Identifying Sensors from Fingerprint Images

Nick Bartlow Nathan Kalka Bojan Cukic
Arun Ross

West Virginia University
Morgantown, WV, USA
{nick.bartlow, nathan.kalka, bojan.cukic, arun.ross}@mail.wvu.edu

Abstract

In this paper we study the application of hardware fingerprinting based on PRNU noise analysis of biometric fingerprint devices for sensor identification. For each fingerprint sensor, a noise reference pattern is generated and subsequently correlated with noise residuals extracted from test images. We experiment on three different databases including a total of 20 fingerprint sensors. Our results indicate that fingerprint sensor identification at unit level is attainable with promising prospects. Our analysis indicates that in many cases identification can be performed even when one only has access to a limited number of samples. For two of the three databases one can train on less than 8 images per device and establish sensor identification with little or no misclassification error. On the third database, high levels of identification performance can be achieved when training on similar amounts of images required for other types of sensor identification such as cameras or scanners.

1. Introduction

As the field of biometrics continues to grow, so does its areas of application. Such areas can include access control in protected sites and border control, remote authentication in commercial applications, and identification of criminal suspects or enemies on the battlefield. Regardless of the intended application, various measures must be taken to ensure the accuracy and integrity of these deployments. Two ways that biometric systems can be compromised include fabrication and alteration of data. Fabrication of biometric data could occur at many points within a biometric system and usually is the result of an act with malicious intent. Whether at the time of data acquisition, matching, or database access, various vulnerabilities may allow raw biometric images to be created and maliciously injected into a system. Similarly, biometric data may also be maliciously altered throughout the course of operation in a biometric system [1]. Besides actions with malicious intent, unintentional alteration of images during the collection, transmission, or storage blocks of a system can take place. To make matters worse, whether intentional or unintentional, there often is no obvious cue that that an image has been fabricated or altered in the first place. This is of particular importance to applications where a "chain of evidence" must be established. Such a chain is useful in assembling cases to prosecute criminal activity, establishing identity dominance in the battlefront, and discovering fraudulent activity in commercial systems. In an effort to minimize the presence of fabricated / altered images in such systems, the notion of source identification is applied. Falling under the field of digital forensics, digital hardware fingerprinting provides the ability to identify and validate the source hardware which captured an image. Whether establishing a chain of evidence or addressing a specific biometric vulnerability, application of digital hardware fingerprinting for biometric image source validation should prove to be very useful. Digital hardware fingerprinting is the process of identifying the source hardware used to capture an image regardless of the scenery or primary image content. The primary method of identifying the source hardware from which an image originated is an-

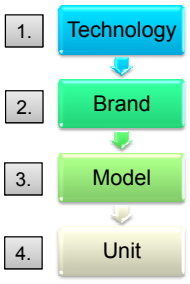


Figure 1. Source Hardware Identification Levels.

alyzing differences in images resulting from imaging sensor imperfections [2]. Due to slight inconsistencies in the production process, all sensors are subject to small manufacturing imperfections. These imperfections lead to the necessary observance of noise (sometimes visually undetectable by humans) in images collected. Although previous work has focused on devices using optical technology to capture images, such noise would also be present in sensors relying on different technologies for image capture such as capacitance, thermal, or piezoelectric. Identification of source hardware can potentially occur at different levels of granularity. Figure 1 shows four levels that may be considered in a source hardware identification problem: technology, brand, model, and unit. Based on the four levels, various questions can be answered:

- Was the image in question captured from a sensor relying on technology X or technology Y? (Technology)
- Was the image in question captured from a device manufactured by vendor X or vendor Y? (Brand)
- Was the image in question captured from a device corresponding to model X or model Y manufactured by vendor Z? (Model)
- Was the image in question captured from a unit A or unit B of model X manufactured by vendor Z? (Unit)

