systems (BTAS)*, pp. 1–6, 2013.

[18] P. Grother and E. Tabassi, "Performance of Biometric Quality Measures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 531–543, 2007.

[19] E. Tabassi and C. L. Wilson, "A Novel Approach to Fingerprint Image Quality," *IEEE Conference on Image Processing 2005*, vol. 2, pp. 37–40, 2005.

[20] R. C. P. Grother, W. Salamon, "Biometric Specifications for Personal Identity Verification," *Special Publication (NIST SP) - 800-76-2*, 2013.

[21] N. K. R. S. P. R. M. Bolle, A. W. Senior, "Fingerprint Minutiae: A Constructive Definition," *Biometric Authentication*, pp. 58–66, 2002.

[22] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *Information Forensics and Security, IEEE Transactions on*, vol. 1, no. 2, pp. 205–214, 2006.

[23] R. Polikar, "Ensemble based systems in decision making," *Circuits and systems magazine, IEEE*, vol. 6, no. 3, pp. 21–45, 2006.

[24] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.