Naturally, unique challenges exist to performing source model identification at different levels. In this paper, we will determine source identification at the brand (2) and unit levels (4) using photo-response nonuniformity noise (PRNU) across multiple units of two different biometric fingerprint readers. The contribution of the work is two-fold. To our knowledge, it is the first work to demonstrate the ability to identify the hardware source used to collect biometric fingerprint images. To do so, we adopt the technique presented by Lukas et al. in [2]. Secondly, we formally establish the effect of varying the amount of images used to arrive at reference templates for readers at the unit and brand level. The remainder of the paper is broken down as follows: Section 2 outlines past work in identification of image source captured by digital cameras or scanners, Section 3 defines the algorithm applied for digital hardware fingerprinting, Section 4 outlines the design of the experiment including a description of the data set and testing methodology, Section 5 presents the results of the identification experiment, Section 6 provides a discussion including considerations of interest, finally Section 8 summarizes the contribution of the work.

2. Related Work

While the authors are not aware of any prior work on digital hardware fingerprinting specific to biometric capture devices, a large bed of research exists on digital image forensics, specifically images captured from photographic cameras and document scanners. There are many proposed approaches to performing source hardware identification at different levels that are well summarized in a survey of the field by Khanna et al. in [3]. In particular, Choi et al. proposed an approach based on lens distortions in [4]. Geradts et al. proposed a technique relying on sensor imperfections such as dead pixels in [5]. Kharrazi et al. outline 34 image features including average pixel values, RGB pair correlations, and neighbor distribution center of mass [6]. Similar to image features, Bayram et al. proposed a technique relying on the measurement of interpolation artifacts in an image due to the use of a color filter array (CFA) [7]. Additionally, although not inherent to the capture devices, the notion of sensor dust characteristics has been explored by Dirik et al. in [8]. The approach proposed by Lukas et al. in [2] based on measuring pixel nonuniformity (PNU) noise is the basis of this work and is arguably the most promising technique to date. For the purposes of brevity, we defer to the cited sources for further details of related approaches.

3. Approach

As a means to identify fingerprint readers at the brand and unit level, we first adopt the approach proposed by Lukas et al. in [2]. This approach is based on estimating pixel nonuniformity (PNU), a portion of the photo-repsonse nonuniformity (PRNU) inherent to every image captured by the readers. The remainder of this section is broken down into two parts: a description of the general framework for identifying hardware sources through PNU noise and a description of the wavelet-based denoising algorithm utilized in [2].

3.1. Identification Process

The process of sensor identification can then be broken down into two main steps:

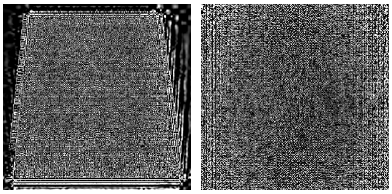
1. **Calculate Reference Patterns**. For each fingerprint reader, calculate a reference pattern by taking an average of the noise residual estimates across multiple training images as seen in Equation 1.

$$\mathbf{R}_i = \frac{\sum_{k=1}^{N} \mathbf{p}^{(k)} - F(\mathbf{p}^{(k)})}{N} \tag{1}$$

Here, N represents the number of images used to generate the reference pattern, $\mathbf{R}_i$, $\mathbf{p}^{(k)}$ represents each image in the training set, and F represents a denoising filter. It should be noted that while F can represent any denoising filter, Lukas et al. found that a wavelet-based approach yielded the best results [2]. The specifics of the wavelet-based denoising is described later in this section.

2. **Correlate Noise Residuals to Reference Patterns**. For each image to be tested, extract the noise residual $\mathbf{p}^{(k)} - F(\mathbf{p}^{(k)})$, and measure the correlation, $\mathbf{C}$, to each reference patterns, $\mathbf{R}_i$, for all of the reference patterns in question. In [2], Lukas et al. propose the correlation measure seen in Equation 2, although in theory, any correlation measure could be applied.

$$\mathbf{C}_i = corr(\mathbf{p}, \mathbf{R}_i) = \frac{(\mathbf{p} - \bar{\mathbf{p}}) \cdot (\mathbf{R}_i - \bar{\mathbf{R}}_i)}{\|(\mathbf{p} - \bar{\mathbf{p}})\| \|(\mathbf{R}_i - \bar{\mathbf{R}}_i)\|} \quad (2)$$



(a) Microsoft (b) BioTouch

Figure 2. Example reference patterns for Microsoft and BioTouch Readers (16 images averaged).

3.2. Wavelet-based Denoising Algorithm

The wavelet-based denoising approach in [2] is described in four steps.

1. Calculate the first through fourth wavelet decompositions of the original noisy image using the 8-tap Daubachies Quadratic Mirror Filters (QMF). The vertical, horizontal, and diagonal subbands are denoted by $v(i,j)$, $h(i,j)$, and $d(i,j)$ respectively. Here (i,j) represents the coefficients for each pixel in each of the three subbands.

2. In each subband, estimate the local variance of the noise-free image for each wavelet coefficient using MAP estimation for four sizes of a $W \times W$ neighborhood N, for $W \in \{3,5,7,9\}$ as seen in Equation 3.

$$\hat{\sigma}_W^2(i,j) = max\left(0, \frac{1}{W^2} \sum_{(i,j)\in N} h^2(i,j) - \sigma_0^2\right) \quad (3)$$

Then apply Equation 4 to arrive at the minimum of the four variances as the final estimate.

$$\hat{\sigma}_W^2(i,j) = min\left(\sigma_3^2(i,j), \sigma_5^2(i,j), \sigma_7^2(i,j), \sigma_9^2(i,j)\right) \quad (4)$$

3. Determine the denoised wavelet coefficients by applying the Wiener filter as seen in Equation 5

$$h_{den}(i,j) = h(i,j)\frac{\hat{\sigma}^2(i,j)}{\hat{\sigma}^2(i,j) + \sigma_0^2} \quad (5)$$

Similarly, the filter is applied to $v(i,j)$ and $d(i,j)$.

4. Steps 1-3 are repeated for each level and color channel.

In [2], the authors used $\sigma_0 = 5$ in the experiments as do we in this work. It should be noted that due to the grayscale nature of the fingerprint images, it is not necessary to perform Step 4 across multiple color channels.

4. Experimental Design

Three different datasets were considered in this experiment. The first is a WVU collection consisting of images from two sensor models (3 units each). The second collection comes from both WVU and Clarkson which consists of images from three sensor models (2 units each). The last set is from the first three years of the Fingerprint Verification Competition (FVC) which consists of images from 8 different units. The WVU data was collected specifically with hardware fingerprinting experiments in mind, while the other two datasets were collected primarily for biometric testing purposes. Along those lines, the WVU dataset fingerprint images resulted from 4 subjects providing 100 images per sensor (25 images from 4 digits) for a total of $2,400$ images. The WVU / Clarkson datasets each have substantially more images and the experiment only used a subset of the fingerprint images available from each. Specifically, we randomly selected 1000 images per sensor pertaining to the right index and thumb. The FVC dataset was comprised of 10 subjects with 8 images per subject, collected from the index and middle finger, totaling 640 images. A summary of the sensors used including details, example images, and noise residuals can be found in Figure 3.

Although the amount of users found in the WVU or FVC data may be prohibitively small for a traditional biometric experiment, we note the experiment is not studying biometric recognition or identification. Instead, we are studying sensor identification. To that effect, we believe the variety subjects and associated fingerprint digits provides sufficient variation for our tests. We applied a cross-validation framework for all three datasets for testing the proposed methods. In our experiments, we tested the success of the digital fingerprinting techniques while varying the amount of images used to generate reference patterns using the methodology described in the previous section. Table 1 lists the training and testing breakdowns for each dataset. It is important to note that 10 fold cross validation was applied for each

Sensor	Fing.	Noise	Sensor	Fing.	Noise
WVU Identix #1			WVU Microsoft #1		
WVU Identix #2			WVU Microsoft #2		
WVU Identix #3			WVU Microsoft #3		
WVU Precise			Clarkson Precise		
WVU Secugen			Clarkson Secugen		
WVU CrossMatch			Clarkson CrossMatch		
FVC KeyTronic			FVC Microelectronics		
FVC Identicator			FVC Identix		
FVC Biometrika			FVC Precise		
FVC CrossMatch			FVC DigitalPersona		

(a) Example fingerprints and noise residuals from three different data sets.

Brand	Model	Tech.	Width	Height
Microsoft (WVU 1-3)	Fingerprint Reader	O	355	390
Identix (WVU 1-3)	BioTouch200	O	256	255
Precise Biometrics (WVU / Clarkson)	AX 100	C	200	200
Secugen (WVU / Clarkson)	Hamster III	O	260	300
CrossMatch (WVU / Clarkson)	Verifier 300 LC	O	640	480
KeyTronic (FVC)	Secure Desktop Scanner	O	300	300
Microelectronics (FVC)	TouchChip	C	256	364
Identicator Technology (FVC)	DF-90	O	448	478
Identix (FVC)	TouchView II	O	388	374
Biometrika (FVC)	FX2000	O	296	560
Precise Biometrics (FVC)	100 SC	C	300	300
CrossMatch (FVC)	V300	O	640	480
DigitalPersona (FVC)	U.are.U 4000	O	328	364

(b) Fingerprint sensor details from three different data sets. Tech. = Technology {O=optical, C=capacitive}.

Figure 3. Fingerprint images, noise residuals, and sensor details for each sensor in the three databases.

WVU									
Train	1	2	4	8	16	32	64	128	256
Test	399	398	396	392	384	368	336	272	144
WVU / Clarkson									
Train	1	2	4	8	16	32	64	128	256
Test	500								
FVC									
Train	1	2	4	8	16	32	64	n/a	n/a
Test	79	78	76	72	64	48	16	n/a	n/a

Table 1. Experimental training variation measured in images / sensor.

test. Therefore, total number of tests in the WVU and FVC datase results range depending on the split with WVU ranging from $3,990$ (399 * 10) to $1,440$ (144 * 10) and FVC results range from 790 (79 * 10) to 160 (16 * 10) tests. On the other hand, the test size was constant for the WVU / Clarkson set due to the availability of images and includes includes $5,000$ tests (500 * 10 folds).

5. Experimental Results

The results in this section reflect attempts to perform sensor identification at the unit level. In this case, we wish to distinguish which unit the image was captured from given a pool of units. Therefore, a test noise residual is compared against reference patterns for each unit from the dataset in consideration. For each dataset, we provide example histograms of match / non-match distributions, confusion matrices for specific train / test sets, and Cumulative Match Characteristic (CMC) curves as the train / test splits vary.

5.1. FVC Dataset

The first set of experiments was performed on the FVC data. Figure 4 displays the difference in correlation between match and non-match comparisons of test noise residuals and sensor reference patterns. As can be seen in the figure, perfect separation is achieved when considering Identicator test residuals against reference patterns produced from 32 training images per sensor. While perfect separation was achieved in this instance, this was not the case across all sensors in all train and test splits. An example of which can be seen in Table 2. In the table, we see that the Identix test residuals are occasionally misclassified as having originated from other sensors. It is interesting to note that the distribution of errors is fairly uniform across the other seven sensors. With the exception of the Identix sensor, no errors are made on experiments training on at least 8 images. Figure 5 shows a Cumulative Match Characteristic (CMC) plot which indicates the overall accuracy across sensors as the train / test splits are varied. Here we see the rank one identification rate when training on 1 image per sensor falls around 85%. This is fairly high considering this is the smallest amount of data that could be used to generate a

reference pattern for a sensor. When 64 images are used to generate reference patterns for each sensor, the rank 1 identification rate exceeds 98%. Furthermore, the only reason the identification rate is not 100% is due to a small amount of errors made on classifying the Identix noise residuals.

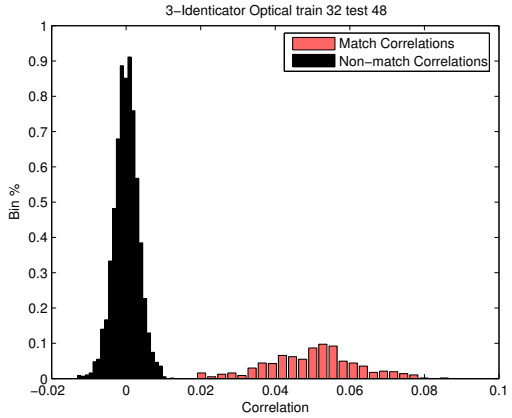


Figure 4. FVC example match and non-match distributions with 32 training images per sensor.

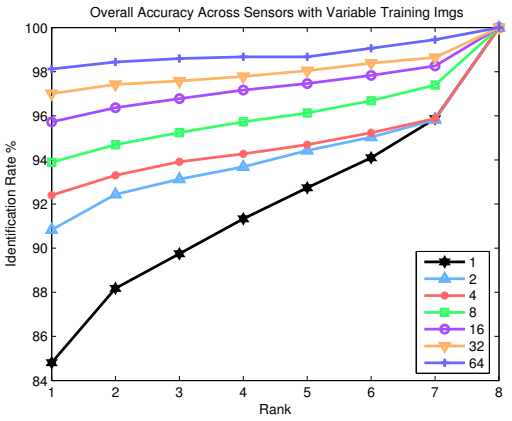


Figure 5. FVC sensor identification as a function of training set size.

5.2. WVU Dataset

The performance on the WVU dataset is the highest among the three datasets considered. Virtually all instances achieve perfect separation, therefore it is not beneficial to display a histogram of match and non-match distributions. However, in Table 3, the confusion matrix resulting from training on 1 image per sensor is displayed. Again considering the minimal training requirement, we see only sporadic errors across the $3,990$ test cases per sensor. This high level of performance is reflected in the CMC curve shown in Figure 6. Here, only 2 train / test scenarios do not achieve perfect rank 1 identification (training on 1 and 2 images per sensor). Even still, the rank 1 identification when using only 1 image per sensor exceeds 99%.

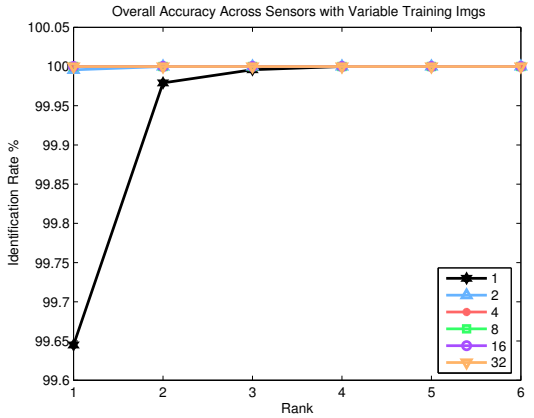


Figure 6. WVU sensor identification as a function of training set size.

5.3. WVU / Clarkson Dataset

The most challenging dataset for sensor identification ending up being the WVU / Clarkson dataset. Observing a notable difference compared to the previous to databases, Figure 7 shows a slight overlap between match and non-match distributions when generating reference templates from 128 training images. In all but one sensor in the WVU and FVC datasets, training on 128 images was unnecessary to achieve perfect separation. This pattern can also be seen in Table 4 which displays the confusion matrix when training on 128 images. While the results can still be considered promising as the overall rank 1 identification accuracy is near 90%, we observe far more errors across the test cases. It is also interesting to note where the errors are made. Somewhat intuitively, more errors are made misclassifying the WVU Secugen noise residuals as Clarkson Secugen noise residuals than any other sensor's residuals. This may be the case as they are the same model sensor. Surprisingly, this pattern does not hold true for the WVU CrossMatch residuals as they are misclassified as Clarkson residuals the fewest number of times when compared to the other sensors. We are still investigating why this may be the case. Figure 8 displays the CMC plots for the final dataset. Once again, this data proved to be the most challenging of the three sets as we see the rank 1 identification accuracy drops to 45%. However, any reasonable application of source identification will likely have access to more than one training image to generate reference patterns. Along

Actual \ Classified	KeyTronic	Microelectronic	Identicator	Identix	Biometrika	Precise	CrossMatch	DigitalPersona
KeyTronic	480	0	0	0	0	0	0	0
Microelectonic	0	480	0	0	0	0	0	0
Identicator	0	0	480	0	0	0	0	0
Identix	14	19	25	357	9	18	13	25
Biometrika	0	0	0	0	480	0	0	0
Precise	0	0	0	0	0	480	0	0
CrossMatch	0	0	0	0	0	0	480	0
Digital Persona	0	0	0	0	0	0	0	480

Table 2. FVC confusion matrix when training on 32 images per sensor

Actual \ Classified	BioTouch #1	BioTouch #2	BioTouch #2	Microsoft #1	Microsoft #2	Microsoft #3
BioTouch #1	3942	14	32	1	1	0
BioTouch #2	6	3967	17	0	0	0
BioTouch #3	3	10	3977	0	0	0
Microsoft #1	0	0	0	3989	1	0
Microsoft #2	0	0	0	0	3990	0
Microsoft #3	0	0	0	0	0	3990

Table 3. WVU confusion matrix when training on 128 images per sensor

those lines, the rank 1 identification rate when training on 256 images is approximately 95% which can once again be considered promising results.

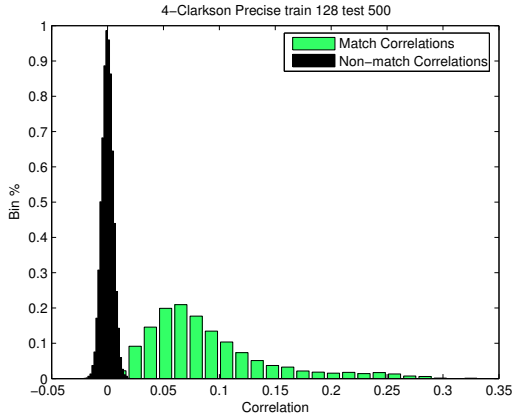


Figure 7. WVU / Clarkson example match and non-match distributions with 128 training images per sensor.

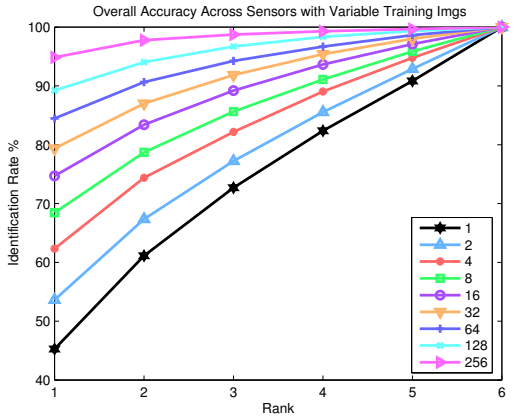


Figure 8. WVU / Clarkson sensor identification as a function of training set size.

6. Discussion

While the results of the experiments clearly demonstrate that inherent PNU noise can be used as a means for performing sensor hardware identification in biometric fingerprint readers, there are a number of considerations which must be mentioned. Most notably, the databases tested only contain limited sets of pairs of identical units (3 sensors of each model in WVU and 2 sensors of each model in WVU / Clarkson). Increasing the number of identical units may result in a decrease in identification performance. At this point, it is unclear if sensor identification with the applied technique would be possible in a pool of 100's or $1,000$'s of the same model sensor at the unit level. Additionally, as different models of fingerprint readers capture images at different resolutions (see Figure 3) cropping of noise residuals was performed when necessary as the chosen method of correlation requires that the noise residuals have the same dimension. Although the results seem to indicate this method of handling dissimilar images sizes is sufficient, there are a number of options one may choose to exercise when dealing with this issue, not the least of which is resizing the original image before denoising. The images could be resized but this may introduce artifacts that could artificially enhance performance. To avoid this, different correlation procedures could be applied such as normalized cross correlation which would not require equally sized images.

Actual \ Classified	WVU Precise	WVU Secugen	WVU CrossMatch	Clarkson Precise	Clarkson Secugen	Clarkson CrossMatch
WVU Precise	4979	0	0	21	0	0
WVU Secugen	364	3146	179	492	596	223
WVU CrossMatch	313	188	3781	366	246	106
Clarkson Precise	32	3	0	4963	2	0
Clarkson Secugen	14	25	7	14	4940	0
Clarkson CrossMatch	27	1	2	1	10	4959

Table 4. WVU / Clarkson confusion matrix when training on 128 images per sensor

7. Future Work

Many areas of this work require further investigation and the work is ongoing. Obvious tasks such as increasing the amount of sensors in the identification pools are being considered but more fundamental issues are the primary areas of concern. For instance, we are working on improving the noise residual extraction method such that it is specifically tailored to the features of fingerprint sensors. Note this is different than attempting to tailor the approach to the signal of fingerprint images (locally approximated 2D sinusoid). The distinction is important because an effort should be made to characterize the scenery independent noise imperfections as opposed to characteristics which may define the foreground signal. It will also be interesting to see if a method can be devised to differentiate between images at other levels such as model, brand, or technology. Additionally, we are testing the technique on other biometric modalities such as iris. Since iris images are captured with devices operating in the infrared region of the electromagnetic spectrum, the optical sensors in iris capture devices are different than those found in optical fingerprint readers which will likely result in a different noise signature. Finally, a framework is being developed such that the technology could be integrated as a method of source validation in a digital chain of evidence.

8. Conclusion

This paper investigated the notion of sensor identification in biometric fingerprint devices. We established the prospects of performing identification based on estimating PNU noise inherent to image through wavelet based denoising as proposed by Lukas et al. in [2]. Beyond presenting the ability to perform sensor fingerprinting on biometric fingerprint devices at the unit level, we established the effect of varying the number of images used in arriving at reference patterns. Having looked at the minimum training requirements, we concluded that sensor identification can be performed with a great deal of accuracy (in two of the databases) even if one has access to only 1 image in which to establish a reference pattern. The application of a digital hardware fingerprinting technique such as the one tested in this work can be used both as a method of counteracting injection attacks at the time of check in biometric systems as well as allowing an individual to establish a "chain of evidence" which is often critical in systems such as assembling cases to prosecute criminal activity, establishing identity dominance in the battlefront, and discovering fraudulent activity in commercial systems.

References

[1] N. K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," *IBM Systems Journal*, vol. 40, no. 3, pp. 614–634, 2001.

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

[3] N. Khanna, A.K. Mikkilineni, A.F. Martone, G.N. Ali, G.T.C. Chiu, J.P. Allebach, and E.J. Delp, "A survey of forensic characterization methods for physical devices," *Proceedings of the 6th Annual Digital Forensic Research Workshop (DFRWS '06)*, vol. 3, pp. 17–28, September 2006.

[4] Kai San Choi, Edmund Y. Lam, and Kenneth K. Y. Wong, "Source camera identification using footprints from lens aberration," *Digital Photography II SPIE*, vol. 6069, no. 1, pp. 172–179, 2006.

[5] Z.J. Geradts, J. Bijhold, M. Kieft, K. Kurosawa, K. Kuroki, and N. Saitoh, "Methods for identification of images acquired with digital cameras," *Enabling Technologies for Law Enforcement and Security SPIE*, vol. 4232, Feb 2001.

[6] M.L. Kharrazi, H.T. Sencar, and N. Memon, "Blind source camera identification," *International Conference on Image Processing (ICIP) 2004.*, vol. 1, pp. 709–712 Vol. 1, 24-27 Oct. 2004.

[7] S. Bayram, H. Sencar, N. Memon, and I. Avcibas, "Source camera identification based on cfa interpolation," *IEEE International Conference on Image Processing (ICIP) 2005.*, vol. 3, pp. III–69–72, 11-14 Sept. 2005.

[8] A. Emir Dirik, Husrev T. Sencar, and Nasir Memon, "Source camera identification based on sensor dust characteristics," *IEEE Workshop on Signal Processing Applications for Public Security and Forensics (SAFE) 2007.*, pp. 1–6, 11-13 April 2007.